VirD2 of Agrobacterium tumefaciens : functional domains and biotechnological applications

Kregten, M. van

Chapter 4

Translocation of novel recombinant effector proteins from *Agrobacterium tumefaciens* to *Arabidopsis thaliana*

Maartje van Kregten, Paulo de Boer, Paul J.J. Hooykaas and Bert J. van der Zaal

Abstract

In this study, novel effector proteins for use in plant transformation were created by fusion of several biologically interesting protein domains to the relaxase VirD2 of *Agrobacterium tumefaciens*. All effector proteins, when equipped with the C-terminus of VirF, were translocated through the Type IV Secretion System of *Agrobacterium* to *Arabidopsis thaliana* roots, albeit with different efficiencies. When artificial DNA-binding zinc fingers were added, the translocation efficiency dropped from 63% of wild type VirD2 for one zinc finger to 1% for six zinc fingers. Addition of a *Fok*I endonuclease domain did not decrease translocation efficiency any further. Compared to a HO nuclease moiety, an I-SceI homing endonuclease moiety was transferred very efficiently. Moreover, expression of the VirD2-I-SceI-F fusion protein in *Arabidopsis* plants demonstrated that this protein still possessed site specific nuclease activity. Fusing proteins-of-interest to VirD2 proved to be a very sensitive way to determine if they can be translocated by simply monitoring T-DNA encoded *GUS* gene activity. Moreover, the fact that zinc finger nucleases as well as homing endonucleases can be translocated to recipient eukaryotic cells, employing a protein moiety which is also able to transfer a single stranded DNA molecule, offers intriguing possibilities for gene targeting purposes.

Introduction

Targeted integration of transgenes at a predetermined locus in a complex genome, also known as gene targeting (GT), would be the ideal way to create organisms with precisely the trait altered that is of interest for study or for biotechnological applications. Unfortunately, in higher eukaryotes and especially in plants, foreign DNA predominantly integrates at random positions in the genome. This may cause the transgene to be expressed at different levels, the expression levels of endogenous genes to be disturbed, and the presence of extra copies of the transgene - or parts thereof - at various positions in the recipient genome. Both for the study of fundamental processes of plant biology and the generation of genetically altered crops, a method for reliable gene targeting would be more than welcome.

The problem regarding gene targeting in plants lies in the pathway employed for

the integration of foreign DNA into the plant genome. In most cases, DNA integration happens via non-homologous end-joining (NHEJ), in which the incoming DNA is ligated to the plant genome in an error-prone way. However, for biotechnological applications, the pathway of homologous recombination (HR), in which DNA is integrated via recombination between incoming and target DNA, would be the pathway of choice since the locus of integration can then be predetermined by inserting homology with the desired integration site. Unfortunately, in plants NHEJ is much more prevalent than HR. The earliest attempts at GT in several plant species resulted in very low efficiencies, estimated at 10^{-6} to 10^{-4} (Paszkowski *et al.*, 1988, Offringa *et al.*, 1990, Halfter *et al.*, 1992, Hrouda and Paszkowski, 1994, Risseeuw *et al.*, 1995, Miao and Lam, 1995). Since then, it has been demonstrated in plants that the presence of a double-stranded break (DSB) in the genome near the integration site will enhance the possibility of integration via HR (Puchta *et al.*, 1996). Since this discovery, methods have been developed to induce DSBs in the genome to optimize the levels of HR, most of them using zinc finger nucleases but also homing endonucleases.

Zinc finger nucleases can be regarded as artificial restriction enzymes consisting of two protein domains: one consisting of artificial zinc finger (ZF) moieties and an endonuclease (N) domain. The ZF domain represents a DNA binding module, which can be made by linking individual Cis_2-His_2 type ZF moieties with known DNA-binding characteristics, thereby forming a polydactyl ZF (PZF) structure binding to a DNA sequence of choice. With a single ZF essentially interacting with three consecutive basepairs of DNA, several ZFs need to be combined to generate the specificity needed to target a specific site in a complex genome. The N domain of ZFNs is thus far represented by the non-specific nuclease domain of the restriction enzyme *Fok*I, which is active after dimerization (Bitinaite *et al*, 1998, Kim *et* al., 1996, reviewed in Carroll, 2008). The development of ZFN technology has already resulted in many successful gene targeting events in a variety of eukaryotic organisms (Bibikova *et al.*, 2002, Wright *et al.*, 2005, Doyon *et al.*, 2008, De Pater *et al.*, 2009, reviewed in Weinthal *et al.*, 2010).

Another method of creating DSBs in the genome is by using natural homing endonucleases (HEs) that have very long, and therefore highly specific, asymmetric recognition sites. As their name implies, HEs are involved in the process of intron homing. HEs are unusual in the sense that they are encoded by elements inside another open reading frame, either in introns or as inteins, intervening protein sequences that are spliced out post-translationally. Their function is solely to copy their coding sequence into other loci, making

HEs a selfish genetic element for which no benefit to the host has been found (Jurica and Stoddard, 1999, Stoddard, 2005). The advantage of HEs is that their recognition sites are often so long that they are mostly absent in a given genome. Insertion of a HE recognition site, such as for I-SceI (Puchta *et al.*, 1996), thus allows for controlled experiments using a single DSB in a genome. Although not yet as flexible as the ZFN system, recent reports that HEs with modified recognition sites can be generated will undoubtedly lead to the use of HEs as important tools in molecular genetics (reviewed in Pâques and Duchateau, 2007, Galetto *et al.*, 2009).

Obviously, for getting protein tools into cells, a delivery system will be required. By using the Type 4 Secretion System (T4SS) of the natural genetic engineer *Agrobacterium tumefaciens*, it should be possible to precisely deliver DNA as well as protein substrates to a wide variety of cells. The T4SS is nowadays considered to be a passageway between the *Agrobacterium* cell and the recipient cell through which both protein and single stranded DNA are translocated (reviewed in Christie, 2004, Ding *et al.*, 2003, Lawley *et al.*, 2003). In nature, *Agrobacterium* transfers part of its tumor-inducing plasmid – the T- or transferred strand, flanked by the right (RB) and left border (LB) repeats – to cells of susceptible dicotyledonous plant hosts, causing the crown gall tumor disease (reviewed in Zhu *et al.*, 2000, Pitzschke and Hirt, 2010). In a laboratory setting, the host range of *Agrobacterium* has been extended to, amongst others, monocotyledonous plants and fungi (resp. Hooykaas-van Slogteren *et al.*, 1984, Bundock *et al.*, 1995, De Groot *et al.*, 1992), thus making *Agrobacterium*-mediated transformation (AMT) a widely used tool in biotechnology.

For genetic transformation of target cells, the protein VirD2 is essential. It is a strand processing relaxase which participates in liberating the T-strand by nicking the RB and LB sequences. During transfer of the single stranded T-strand via the T4SS, the VirD2 protein remains covalently bound at the 5' end (Dürrenberger *et al.*, 1989, Scheiffele *et al.*, 1995, Vogel and Das, 1992).

Besides DNA-bound VirD2, several other virulence (Vir) proteins can pass through the T4SS. This has been shown for VirE2 (Citovsky *et al.*, 1992), VirF (Vergunst *et al.*, 2000), VirD5 (Vergunst *et al.*, 2005), and VirE3 (Schrammeier *et al.*, 2003). These proteins share a C-terminal stretch of positively charged amino acids, which are likely to function as a T4SS translocation signal (Vergunst *et al.*, 2005). In the seminal study regarding Vir protein translocation, the recombinase protein Cre – which is unrelated to any of the Vir proteins - was shown to be able to translocate to *Arabidopsis* root cells when fused to the C-terminus of VirF (Vergunst *et al.*, 2000). In the same study, the CRAfT (Cre Recombinase reporter

Assay for Translocation) assay was developed, allowing detection of Cre protein transfer in an indirect manner by monitoring Cre-mediated recombination events in a special *Arabidopsis* target line.

In order to develop a system for the translocation of proteins via the T4SS of *Agrobacterium*, the most straightforward solution would thus be to equip proteins of interest with a C-terminal sequence derived from the C-terminus of VirF. Applying this strategy enabled us recently to present evidence for the translocation of severely truncated versions of the relaxase VirD2. Of these truncations, we showed that they are still functional in T-strand processing, and lack only the translocation function necessary to translocate the VirD2-T-strand to the recipient cell. When the VirD2 truncations are fused to the C-terminus of VirF, translocation is restored, thus providing definitive proof that T-strand transfer is protein-driven (Van Kregten *et al*., 2009).

Our finding that the VirD2 protein can be manipulated rather extensively while maintaining the natural biological function of interest, prompted us to investigate whether a VirD2 framework might be a useful vehicle for accommodation of other protein domains while preserving all biological functions of interest. Detection of VirD2 translocation is easy; genes on the T-strand, which is covalently bound to VirD2, will make detection simple and does not require a special reporter line of the target organism. In that manner, starting from the natural VirD2 function, it would be possible to construct novel protein tools for molecular genetics. Concerning the goal of GT in higher eukaryotes, in particular plants, such protein tools should combine (1) an ability to transfer T-strand molecules to the large range of organisms amendable for AMT, with (2) additional functions that will facilitate site-specific integration of co-transferred DNA into the recipients genome.

As a first step towards VirD2-mediated genome engineering (VirGEN) we investigated whether functional VirD2 derivatives can be combined with biotechnologically interesting protein domains. We thus demonstrate that ZF domains as well as complete ZFNs can in principle be translocated, together with incoming T-strands, although the efficiency declines severely with more complicated PZF domains. Two HEs, I-SceI and HO can also be translocated in fusion to VirD2; in particular the I-SceI moiety was were translocated very efficiently, at the high level of 71% of wild type VirD2. The implications of these findings are discussed.

Results

Translocation of VirD2 fusion proteins

To create novel fusion proteins with multiple functions, both in T-strand production and translocation from *Agrobacterium* to the recipient cell, and as effector protein in the recipient cell, we made fusion proteins containing VirD2 or VirD2-204, the relaxase domain of VirD2, of which we have shown that it is proficient in T-strand translocation (Van Kregten *et al.*, 2009). To easily detect fusion protein expression, we added the FLAG tag N-terminally, and to ensure nuclear localization, the NLS of SV40 was added just downstream of the FLAG tag. To enable translocation, the C-terminal 37 amino acid residues of VirF constituting a T4SS translocation signal (Vergunst *et al.*, 2000) were added at the C-terminus. While addition of the short N-terminal FLAG tag and NLS sequence did not seriously hamper VirD2 activity (Van Kregten *et al.*, 2009), preliminary experiments indicated that much longer N-terminal protein sequences could be inhibitory, possibly for the reason that the catalytic Tyr residue at position 29 that nicks the DNA in T-strand processing is normally present at the N-terminus (Vogel and Das, 1992). Moreover, it can easily be envisaged that N-terminal fusion of protein domains to VirD2 might lead to biased conclusions regarding their effects upon the natural VirD2 activity; in case that the hybrid protein should fall apart, a C-terminal fragment comprised of a VirD2 domain directly followed by a C-terminal VirF sequence is likely to possess relaxase as well as translocation activity. With structural or enzymatic protein modules inserted in between the VirD2 proper and the F sequence, transfer of the complete proteins is close to guaranteed since the VirD2 module has then become dependent upon the C-terminus of VirF. VirD2 was thus fused to (P)ZF moieties comprised of 1 to 6 ZFs, to 1 to 3 (P)ZF moieties combined with the nuclease domain of *Fok*I, creating ZFN-type fusion domains, and to the HEs I-SceI and HO. Of the latter two, a fusion to just the relaxase domain of VirD2 (VirD2-204) was also made. A schematic drawing of the novel recombinant proteins can be found in Fig. 1. Expression of constructs in *Agrobacterium* was confirmed by Western blot, using an anti-FLAG antibody, except for VirD2-HO-F (Fig. 2).

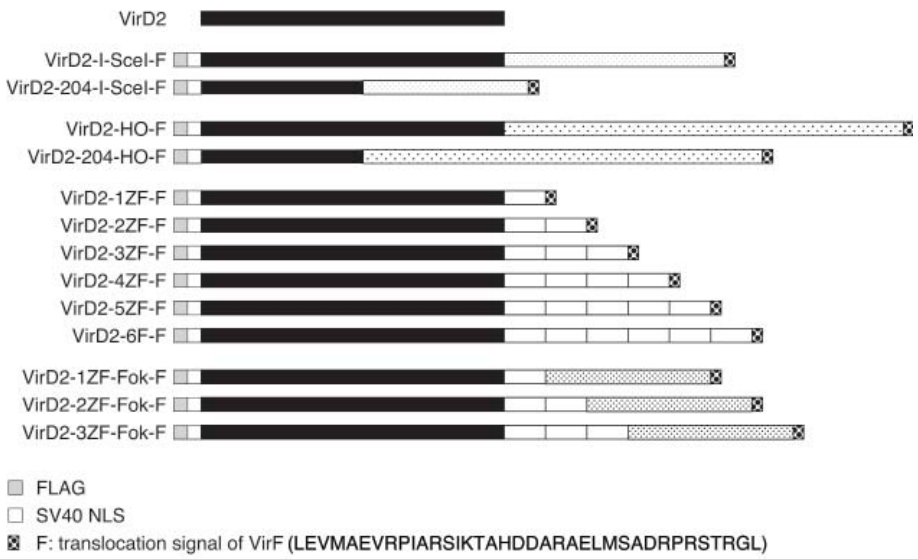

Figure 1: Schematic drawing of the novel recombinant effector proteins. Drawing is not to scale.

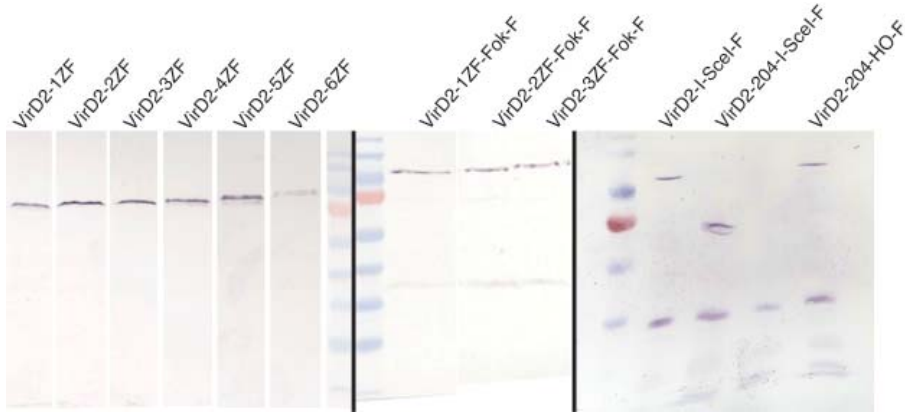

Figure 2: Western blot indicating expression of the novel recombinant effector proteins, expressed in *Agrobacterium* strain LBA2585 (Δ*virD2*, ΔT-DNA). The red band in the marker indicates 70 kDa (Fermentas).

To investigate the ability of the novel hybrid proteins to pass through the T4SS, we cocultivated *Agrobacterium* strain LBA2585 (Δ*virD2*, ΔT-DNA; Bravo-Angel *et al*., 1998) containing the expression vector for the relevant construct as well as the wide-host range vector pCAMBIA2301, the donor for a T-strand sequence with *in planta* expressed *GUS* and *NPTII* (kanamycin) marker genes, with *Arabidopsis* C-24 root explants. The number of GUS positive spots was determined after three days of cocultivation and taken as an

indication of transfer of T-strand molecules to the root cells, either transiently present or integrated into the genome, as described previously (Tinland *et al.*, 1995, Van Kregten *et al.*, 2009). All data were normalized to the level obtained with an LBA2585 strain harboring an expression plasmid providing wild-type (WT) VirD2, thus without any extra tags or protein domains. For all constructs, at least three independent experiments were performed.

As can be seen in Table 1, the efficiency of translocation differs widely between the different constructs. For VirD2 fusions with an increasing number of ZFs, a gradual decline in translocation efficiency was observed. While a construct consisting of VirD2 and 1 ZF still translocated at the relatively high level of 63% of wild type VirD2, addition of a second ZF and third ZF lowered the efficiency to 48% and 30%, respectively. Further increasing the number of ZFs about halved the transfer efficiency with every extra ZF added, until only 1% of transfer efficiency was observed when six ZFs were present.

To determine if the translocation of an entire ZFN could be a possibility, we created fusion proteins consisting of VirD2, 1 to 3 ZFs, the nuclease domain of *Fok*I, and the C-terminus of VirF. Remarkably, these constructs translocated at the same level as their counterparts that did not contain a *Fok*I nuclease domain (Table 1). Hence, the *Fok*I protein domain in itself does not at all hamper translocation of a VirD2 fusion via the T4SS.

Translocation of recombinant VirD2-ZFN-F proteins is thus possible, but for biotechnological applications the levels of translocation might be insufficient, especially for those proteins with 5 or 6 ZFs which should possess superior DNA-binding characteristics (Neuteboom *et al.*, 2006). Therefore, we set out to investigate translocation of the other group of DSB-inducing proteins that are widely used *in planta*, the homing endonucleases. We created fusion proteins of the monomeric HEs I-SceI, the use of which in plant research has already been documented (e.g. Puchta *et al.*, 1993, Puchta *et al.*, 1996, Orel *et al.*, 2003), and HO (Chiurazzi *et al.* 1996). Recombinant proteins consisting of VirD2, I-SceI and the C-terminus of VirF translocated at 71% of the level of wild type VirD2 (Table 1). A similar HO fusion was transferred to a lesser extent, only at 29% of wild type VirD2. To investigate whether the level of translocation could be further increased when the size of the fusion protein was diminished, we also made recombinant proteins using only the N-terminal relaxase half of VirD2, up to amino acid residue 204. In our previous study (Van Kregten *et al.*, 2009), VirD2-204 fused to the C-terminus of VirF was shown to be able to direct T-strands to recipient cells with only a modest loss in efficiency compared to the two times larger protein possessing the complete VirD2 domain. As can be seen in Table 1, we found that the translocation efficiencies of VirD-204-based fusions with I-SceI and HO

were lower than those of the similar constructs containing full length VirD2; for the I-SceI fusion, translocation dropped from 71% to 51% of the level of wild type VirD2. We thus did not detect a net benefit regarding protein and T-strand translocation by reducing the bulkiness of the fusion proteins by deleting the dispensable C-terminal half of VirD2.

To determine whether stable integration of the T-strand into the genome of *Arabidopsis* cells was still possible using novel recombinant proteins, stable integration was assessed by determining the number of transgenic kanamycin-resistant calli developing on root segments cocultivated with *Agrobacterium* strains expressing VirD2-I-SceI-F, VirD2-204-I-SceI-F, and VirD2-HO-F (Table 2). When comparing the stable integration results with the previously obtained data regarding the amount of GUS-positive spots after three days of cocultivation, no significant differences were found (see Table 1 and 2). Hence, most supposedly transient events finally led to stable presence of one or more T-DNAs within the recipient genome, demonstrating that the addition of a large protein domain to VirD2 does not hamper integration of the T-DNA.

All the biologically interesting proteins tested in this study could be made translocation-competent by the addition of VirD2 and the C-terminus of VirF. However, thus far the translocation of in particular the I-SceI domain was most compatible with the T4SS of *Agrobacterium*. For further proof of principle that novel recombinant effect proteins can be designed which combine DNA translocating potential with DSB-inducing capacity, it is essential to demonstrate that the fusion proteins retain their nuclease activity in the recipient cells. Therefore, we tested the *in planta* activity of the constructs containing I-SceI and HO.

Activity of the nuclease moiety

The activity of the VirD2 moiety of the fusion protein is warranted by its translocation and the subsequent detection by means of expression of the genes on the T-strand; if the VirD2 domain would be non-functional, this could not occur. For obtaining evidence that fusion of enzyme domains within a VirD2 framework does not affect the enzymatic properties of interest, in this case nuclease activity on chromatin-embedded DNA *in planta*, further tests are required. The activity of the specific nuclease moieties in the set of constructs mentioned above can be assessed in *Arabidopsis* lines that have a specific target site inserted into their genome. Expression of an active, site specific nuclease in such plant lines should induce a DSB in the DNA at its cognate recognition site. To a certain extent, erroneous repair will

Table 1: Transient transformation by novel recombinant effector proteins

construct	transient transformation[a]
wild type VirD2	1
FLAG-NLS-VirD2-1ZF-F	0.63 +/- 0.04
FLAG-NLS-VirD2-2ZF-F	0.48 +/- 0.09
FLAG-NLS-VirD2-3ZF-F	0.30 +/- 0.08
FLAG-NLS-VirD2-4ZF-F	0.12 +/- 0.04
FLAG-NLS-VirD2-5ZF-F	0.06 +- 0.03
FLAG-NLS-VirD2-6ZF-F	0.01 +/- 0.01
FLAG-NLS-VirD2-1ZF-Fok-F	0.59 +/- 0.07
FLAG-NLS-VirD2-2ZF-Fok-F	0.48 +/- 0.05
FLAG-NLS-VirD2-3ZF-Fok-F	0.34 +/- 0.10
FLAG-NLS-VirD2-I-SceI-F	0.71 +/- 0.09
FLAG-NLS-VirD2-204-I-SceI-F	0.51 +/- 0.11
FLAG-NLS-VirD2-HO-F	0.29 +/- 0.13
FLAG-NLS-VirD2-204-HO-F	0.17 +/- 0.08

[a] n=3 or greater, at least 100 explants were scored for GUS-positive spots per experiment.

Table 2: Stable integration of T-DNA by novel recombinant effector proteins

construct	stable transformation[a]
wild type VirD2	1
VirD2-I-SceI-F	0.63 +/- 0.07
VirD2-204-I-SceI-F	0.52 +/- 0.14
VirD2-HO-F	0.28 +/- 0.14

[a]n=3 or greater, at least 100 explants were scored for formation of kanamycin-resistant calli per experiment.

lead to mutations at this restriction site, rendering it insensitive to further cleavage by the nuclease moiety. As mentioned above, we focused on I-SceI fusion proteins as their transfer through the T4SS was very efficient. Furthermore, detection of erroneously repaired I-SceI target sites is very much facilitated by commercially available I-SceI enzyme; the use of HO is less attractive in this respect as such an enzyme source is lacking. For ZFN target sites, any commercially available enzyme can in principle be used for digestion of a cognate 6 bp target site chosen to be present in between the recognition half sites for the ZFN (Lloyd *et al.*, 2005). In this study, however, we refrained from further testing ZFNs for a combination of two reasons. Firstly, as demonstrated above, ZFNs with complex PFZ

domains consisting of five or six ZFs apparently are much less suitable substrates for T4SS transfer. Secondly, although three fingered PZF domains can be transferred via VirD2, the ZFNs based upon such simple PZF domains that were produced in our lab did not lead to site directed mutations in a chromosomal yeast reporter system (De Pater *et al*, 2009).

To assess the activity of the nuclease moiety of the VirD2-I-SceI-F, VirD2-204-I-SceI-F, VirD2-HO-F or VirD2-204-HO-F proteins upon chromatin-embedded DNA in plant cells, we produced the relevant VirD2 fusion protein under control of the *RPS5a* promoter, which is primarily active in dividing cells, in plant lines that contained a single copy of a locus with a recognition site for either I-SceI or HO. A similar approach was followed previously for complex ZFNs (De Pater et al., 2009).

At least 10 primary transformed plants were grown and their seeds were sown on selective medium. From the second generation T2 seedlings, DNA was extracted and the part of the locus containing the I-SceI or HO site was PCR-amplified. To screen for mutated recognition sites, the PCR product was digested with I-SceI (New England Biolabs), or *Hind*III (Fermentas). *Hind*III is the restriction enzyme cutting closest to the HO recognition site, since HO is not commercially available (see Fig. 3). In PCR products derived from plant lines expressing VirD2-I-SceI-F and VirD2-204-I-SceI-F, enzyme-resistant DNA was readily found and sequenced (Fig 3). The slower migrating DNA bands, each corresponding to a different transgenic plant line expressing VirD2-I-SceI-F or VirD2-204-I-SceI-F, in which the I-SceI recognition site is destroyed, vary in intensity (Fig. 3 and data not shown). This indicates that the fraction of damaged I-SceI recognition sites within DNA samples derived from different plant lines is variable, possibly due to different expression levels of the proteins. Since VirD2-I-SceI-F and VirD2-204-I-SceI-F are expressed under the control of the *RPS5a* promoter, the protein is mainly expressed in dividing cells. Initial damage occurring in dividing cells will thus be inherited in clonal sections, the size of each section depending on the number of daughter cells generated after the mutagenic event. On average, under the given circumstances, about one third of the total PCR-product displayed I-SceI-resistance, indicating that about one third of all available sites in T2 seedlings have been erroneously repaired. Of course, it is impossible to determine the amount of breaks repaired correctly, but it may very well be that only a minority of breaks will lead to a detectable footprint.

Sequence analysis damage induced by VirD2-I-SceI-F or VirD2-204-I-SceI-F showed that most events had led to small deletions or insertions (Fig. 3). Of course, very large deletions or insertions will be missed by PCR screening; very large deletions will

remove the binding site of a primer, and very large insertions will cause the formation of products that will not be amplified in the PCR reaction.

A similar strategy for finding evidence of mutations induced by expression of VirD2-HO-F and VirD2-204-HO-F was unsuccessful. This might indicate that HO is inactive when fused to VirD2 or, alternatively, it might reflect a technical difficulty when repair of HO-induced breaks in dividing plant cells would produce mainly small (6 bp or less) deletions that do not affect the neighboring *Hind*III-site (Fig. 3). Nevertheless, the fact that VirD2-(204)-I-SceI-F fusion proteins which can process and transfer T-DNA to a recipient plant cell do also posses nuclease activity when expressed *in planta* strongly indicates novel VirD2 based protein tools are feasible.

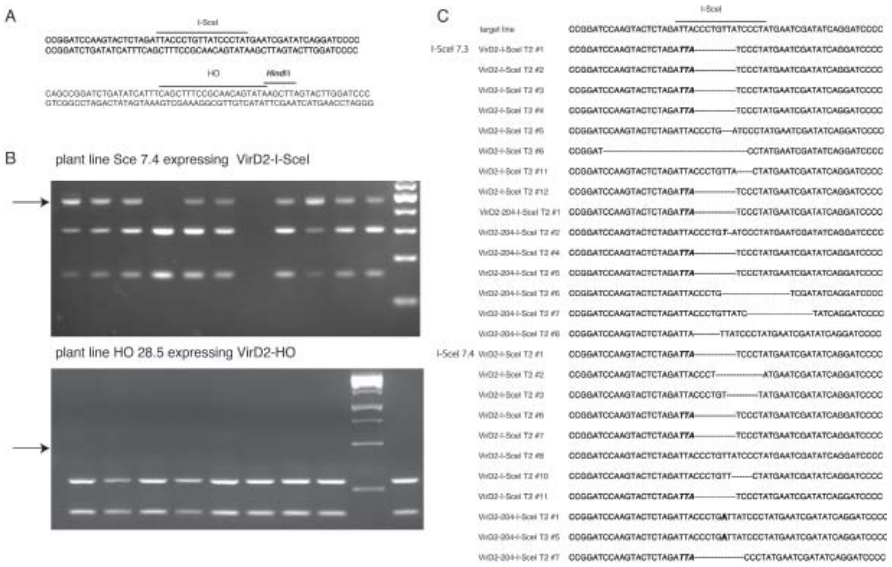

Figure 3: evidence for nuclease activity of VirD2-fusion proteins, **A**: recognition sites of I-SceI and HO, **B**: examples of digested PCR product, with the position of resistant bands indicated by an arrow, **C**: sequences of resistant bands. **Figure 3**: evidence for nuclease activity of VirD2-fusion proteins, **A**: recognition sites of I-SceI and HO, **B**: examples of digested PCR product, with the position of resistant bands indicated by an arrow, **C**: sequences of resistant bands.

Discussion

In this study, we report that the creation of a novel class of recombinant effector proteins for use in *Agrobacterium*-mediated transformation is feasible. In their most desirable form, these proteins posses T-strand processing activity in *Agrobacterium*, can be translocated to a plant cell recipient together with the T-strand, which can contain DNA sequences of interest for delivery to the recipient cells, and, when present in the plant cell, they can induce the formation of a site specific DSB in the recipients genome

To create these proteins, we made fusions of three protein domains: 1) the *Agrobacterium* relaxase protein VirD2, which is responsible for T-strand processing within *Agrobacterium* cells, 2) a biologically interesting effector moiety which can exert a function in a recipient cell, and 3) the T4SS translocation signal of the C-terminus of the *Agrobacterium* VirF protein (see Fig. 1 for a schematic drawing of the proteins). The effector domains used in the present work share a common feature: they are in principle able to induce a site-specific DSB in a complex genome, an event that can greatly GT (reviewed in Pâques and Duchateau, 2007, Carroll, 2008). Since the main tools for the induction of DSBs are ZFNs and HEs, we have included members of both classes in our study.

Regarding the possibility to include a complete ZFN moiety in between the N-terminal VirD2 domain and the C-terminal T4SS translocation signal, the situation is rather complex. Our data demonstrate that the *Fok*I nuclease domain passes the T4SS without any difficulty and, in addition to that, its fusion to VirD2 also leaves the essential VirD2 functions regarding T-strand processing and transfer unaffected. However, complete ZFN transfer seems to be hampered by the required ZF domains; while addition of one ZF still allows for about 68% of WT VirD2 translocation, the addition of more ZFs results in a steady decrease, with a PZF domain consisting of six ZFs being transferred at only 1% of WT VirD2 translocation (see Table 1). The experiments described in the present study do by themselves not yet provide definitive proof that the actual passage through the T4SS is the limiting factor for transfer of complex PZF domains since an inhibitory action of these domains upon the VirD2 functions associated with T-strand processing cannot be excluded. However, using the CRAfT assay, previous results indicated that detectable Cre activity in recipient plant cells was already severely diminished when only a single artificial ZF moiety was fused in between Cre and the C-terminus of VirF. Addition of more complex PZF

domains, consisting of three or more ZFs, nearly nullified detectable Cre activity in target cells (Lindhout and Van der Zaal, unpublished data). Since these Cre fusion proteins had nothing to do with T-strand processing, it thus appears that Cis_2-His_2 type ZF moieties indeed hamper translocation of proteins via the T4SS of *Agrobacterium*. Nevertheless, in contrast to Cre fusions, a VirD2 framework still allows for considerable transfer of three fingered PZF domains at about 30% of WT VirD2 levels, with or without a *Fok*I nuclease domain (Table 1). This might still hold promise for VirD2-mediated ZFN transfer.

In the present study, we did not pursue further investigation of ZFN transfer since three fingered ZFNs currently present in our lab do not suffice to generate footprints in yeast cells (De Pater *et al.*, 2009, Van der Zaal *et al.*, unpublished data). Obviously, it will be very interesting to explore the biotechnological applications of VirD2-ZFN fusions with more potent three-fingered ZFNs, possibly produced using the OPEN system (Maeder *et al.*, 2008), instead of modular assembly as was done in our lab. However, it should be realized that the commonly used ZFNs function as a dimer and that naturally occurring recognition sites are very rarely palindromic, requiring a different PZF domain for each half site. Hence, when the proteins constituting a ZFN are administered in small quantities, as it is most likely the case when using T4SS transfer during AMT, the correspondingly reduced chance of two ZFN subunits encountering each other is likely to diminish ZFN activity. Creating a ZFN that functions as a monomer might therefore be highly advantageous. The recent description of a monomeric ZFN in which the *Fok*I domain already consists of a covalently bound dimer which is `sandwiched` between two PZF domains (Mori *et al.*, 2009) is therefore very promising, although an extra PZF domain for sequence specification might very well again complicate T4SS-mediated transfer.

While the overall picture regarding the option that VirD2-ZFN-F fusions can be used for VirD2-mediated genome engineering (VirGEN) is not yet clear, the possibilities for VirD2-HE-F fusions are very promising. The homing endonuclease HO, in fusion to VirD2 and the C-terminus of VirF, translocates at 29% of the efficiency of wild type VirD2. Furthermore, the HO moiety is not inhibitory to integration of the T-strand; stable integration is approximately equal to transient transformation (see Tables 1 and 2). Translocation of HO is remarkable, considering the fact that HO contains 5 putative zinc fingers at its C-terminus (Russell *et al.*, 1986). Apparently, these ZFs are in a conformation that is less inhibitory to translocation than the ZFs stitched together in a PZF domain of similar size. Hence, simply attributing failure of translocation to increased bulkiness of fusions proteins provides insufficient reason for translocation difficulties with complex

PZFs. Individual assessment of the translocation ability of proteins remains is required until the general requirements for protein translocation are more clear. A great step toward possibly defining these requirements has been taken with the elucidation of the structure of the inner and outer membrane complex of the T4SS of plasmid pKM101 (Fronzes *et al.*, 2009, Chandran *et al.*, 2009).

Of all novel protein tools tested, VirD2-I-SceI-F translocates at the highest level of 70% of WT VirD2. Reducing the bulkiness of the total fusion protein, by using only the N-terminal half of the VirD2 protein which constitutes the active relaxase moiety, did not present any advantage over fusion to full length VirD2. In fact, the translocation rate of VirD2-204-I-SceI-F dropped to 51% of wild type. This might indicate that, although not required for relaxase activity, part of the deleted C-terminal half of VirD2 can positively contribute to transfer of fusion proteins, as was suggested previously (Van Kregten *et al.*, 2009). Also for HO fusions, the use of the VirD2-204 domain was not advantageous for transfer efficiency (see Table 1). As can be seen in Table 2, using VirD2 or VirD2-204 as a framework for the fusion protein does not make much difference in terms of integration efficiency. Just as in earlier work (Tinland *et al.*, 1995, Bravo-Angel *et al.*, 1998), an active role for VirD2 in integration of the T-DNA remain unlikely.

Whatever the future may hold for VirD2-mediated transfer of ZFN domains, recombinant effector proteins that involve HEs are highly promising to develop as DSB-inducing agents for GT in plants, provided that proof of nuclease activity of these fusion proteins can be obtained. As mentioned above, we did not succeed in finding footprints indicative of DSB-formation after expressing HO-containing fusion proteins *in planta*. This might reflect technical difficulties, as mutations other than larger than 6 bp deletions or the gain of large insertions in a HO site remained unnoticed with the procedures used. Otherwise, HO might be a difficult enzyme for use in plants. To our knowledge, HO has been used as DSB-creating agent in plants once and then under the control of the strong, constitutive *35S* promoter (Chiurazzi *et al.*, 1996). Since we expressed VirD2-(204)-HO-F under the control of the relatively weak *RPS5a* promoter, protein levels might have remained too low. Furthermore, it has been shown that HO is a rather unstable protein (Kaplun *et al.*, 2000), another possible complication for successful application in our experimental set-up. In contrast to HO fusion proteins, footprints confirming the activity of VirD2-(204)-I-SceI-F were readily found, mostly as small deletions or insertions (Fig. 3). This is consistent with findings in experiments performed in mammalian cells (Rouet *et al.*, 1994, Liang *et al.*, 1998).

Altogether, we have shown that several novel biologically interesting proteins that can in principle induce site specific DSBs can be translocated through the T4SS of *Agrobacterium* by fusion to VirD2 and F, although to a different extent. For I-SceI-containing fusion proteins, the T-strand processing and translocation ability of VirD2 remained largely intact and the nuclease activity was detectable when fusions proteins were expressed in plants. With this proof of principle, VirGEN could be realized, holding great promise for applications in GT. Although naturally occurring HE type nucleases like I-SceI have a fixed recognition site, different groups are working on strategies to accommodate the required flexibility within the HE scaffold in order to be able to generate proteins that can recognize chosen sequences of interest (Chen *et al.*, 2009, reviewed in Pâques and Duchateau, 2007, Galetto *et al.*, 2009). When the initial successes result in a generic strategy for selecting-site specific DNA-binding domains, VirGEN as we envisage it will easily come of age. By linking the T-strand and the protein of interest, it should become possible to deliver a T-strand to a specific site in a complex genome which is being cleaved in the same time frame, thus increasing the chances of GT and HR when the T-strand has homology to the locus of interest.

Materials and methods

Cloning of expression vectors for VirD2 fusion proteins

Cloning was performed using standard techniques in *Escherichia coli* strain DH5α. Both *E. coli* and *Agrobacterium* were cultured in LC medium containing the appropriate antibiotics. Primer sequences are listed in Table 3.

I-SceI was amplified from pUC18-I-SceI, a kind gift of Dr. H. Puchta, University of Karlsruhe, Germany, using primers I-SceI fw and I-SceI rv. The product was digested with *Spe*I and *Bgl*II and ligated to similarly digested pSKN-VirD2 or pSKN VirD2-204 (described in Van Kregten *et* al., 2009) to yield respectively pSKN-VirD2-I-SceI, or pSKN-VirD2-204-I-SceI.

HO was amplified from *Saccharomyces cerevisiae* using primers HO start and HO end. The product was digested with *Eco*RI and *Spe*I and ligated to similarly digested pSKN-VirD2 or pSKN-VirD2-204 to yield respectively pSKN-VirD2-HO and pSKN-VirD2-204-HO. The plasmids pSKN-VirD2 and pSKN-VirD2-204 are high copy *E. coli* cloning

vectors, derived from pSKN, in which the gene of interest becomes fused to the SV40 NLS sequence (for details, see Van Kregten *et al*, 2009).

The fragment containing the gene of interest of all pSKN-based vectors was released by digestion with *Not*I and ligated to *Not*I-digested *Agrobacterium* expression vector pBFF. As described previously, the pBFF plasmid allows for *virF* promoter-driven expression of N-terminal FLAG-tagged and SV40 NLS-containing fusion proteins, equipped with the 37 C-terminal amino acid translocation signal of the VirF protein (Van Kregten *et al.*, 2009). A schematic drawing of the proteins expressed by these constructs can be found in Fig. 1.

For expression of the same proteins in *Arabidopsis* plants under control of the *RPS5a* promoter, primarily active in embryos and meristems, *Not*I fragments were cloned into the previously described pGPTV-derived wide host range T-DNA vector pRF (Lindhout *et al.*, 2006, De Pater *et al.*, 2009) to create pRF-VirD2-I-SceI, pRF-VirD2-204-I-SceI, pRF-VirD2-HO and pRF-VirD2-204-HO. A schematic drawing of these proteins can be found in Fig. 1.

The zinc finger domains used were obtained from constructs in pSKN-SgrAI (Neuteboom *et al.*, 2006) by digestion with *Sfi*I and ligated into similarly digested pSKN-VirD2. The fragment containing the gene of interest was released by digestion with *Not*I and ligated to *Not*I-digested *Agrobacterium* expression vector pBFF.

The nuclease domain of *Fok*I was amplified from pSDM3835 (De Pater *et al.*, 2009) using primers Fok fw and Fok rv. The PCR product was cloned into pJET1.2 (Fermentas), excised by digestion with *Spe*I and ligated to relevant similarly digested pBFF VirD2-ZF vectors. A schematic drawing of the proteins expressed by these constructs can be found in Fig. 1.

Cloning of T-DNA vectors

To create binary vectors with an HO or I-SceI restriction site on T-DNA, pSDM3832 (De Pater *et al.*, 2009) was digested with *Bam*HI. Double-stranded oligo's were created from Scesite fw and Scesite rv or HOsite fw and HOsite rv, and ligated to *Bam*HI-digested pSDM3832, creating pSDM3832-I-SceI and pSDM3832-HO.

Root transformation

Root transformation experiments were performed as described earlier (Vergunst et al.,

2000). In short, seedlings of *Arabidopsis thaliana* ecotype C-24 were grown for 10 days in liquid medium, after which roots were harvested and pre-incubated for 3 days on medium containing 2,4-dichlorophenoxyacetic acid at 0.5 mg/liter and kinetin at 0.05 mg/liter. Subsequently, root segments were cocultivated for two days on the same medium but supplemented with 20 μM acetosyringone (Sigma-Aldrich) with *Agrobacterium tumefaciens* strain LBA2585, which has a deletion for *virD2* and the T-DNA (Bravo-Angel *et al.*, 1998), supplied *in trans* with the appropriate pBFF-derived VirD2 fusion protein expression vector and pCAMBIA2301 (www.cambia.org) as T-DNA donor. After this period, root segments were stained in staining solution containing X-gluc. For at least 100 root segments per construct per experiment, the amount of spots were counted. Data were normalized against data for T-DNA transfer by WT VirD2, as described previously (Van Kregten *et al.*, 2009). For the determination of stable integration, root segments were placed on medium containing N6-(isoprofenyl)adenine at 5 mg/liter and indole 3-acetic acid at 0.15 mg/liter. For selection of transformed plant cells, kanamycin was present at 50 mg/liter. Timentin (100 mg/liter) was used to kill remaining *Agrobacterium* cells. After two to three weeks, kanamycin-resistant calli were counted. Data were normalized against data for kanamycin-resistant callus formation by wild type VirD2, as described before (Van Kregten *et al.*, 2009).

Footprinting

Arabidopsis thaliana ecotype Col-0 lines containing an HO or I-SceI recognition site were created using floral dip (Clough and Bent, 1998) with *Agrobacterium* strain AglI (Lazo *et al.*, 1991) harboring pSDM3832-HO or pSDM3832-I-SceI as T-DNA donor. Homozygous phosphinotricin-resistant lines containing a single copy of the site of interest, namely Sce 7.3 and 7.4, and HO 28.5, were selected and were subsequently transformed using AglI containing pRF-VirD2-I-SceI, pRF-VirD2-204-I-SceI, pRF-VirD2-HO or pRF VirD2-204-HO. For every construct, at least 10 kanamycin-resistant transformants were allowed to grow and set seed. From their seedlings, DNA was extracted and footprinting was determined. The locus containing the restriction site was amplified with primers SP250 and SP251, as reported previously (De Pater *et al.*, 2009). For loci containing an I-SceI site, the PCR product was digested with I-SceI (NEB). For loci containing an HO site, the PCR product was digested with *Hind*III (Fermentas), because the enzyme HO is not commercially available. Enzyme-resistant bands were isolated from gel, cloned into pJET1.2 (Fermentas) and sequenced.

Western blot

Agrobacterium cultures were induced with 200 μM acetosyringone (Sigma-Aldrich) for at least 4 h before being harvested. A crude protein extract was made by boiling *Agrobacterium* cells in Laemmli buffer with β-mercapto-ethanol for 10 min. Insoluble proteins and debris were spun down and the supernatant was used for Western blotting. Proteins were separated on 10% sodium dodecyl sulfate polyacrylamide gel electrophoresis gels, transferred to nitrocellulose membrane (Immobilon-P, Millipore) and treated with FLAG antibody conjugated to alkaline phosphatase (cat no. A9469; Sigma-Aldrich). Detection of alkaline phosphatase activity was performed using nitroblue tetrazolium–5-bromo-4-chloro-3-indolyl phosphate solution (Sigma-Aldrich) according to instructions from the manufacturer.

Table 3: primer sequences

I-SceI fw:	GGATACTAGTGATGAAAAACATCAAAAAAAACCAGGTAATG
I-SceI rv:	TGATAGAGGAGGCTTTGAAAGGACTTGTAGATAGG
HO start:	GGATACTAGTTATGCTTTCTGAAAACACGACTATTCTGAT
HO end:	GGATCGAATTCTTGCAGATGCGCGCACCT
Fok fw:	TTCAACTAGTCAAAAGTGAACTGG
Fok rv:	AAACTAGTTGAAAGTTTATCTCGCCGTTATT
Scesite fw:	GATCTGATATCGAATTCATAGGGATAACAGGGTAATCTAGAGTACTTG
Scesite rv:	GATCCAAGTACTCTAGATTACCCTGTTATCCCTATGAATTCGATATCA
HOsite fw:	GATCTGATATCATTTCAGCTTTCCGCAACAGTATAAGCTTAGTACTTG
HOsite rv:	GATCCAAGTACTAAGCTTATACTGTTGCGGAAAGCTGAAATGATATCA
SP250:	CTCTGCCGTCTCTCTATTCG
SP251:	CTTGAAGAAGTCGTGCTGCTT

References

Bravo-Angel, A.M., Hohn, B., and Tinland, B. 1998. The omega sequence of VirD2 is important but not essential for efficient transfer of T-DNA by *Agrobacterium tumefaciens*. Mol. Plant-Microbe

Interact. 11:57-63.
Bibikova, H., Golic, M., Golic, K.G., and Carroll, D. 2002. Targeted chromosomal cleavage and mutagenesis in *Drosophila* using zinc-finger nucleases. Genetics. 161(3):1169-1175.
Bitinaite, J., Wah, D. A., Aggarwal, A. K., and Schildkraut, I. 1998. *Fok*I dimerization is required for DNA cleavage. Proc. Natl. Acad. Sci. 95(18):10570-10575.
Bundock, P., den Dulk-Ras, A., Beijersbergen, A., and Hooykaas, P. J. J. 1995. Trans-kingdom T-DNA transfer form *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO (Eur. Mol. Biol. Organ.) J. 14:3206-3214.
Carroll, D. 2008. Zinc finger nucleases as gene therapy agents. Gene Ther. 15(22):1463-1468.
Chen, Z., Wen, F., Sun, N., and Zhao, H. 2009. Directed evolution of homing endonuclease I-SceI with altered sequence specificity. Protein Eng. Des. Sel. 22(4):249-256.
Chandran, V. Fronzes, R., Duquerroy, S., Cronin, N., Navaza, N., and Waksman, G. 2009. Structure of the outer membrane of a type IV secretion system. Nature. 462:24-31.
Chiurazzi, M, Ray, A., Viret, J. F., Perera, R., Wang, X. H., Lloyd, and A. M., and Signer, E. R. 1996. Enhancement of somatic intrachromosomal homologous recombination in Arabidopsis by the HO endonuclease. Plant Cell. 8(11):2057-2066.
Christie, P.J. 2004. Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. Biochem. Biophys. Acta. 1693:219-234.
Citovsky, V., Zupan, J., Warnick, D., and Zambryski, P. 1992. Nuclear localization of *Agrobacterium tumefaciens* protein in plant cells. Science. 256:1802-1805.
Clough, S.J. and Bent, A.F. 1998. Floral dip: a simplified method for *Agrobacterium tumefaciens*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16(6):735-743.
De Groot, M. J., Bundock, P., Hooykaas, P. J., and Beijersbergen, A. G. 1998. *Agrobacterium*-mediated transformation of filamentous fungi. Nat. Biotechnol. 16:839-842.
De Pater, S., Neuteboom, L.W., Pinas, J.E., Hooykaas, P.J.J., and van der Zaal, B.J. 2009. ZFN-induced mutagenesis and gene targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. Plant Biotech. J. 7:821-835.
Ding, Z., Atmakuri, K., and Christie, P.J. 2003. The ins and outs of bacterial type IV secretion substrates. Trends Microbiol. 11(11):527-535.
Doyon, Y., McCammon, J.M., Miller, J.C., Faraji, F., Ngo, C., Katibah, G.E., Amora, R., Hocking, T.D., Zhang, L., Rebar, E.J., Gregory, P.D., Urnov, F.D., and Amacher, S.L. 2008. Heritable targeted disruption in zebrafish using designed zinc-finger nucleases. Nat. Biotechnol. 26(6) :702-708.
Dürrenberger, F, Crameri, A., Hohn, B., and Koukolikova-Nicola., Z. 1989. Covalently bound VirD2 from *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. Proc. Natl. Acad. Sci. 86(23):9154-9158.
Fronzes, R., Schäfer, E., Wang, L., Saibil, H.R., Orlova, E.V., and Waksman, G. 2009. Structure of a Type IV Secretion System Core Complex. Science. 323:266-268.
Galetto, R., Duchateau, P., and Pâques, F. 2009. Targeted approaches for gene therapy and the

emergence of engineered meganucleases. Expert Opin. Biol. Ther. 9(10):1289-1303.
Halfter, U., Morris, P.C., and Willmitzer, L. 1992. Gene targeting in *Arabidopsis thaliana*. Mol. Gen. Genet. 231:186-193.
Hrouda, M., and Paszkowski, J., 1994. High fidelity extrachromosomal recombination and gene targeting in plants. Mol. Gen. Genet. 243:106-111.
Jurica, M.S., and Stoddard, B.L. Homing endonucleases: structure, function and evolution. 1999. CMLS Cell. Mol. Life Sci. 1304-1326.
Hooykaas-van Slogteren, G. M., Hooykaas, P. J. J., and Schilperoort, R. A. 1992. Expression of Ti plasmid genes in monocotyledonous plants infected with *Agrobacterium tumefaciens*. Biotechnology 24:382-383.
Kaplun, L, Ivantsiv, I., Kornitzer, D. and Raveh, D. 2000. Functions of the DNA damage response pathway target Ho endonuclease of yeast for degradation via the ubiquitin-26S proteosome system. Proc. Natl. Acad. Sci. USA 97:10077-10082.
Kim, Y.G., Cha, J., and Chandrasagaran, S. 1996. Hybrid restriction enzymes: zinc finger fusions to *Fok*I cleavage domain. Proc. Natl. Acad. Sci. 93:1156-1160.
Lawley, T.D., Klimke, W.A., Gubbins, M.J. and Frost, L.S. 2003. F factor conjugation is a true Type IV secretion system. FEMS Microbiol. Lett. 224(1):1-15.
Lazo, G.R., Stein, P.A., and Ludwig, R.A. 1991. A DNA transformation-competent *Arabidopsis* genomic library in *Agrobacterium*. Biotechnology. 9:963-967.
Liang, F., Han, M., Romanienko, P.J., and Jasin, M. 1998. Homology-directed repair is a major double-strand repair pathway in mammalian cells. Proc. Natl. Acad. Sci. 95(9):5172-5177.
Lloyd, A., Plaisier, C.L., Carroll, D., and Drews, G.N. 2005. Targeted mutagenesis using zinc finger-nucleases in *Arabidopsis*. Proc. Natl. Acad. Sci. 102(6):2232-2237.
Lindhout, B.I., Pinas, J.E., Hooykaas, P.J.J., and van der Zaal, B.J. 2006. Employing libraries of artificial zinc finger transcription factors to screen for homologous recombination mutants in *Arabidopsis*. Plant J. 48(3):475-483.
Maeder, M.L., Thibodeau-Beganny, S., Osiak, A., Wright, D.A., Anthony, R.M., Eichtinger, M., Jiang, T., Foley, J.E., Winfrey, R.J., Townsend, J.A., Unger-Wallace, E., Sander, J.D., Müller-Lerch, F., Fu, F., Pearlberg, J., Go¨ bel, C., Dassie, J.P., Pruett-Miller, S.M., Porteus, M.H., Sgroi, D.C., Lafrate, A.J., Dobbs, D., McCray Jr, P.B., Cathomen, T., Voytas, D.F. and Joung, J.K. 2008. Rapid "open-source" engineering of customized zincfinger nucleases for highly efficient gene modification. Mol. Cell. 31:294–301
Miao, Z.H., and Lam, E. 1995. Targeted disruption of the TGA3 locus in *Arabidopsis thaliana*. Plant J. 7:359-365.
Mori, T., Kagatsume, I., Shinomiya, K., Aoyama, Y., and Sera, T. 2009. Sandwiched zinc-finger nucleases harboring a single-chain *Fok*I dimer as a DNA-cleavage domain. Biochem. Biophys. Res. Commun. 390:694-697.
Neuteboom, L.W., Lindhout, B.I., Saman, I.L., Hooykaas, P.J.J., and van der Zaal, B.J. 2006. Effects

of different zinc finger transcription factors on genomic targets. Biochem. Piophys. Res. Commun. 339:263-270.

Offringa, R., de Groot, M.J., Haagsman, H.J., Does, M.P., van den Elzen, P.J., and Hooykaas, P.J.J. 1990. Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium*-mediated transformation. EMBO J. 9:3077-3084.

Orel, N, Kyryk, A., and Puchta, H. 2003. Different pathways of homologous recombination are used for the repair of double-strand breaks within tandemly arranged sequences in the plant genome. Plant J. 35(5):604-612.

Paszkowski, J., Baur, M., Bogucki, A., and Potrykus, I. 1988. Gene targeting in plants. EMBO J. 7:4021-4026.

Pâques, F. and Duchateau, P. Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. 2007. Curr. Gene Ther. 7:49-66.

Pitzschke, A, and Hirt, H. 2010. New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. EMBO J. 29(6):1021-1032.

Puchta, H., Dujon, B., and Hohn, B. 1993. 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. 21(22):5034-5040.

Puchta, H., Dujon, B., and Hohn, B. 1996. Two different but relate mechanisms are used in plants for the repair of genomic double-strand breaks by homologous recombination. Proc. Natl. Acad. Sci. 93:5505-5060.

Risseeuw, E., Offringa, R., Franke-van Dijk, M.E.I., and Hooykaas, P.J.J. 1995. Targeted recombination in plants using *Agrobacterium* coincides with additional rearrangements at the target locus. Plant J. 7:109-119.

Rouet, P., Smith, F., and Jasin, M. 1994. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol. Cell. Biol. 14(12):8096-8106.

Russell, D.W., Jensen, R., Zoller, M.J., Burke, J., Errede, B., Smith, M., and Herskowitz, I. 1986. Structure of the *Saccharomyces cerevisiae HO* gene and analysis of its upstream regulatory region. Mol. Cell. Biol. 6(12):4281-4294.

Scheiffele, P., Pansegrau, W., and Lanka, E. 1995. Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and trans-specific cleavage of superhelical T-border DNA *in vitro*. J. Biol. Chem. 270:1269-1276.

Schrammeier, B., den Dulk-Ras, A., Vergunst, A. C., Jurado Jacome, E., and Hooykaas, P.J.J. 2003. 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. 31(3):860-868.

Shukla, V.K., Doyon, Y., Miller, J.C., DeKelver, R.C., Moehle, E.A., Worden, S.E., Mitchell, J.C., Arnold, N.L., Gopalan, S., Meng, X., Choi, V.M., Rock, J.M., Wu., Y.Y., Katibah, G.E., Zhifang, G., McCaskill, D., Simpson, M.A., Blakeslee, B., Greenwalt, S.A., Butler, H.L., Hinkley, S.J., Zhang, L.,

Rebar, E.J., Gregory, P.D., and Urnov, F.D. 2009. Precise genome modification in the crop species *Zea mays* using zinc-finger nucleases. Nature. 459(7254):437-441.

Stoddard, B. L., 2005. Homing endonuclease structure and function. Q. Rev. Biophys. 38(1). 49-95.

Tinland, B., Schoumacher, F., Gloeckler, V., Bravo-Angel, A. M., and Hohn, B. 1995. The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. EMBO (Eur. Mol. Biol. Organ.) J. 14:3585-3595.

Townsend, J.A., Wright, D.A., Winfrey, R.J., Lu, F., Maeder, M.L., Joung, J.K., and Voytas, D.F. 2009. High-frequency modification of plant genes using engineered zinc-finger nucleases. Nature. 459:442-446.

Van Kregten, M, Lindhout, B.I., Hooykaas, P.J.J., and Van der Zaal, B.J. 2009. *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. 22(11): 1356-1365.

Vergunst, A.C., Schrammeier, B., den Dulk-Ras, A., de Vlaam, C.M., Regensburg-Tuïnk, T.J., and Hooykaas, P.J.J. 2000. VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. 2000. Science. 290(5493):979-982.

Vergunst, A.C., van Lier, M. C., den Dulk-Ras, A., Stuve, T. A., Ouwehand, A., and Hooykaas, P.J.J. 2005. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of Agrobacterium. 2005. Proc. Natl. Acad. Sci. 102(3): 832-837.

Vogel, A. M., and Das, A. 1992. Mutational analysis of *Agrobacterium tumefaciens virD2*: tyrosine 29 is essential for endonuclease activity. J. Bacteriol. 174:303-308.

Weinthal, D., Tovkach, A., Zeevi, V., and Tzfira, T. 2010. Genome editing in plant cells by zinc finger nucleases. Trends Plant Sci. 15(6):308-321.

Wright, D.A., Townsend, J.A., Winfrey, R.J. Jr., Irwin, P.A., Rajagopal, J., Lonosky, P.M., Hall., B.D., Jondle, M.D., and Voytas, D.F. 2005. High-frequency homologous recombination in plants mediated by zinc-finger nucleases. Plant J. 44(4):693-705.

Zhu, J, Oger, P.M., Schrammeijer, B., Hooykaas, P.J.J., Farrand, S.K., and Winans, S.C. 2000. The bases of crown gall tumorigenesis. J. Bacteriol. 182(14):3885-3895.