VirD2 of Agrobacterium tumefaciens : functional domains and biotechnological applications

Kregten, M. van

Chapter 1

Introduction

1.1 General introduction

Agrobacterium tumefaciens is a pathogen that genetically transforms its hosts, dicotyledonous plants, causing crown gall disease. Plant wounds not only form the entry points for the bacteria, but also release phenolic compounds like acetosyringone that trigger the virulence genes of the bacterium. Recognition of these compounds by *Agrobacterium* triggers a series of events finally leading to genetic transformation of plant cells. Crucial for this process is the presence of a tumor-inducing plasmid, the Ti-plasmid. During transformation by *Agrobacterium*, a strand of DNA derived from the Ti-plasmid is transported to the recipient cell. This T-(transferred) strand can integrate into the recipient cell genome, after which the genes encoded on it become expressed. The expression of genes of plant hormone synthesizing enzymes that are present on the T-DNA leads to the uncontrolled tumorous growth of plants cells, thus forming the crown gall. Depending on the *Agrobacterium* strain causing the infection, T-DNA encoded genes enable crown gall cells to synthesize specific opines, amino acid derivatives which *Agrobacterium* can use as its sole source of carbon and nitrogen (Zhu *et al.*, 2000). Well-known opines include compounds such as nopaline, octopine, and mannopine. By having food produced by another organism, *Agrobacterium* creates an ecological niche for itself. Although the normal targets of *Agrobacterium*-mediated transformation (AMT) are dicotyledonous plants, transformation of a range of other species has been made possible in a laboratory setting. Hosts include, but are not limited to, monocotyledonous plants (Hooykaas – van Slogteren *et al.*, 1991), yeast (Bundock *et al.*, 1995) and other fungi (de Groot *et al.*, 1998).

Insights into the mechanism by which *Agrobacterium* processes and transfers its T-DNA in the form of a single stranded T-strand to the recipient cells have led to its development as a very important tool in plant biotechnology. The T-DNA is delimited by two imperfect repeating sequences in the Ti plasmid, the Right Border and Left Border (RB and LB) (Yadav *et al.*, 1982). None of the genes located between the repeats is necessary for T-DNA transfer, but transfer is mediated by *vir* genes located in a region of the Ti plasmid that is adjacent to the T-region. Any DNA sequence naturally present between these borders will be translocated to the recipient. In the natural situation, the genes encoding the Vir proteins are located within the *vir* region of the Ti-plasmid, usually adjacent to the T-DNA. While the *vir* region is required for processing and transfer of a T-strand, the *vir* genes

themselves are located outside of the RB and LB sequences and are thus not transferred to recipient cells. In genetic terms, *vir* genes act *in trans* upon the formation and transfer of a T-strand. The invention of the binary vector system, in which the *vir* genes are located on a different plasmid than the T-DNA, opened the way for easy transformation of a range of plant and other species, as any gene and any marker can be transferred in a relatively easy way to the recipient (Hoekema *et al.*, 1983). Genetic transformation of plants via AMT has thus become the most widely used tool in fundamental as well as applied plant research; not only for gaining insight into plant gene functioning, but also for breeding purposes, to generate crops that are, for example, drought- or disease-resistant. In the decades following these discoveries, much effort has been dedicated to unraveling the mechanism behind AMT. More recent findings that *Agrobacterium* is also able to transfer proteins to eukaryotic cells has further inspired the research community, making *Agrobacterium* and especially its *vir* region a paradigm for all kinds of transfer events from bacterial to eukaryotic cells (Vergunst *et* al., 2000, Juhas *et al.*, 2008, Pitzschke and Hirst, 2010).

The molecular mechanism enabling DNA and protein transfer by *Agrobacterium* resembles both bacterial plasmid conjugation and pathogenesis of a variety of bacterial species in its utilization of a Type IV Secretion System (T4SS). This large protein complex serves as a passage between *Agrobacterium* and the recipient. The T-DNA and Vir proteins enter the recipient cell via a T4SS encoded by genes within the *vir* region of the Ti plasmid. Such T4SSs are common in bacteria and quite conserved. The genes and operons encoding them show colinearity, i.e. their order in the genome is similar (reviewed in Schröder and Lanka, 2005). In *Agrobacterium* a second T4SS distinct from the T4SS used in pathogenesis is used to conjugate the pTi between *Agrobacterium* cells. Analysis of its origins shows that the different components have probably been acquired from a different transmissible plasmid, rather than from a duplication of the *vir* region (Alt-Morbe *et al.*, 1996). The complex multi-component structure of the T4SS requires a large investment of energy in order to be assembled. It stands to reason that they are not constitutively expressed, but are expressed when the chance of a successful transfer event is greatest. In *Agrobacterium*, the VirB/D4 T4SS becomes expressed when *Agrobacterium* detects substances from a wounded plant cell in the vicinity (see 1.2), while the T4SS involved in the conjugation of the pTi becomes expressed when opines are present (White and Winans, 2007).

The exact mechanism of T-strand and concomitant protein translocation by *Agrobacterium* has yet to be elucidated. Since similar translocation events are found in a variety of pathogenic bacteria, the biology of T4SSs becomes very relevant for health care.

For instance, several pathogens that target humans, but also other mammals, use their T4SS to translocate virulence factors called effector proteins to their hosts. Well-known pathogens utilizing this mechanism include *Helicobacter pylori* (Backert and Selbach, 2008), *Bartonella* species (Dehio, 2008) and *Legionella pneumophilia* (Swanson and Hammer 2000). In addition, plasmid exchange between bacteria via T4SSs (plasmid conjugation) leads to a major challenge in modern medicine by causing the rapid spreading of antibiotic resistance genes (Hawkey and Jones, 2009). Apart from conjugative plasmids, which encode a T4SS, there are also mobilizable plasmids that do not encode a T4SS, but only have the genes required for DNA processing; they use the T4SS encoded by other plasmids to "hitchhike" into a recipient cell (Francia *et al.*, 2004). T4SSs are still being discovered in ever more distant bacterial species, indicating that they are highly important for the exchange of genetic information between bacteria and for manipulation of their environment.

1.2 The mechanism of *Agrobacterium*-mediated transformation

As briefly mentioned above, when a plant cell is wounded, it releases phenolic compounds such as acetosyringone. These compounds are detected by *Agrobacterium* via the VirA protein, which is a receptor kinase spanning the inner membrane (Figure 1) (Lee *et al.*, 1995). Upon detection of phenolic compounds, VirA phosphorylates VirG, which then binds to the *vir* boxes within the regulatory region of other *vir* genes, thus stimulating their expression (Lee *et al.*, 1996). The newly expressed Vir proteins perform several tasks in the cell. The T-strand is prepared by the nicking reaction of VirD1 and VirD2, and this process is enhanced by binding of VirC1 to the overdrive sequence, adjacent to the RB (see 1.3 for details). VirD2 stays covalently attached to the T-strand. This protein-DNA complex is referred to as the T-complex. VirD2 interacts with VirC1 and the 3 VirD2-Binding Proteins (VBP1-3) to enhance recruitment to the T4SS (see 1.4 and Chapter 3) (Atmakuri *et al.*, 2007, Guo *et al.*, 2007a and b).

Proteins VirB1-11, encoded by the *virB* operon, and VirD4, form the acetosyringone-inducible T4SS of *Agrobacterium* (see 1.5 for details), via which the T-complex and several vir effector proteins are translocated to the cytosol of the recipient cell. In the recipient cell, the T-complex is imported into the nucleus, where it is released from its interactions (see 1.6 for details). Finally, the T-strand can integrate into the genome of the recipient cell and the genes on the T-strand can become expressed (see 1.7 for details).

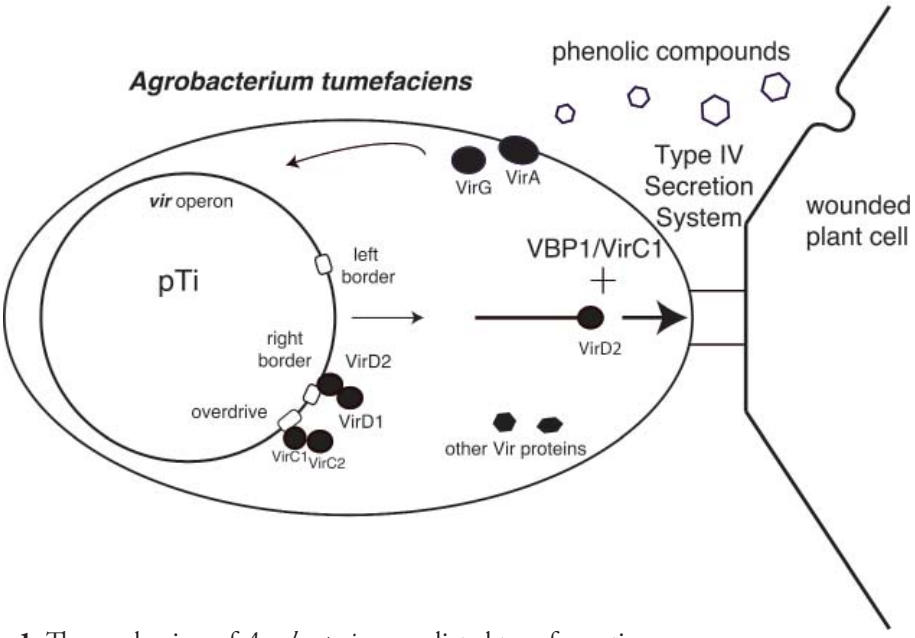

Figure 1: The mechanism of *Agrobacterium*-mediated transformation

1.3 Generation of the T-strand

For the release of the T-strand, a single strand break (nick) is generated at the border sequences that flank the T-DNA. This border nicking is performed by VirD2, a protein belonging to a large family of relaxases which are essential for conjugation and mobilization of plasmids. Relaxases can be placed in several different families and the two families discussed in this thesis can be classified according to their catalytic sites, which contain one or two tyrosine residues. (Frost *et al.*, 2005, Garcillan-Barcia *et al.*, 2009). Examples of relaxases that contain one catalytically active tyrosine residue are MobA of plasmid RSF1010, TraI of plasmid RP4, and VirD2, while examples of relaxases that have two such residues are TrwC of plasmid R388, and TraI of the F plasmid. (Francia *et al.*, 2004). Usually, relaxases contain other well-characterized functional domains C-terminal from their relaxase domain, such as primase or helicase domains (Garcillan-Barcia *et al.*, 2009). VirD2 differs in that respect as it contains a domain of unknown function at its C-terminus (Shurvinton *et al.*, 1992).

At the site of initiation (the RB in the T-region of *Agrobacterium*, *oriT* in conjugative and mobilizable plasmids), the complex of proteins required for nicking of the borders/ *oriT*, called the relaxosome, is assembled. In *Agrobacterium*, the relaxosome consists of the

cofactor VirD1 and the relaxase VirD2, but other systems require different co-factors besides the relaxase (see below). The role of VirD1 in processing of the T-strand is much less clear than that of VirD2. From *in vitro* experiments it is known that, in addition to VirD2, VirD1 is required to nick double-stranded DNA, while VirD2 alone is sufficient to nick single-stranded DNA (Scheiffele *et al.*, 1995).

VirC1, the first of the two proteins encoded by the *virC* operon, binds to the overdrive sequence, a sequence near the RB, but outside of the T-DNA (Toro *et al.*, 1989). This sequence acts as an enhancer of the nicking reaction and of T-strand formation (Peralta *et al.*, 1986, Van Haaren *et al.*, 1987, Toro *et al.*, 1989). VirC2 is also a DNA-binding protein, with a ribbon-helix-ribbon DNA binding domain (Lu *et al.*, 2009). VirC1 and VirC2 are not essential for T-strand processing, but they increase the amount of T-strands produced (Atmakuri *et al.*, 2007).

The only domain of VirD2 important for T-strand processing is the relaxase domain. However, it is not uncommon for relaxases to contain a helicase domain C-terminally from the relaxase domain, which also plays a role in T-strand processing (e.g. relaxase/helicase TrwC of plasmid R338 and relaxase/helicase MobA of plasmid RSF1010 (Garcillan-Barcia *et al.*, 2009)). The presence of the helicase domain can facilitate the separation of the DNA strands prior to making a single strand break or, alternatively, aid in releasing the T-strand once this break has been made. Cofactors with other auxiliary functions are also known; plasmid RP4 has in its relaxosome the cofactors TraJ, to recruit its cognate relaxase/helicase TraI, and TraH and TraK to stabilize the interactions between the relaxase and the DNA. The F plasmid relaxosome requires besides relaxase TraI, the TraY protein to bend the DNA into a conformation that can be more easily processed by the relaxase, and Integration Host Factor to aid in assembly of the relaxosome (reviewed in Byrd and Matson, 1997).

Border cleavage occurs by attack of the catalytically active tyrosine residue (Tyr-29) of VirD2 on the phosphodiesterbond of the *nic* position within the RB. Other relaxases also have a conserved tyrosine residue at a position quite close to the N-terminus performing this action on the *nic* site of their cognate *oriT* (classification of the relaxase family is based on this feature; see Garcillan-Barcia *et al.*, 2009). This reaction creates a covalent bond between the relaxase protein and the DNA (Byrd and Matson, 1997). Relaxase-mediated cleavage of single stranded DNA requires Mg^{2+} ions and water and is isoenergetic; therefore, the reaction is reversible. Ironically, the formation of the covalent protein-DNA bond itself does not actually relax the DNA in the sense that tension due to DNA supercoiling is not decreased; another factor is required to release the tension (Byrd and Matson, 1997).

In all systems but the T-strand transferring mechanism of *Agrobacterium*, a single relaxase-mediated cleavage is enough to create the protein-DNA intermediate. However, to release a perfect pTi-derived T-strand, a second cleavage needs to be performed at the LB. In that manner, only the region in between the RB and LB is transferred as a T-strand to recipient cells, rather than the whole T-DNA containing plasmid. Models for the initiation and termination of T-strand processing are presented in 1.4. Although T-strand transfer is the norm, the whole T-DNA containing plasmid can be transferred when the LB is deleted (Caplan *et al.*, 1985, Miranda *et al.*, 1992). Also when both borders are present, sometimes border skipping is observed, resulting in transfer of the entire plasmid (Van der Graaff *et al.*, 1996, Kononov *et al.*, 1997).

It is nowadays generally believed that uncoupling of the relaxase only occurs after transfer of the relaxase moiety of the T-complex to the recipient cell, although actual experimental evidence is scarce. For the relaxase TrwC from plasmid R388, nicking activity after translocation has been demonstrated in the recipient cell (Draper *et al.*, 2005). In chapter 2, experimental evidence of the translocation of VirD2 as part of the T-complex further confirms this model. These data, demonstrating protein transfer, fit well into the emerging picture of the T4SS as a protein secretion machine, rather than a DNA-transferring machine (Vergunst *et al.*, 2000, Christie and Vogel, 2000, Juhas *et al.*, 2008, Schröder and Lanka, 2005).

1.4 Models for the initiation and termination of T-strand processing

Initiation of T-strand processing has been well described: the catalytic Tyr residue of the relaxase in question creates a covalent bond between the relaxase and the 5' end of the DNA at the *nic* site within oriT or at the RB. This reaction is reversible. (Pansegrau and Lanka, 1996, reviewed in Byrd and Matson, 1997). In the case of plasmids, this one reaction and subsequent processing is all that is required, while for *Agrobacterium*, processing starts at the RB and needs to be terminated at the LB. The exact mechanism of LB cleavage is not yet clarified, although it is clear that the main product of processing in *Agrobacterium* is the T-strand. In the case of *Agrobacterium*, transfer of the entire plasmid can happen when the LB is skipped, which can result in the integration of vector backbone into the genome of the recipient cell (Ramanathan and Veluthambi, 1995, Van der Graaff *et al.*, 1996, Kononov *et al.*, 1997). DNA sequences from the plasmid backbone have been found in 75% of a

population of AMT-generated tobacco plants have (Kononov *et al.*, 1997). Most of these backbone sequences were linked to (part of) the T-DNA, via the LB or RB. Ramanathan and Veluthambi (1995) observed T-DNA integrations originating from the LB of a Ti plasmid, indicating that processing can also accidentally start from the LB. This is also supported by earlier descriptions of processed T-strand intermediates originating from both borders, with protein bound to the 5' end (Dürrenberger *et al.*, 1989, Steck *et al.*, 1989).

After initiation of T-strand processing, the T-strand needs to be separated from the plasmid. There are two possibilities for this process. One model states that, in analogy to rolling circle replication, which occurs in bacteriophages and some plasmids, the T-strand is released by displacement synthesis, mediated by DNA polymerase starting from the 3' OH end of the RB/oriT (reviewed in Lanka and Wilkins, 1995, reviewed in Byrd and Matson, 1997, reviewed in Khan, 2005). Another possibility is that the T-strand is separated from its complementary strand by helicase activity (reviewed in Lanka and Wilkins, 1995, reviewed in Byrd and Matson, 1997). In both cases, the single stranded T-strand must be relatively stable, allowing for its translocation to the recipient cell (reviewed in Byrd and Matson, 1997). Before the relaxase-T-strand bond is broken, the relaxase needs to be translocated, so that release of the T-strand takes place within the recipient cell. This implies that the donor and recipient cell should be in close contact during the processing of the T-strand. In that manner, T-strand processing should be coupled to T-strand translocation, with the relaxase protein at the 5' end piloting the T-strand into the recipient cell as current models suggest (reviewed in Chen *et al.*, 2005 Christie, 2004, Llosa *et al.*, 2002). Evidence has been found for the translocation of TrwC of plasmid R388 (Draper *et al.*, 2005). Additional evidence for the translocation of VirD2, piloting the T-strand, is presented in Chapter 2, while Vergunst and co-workers have shown that low-level translocation of VirD2 and MobA (of plasmid RSF1010) is possible in the absence of a T-strand.

While the formation of the T-strand-relaxase complex takes place in the donor cell, breakage of the relaxase-T-strand bond would then occur in the recipient cell. Efficient bacterial conjugation therefore requires cell-cell contact between donor and recipient (reviewed in Alvarez-Martinez and Christie, 2009). Only for *Agrobacterium*, VirD2-T-strand complexes can be precipitated from *Agrobacterium* cells that are not in contact with recipient cells (but induced to express *vir* genes by administration of acetosyringone) (Cascales and Christie, 2004). After the T-strand is translocated to the recipient cell, a combination of transferred and resident enzymatic activities will determine the final outcome of the DNA transfer. In the case of conjugation, a double-standed circular plasmid will be formed. For

Agrobacterium-mediated DNA transfer, the T-strand finally integrates into the genome of the recipient. In either case, the relaxase is eventually released from the T-strand. Evidence suggests that VirD2 is not responsible for ligating the T-strand to the host genomic DNA (Zieniemowicz *et al*., 2000). Therefore, VirD2 is probably released from the T-strand prior to integration, but the exact mechanism behind this process remains obscure.

1.5 The Type IV Secretion System

After the T-strand is released, it is recruited to the Type IV Secretion System. The T4SS is a multiprotein complex which spans the inner membrane, the periplasm and the outer membrane. In *Agrobacterium,* it extends into the pilus, which protrudes outside of the cell. As already mentioned in section 1.1, the T4SS of *Agrobacterium* that is induced by plant phenolic compounds and which is responsible for the translocation of Vir proteins and the T-strand to the recipient plant cells is composed of the products of the *virB* operon, VirB1-11, in combination with the coupling protein VirD4. *Agrobacterium* also has a second T4SS, encoded by its *tra* operon, which is used for the conjugative transfer of the pTi between *Agrobacterium* cells (Alt-Mörbe *et al*., 1996). For many years the T4SS, such as that encoded by the F plasmid of *E. coli*, was viewed as a channel that must have evolved for translocation of DNA (Frost *et al*., 1994). The coupling protein could serve to "push" DNA through the secretion channel (Llosa *et al*., 2002).

In a seminal paper describing the translocation of the virulence proteins VirE2 and VirF protein of *Agrobacterium* independently of the T-strand, it was proposed that T4SSs might in principle be regarded as protein secretion channels, through which DNA is only transported when bound to the relaxase (Vergunst *et al*., 2000). Evidence for translocation of the relaxases VirD2 and MobA in the absence of T-DNA transfer was later found, as well as translocation of VirE2, VirE3, VirD5 and VirF (Vergunst *et al*, 2005, reviewed in Christie, 2004, Ding *et al*., 2003, Lawley *et al*., 2003).

More recently, T4SS have been identified in pathogenic microorganisms that do not transfer DNA at all, but exclusively mediate the translocation of virulence factors into host cells. Such T4SS are essential for virulence.

T4SS have thus two main functions: conjugation and effector protein translocation. For conjugation to occur, the relaxase-processed DNA intermediate needs to be formed and recognized by the T4SS which then provides it with a passageway into the recipient

cells (Christie, 2004, Llosa *et al.*, 2002). When the T4SS is only used for the translocation of effector proteins, the requirements for the T4SS are approximately the same as for conjugation: a coupling protein to mediate interaction between the effector protein and the T4SS secretion channel and facilitation of transfer (in)to the recipient cell.

A third class of T4SS was more recently discovered in some bacteria, which are involved in the uptake from or release of DNA into the environment. In *Helicobacter pylori*, a DNA uptake system employs a T4SS lacking a coupling protein (Hofreuter *et al.*, 2001). The mechanism of DNA uptake is not yet clear, although it has recently been suggested that the *Helicobacter* homolog of VirB4 energizes DNA transport (Stingl *et al.*, 2010)

In *Neisseria gonorrhea*, DNA is released into the extracellular space via a T4SS (Hamilton *et al.*, 2005). The very complex structure of the T4SS has been the subject of many reviews (e.g. Alvarez-Martinez and Christie, 2009, Gomez-Ruth and Coll, 2006, Schroder and Lanka, 2005). Interestingly, employing the DNA moiety as a detectable target, the order in which the T-strand encounters the structural proteins of the T4SS of *Agrobacterium* has been elucidated by immunoprecipitation of cross-linked Vir protein containing complexes (Cascales and Christie, 2004). Determination of the order of contacts of just the secreted proteins within the T4SS is technically extremely challenging and has not yet been reported.

What is currently known about translocated proteins is that they contain a translocation signal. This T4SS secretion signal has been identified in the C-terminal end of translocated Vir proteins, containing several arginine residues, making the signal positively charged. The relaxase of mobilizable plasmid RSF1010, MobA, also contains such a signal at its C-terminus. The processed relaxase-T-DNA complex of RSF1010 is known to "hitchhike" through the T4SS of *Agrobacterium* (Bravo-Angel *et al.*, 1999). The translocation signal of VirF can even be used to aid the translocation of unrelated proteins that normally cannot pass through the T4SS, such as Cre (Vergunst *et al.*, 2000 and 2005).

1.6 Travel into the nucleus

The T-complex produced by *Agrobacterium*, consisting of VirD2 covalently bound to the T-strand, is delivered into the cytosol of the recipient cell. There it becomes coated by VirE2, which is delivered separately from the T-complex to the recipient cell (Otten *et* al., 1984, Figure 2). It has been postulated that the function of this coating could be to protect the T-strand from degradation by recipient cell nucleases (Citovsky *et al.*, 1989). However, this

has never been shown *in vivo*, only *in vitro* using λ exonuclease (Dürrenberger *et al.*, 1989). The T-complex is then imported into the recipient cell nucleus, due to the presence of a nuclear localization signal (NLS) in VirD2.

In *Arabidopsis thaliana*, nuclear import is mediated by the importin AtKapα, which interacts with VirD2 (Ballas and Citovsky, 1997). Several other interactors have been described for VirD2 in *Arabidopsis*: several cyclophilins (Deng *et al.*, 1998) and the TATA-box Binding Protein (TBP) (Bako *et al.*, 2003). Interaction of VirD2 with cyclophilins is discussed in more detail in Chapter 2. Although the interaction between VirD2 and TBP suggests that a T-complex, when arriving within the nucleus, might by targeted to sites where TBP is present, this is probably not the case. It has been shown that integration sites of T-strands are distributed randomly throughout the genome (Kim *et al.*, 2007). While VirD2 can mediate border fusion *in vitro* (Pansegrau *et al.*, 1996), VirD2 does not appear to ligate the T-strand to the plant genome (Ziemienowicz *et al.*, 2000).

Another factor aiding in nuclear import is VirE2, which interacts with VirE2 Interacting Protein 1 (VIP1). VIP1 is a transcription factor in the *Arabidopsis* defense pathway, and is phosphorylated by MPK3 when plants perceive the presence of *Agrobacterium* or other pathogens. VIP1 then travels to the nucleus, may bind VirE2 and may take the T-complex along. In this way, the plant's own defense mechanism may help bring the T-complex into the nucleus (Djamei *et al.*, 2007). *In vitro*, VIP1 also binds to plant nucleosomes, which may link the T-DNA to the chromatin via a VIP1-VirE2-T-DNA complex (Lacroix *et al.*, 2008).

Before or during integration into the recipient's genome, the T-strand needs to be freed from its binding partners. How VirD2 is released from the T-strand is currently unknown. The F-box protein VirF may play a role in the decoating of VirE2, with the help of the VirE2-interacting protein VIP1. VIP1 brings VirE2 into contact with VirF, which, interacts with the tobacco proteolysis machinery. Through this interaction, both VIP1 and VirE2 are destabilized. Thus this mechanism may help to uncoat the T-DNA (Tzfira *et al.*, 2004). Interestingly, VirF is a host range factor: requirement for this factor is not absolute, but it is needed to induce larger tumors on certain plant species. For instance, VirF is not needed for the transformation of *Saccharomyces cerevisiae* (Bundock *et al.*, 1995). Also, *Arabidopsis* was found to contain an F-box protein that can functionally replace VirF (Zaltsman *et al.*, 2010).

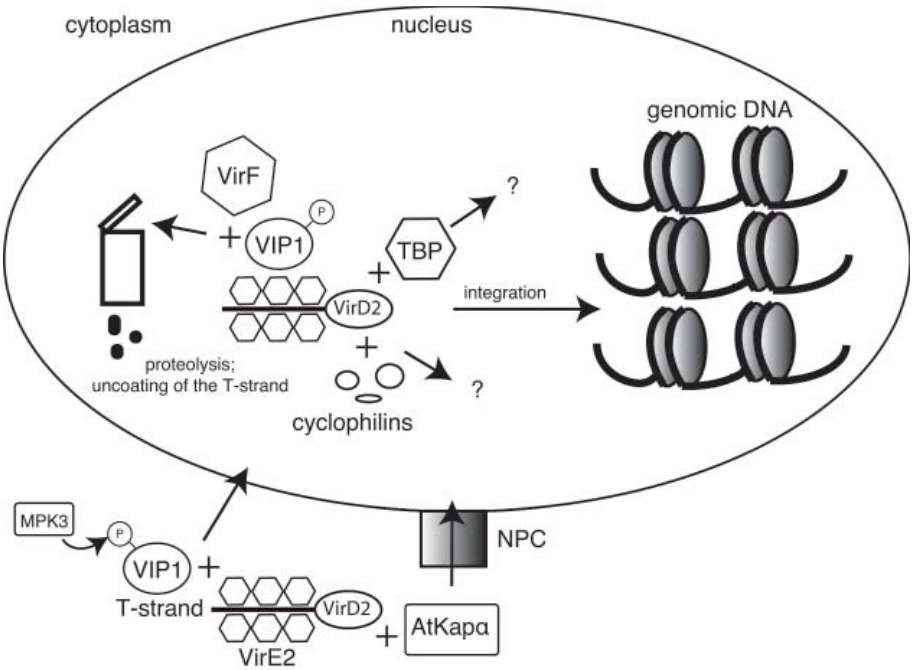

Figure 2: Travel of the VirD2-T-strand complex into the nucleus

1.7 T-DNA integration

In the case of AMT of higher plants, T-DNA insertions have been found to be randomly dispersed over the genome of the recipient (Kim *et al.*, 2007). Although transferred as a single strand, evidence has been presented that indicates that T-DNA can become double-stranded before integration (Tzfira *et al.*, 2003). However, other integration events that were analyzed in detail can only be explained by assuming that the T-DNA was single-stranded at the time of integration (Tzfira *et al.*, 2004).

The main mechanism via which any exogenous DNA integrates in higher eukaryotes is non-homologous recombination (NHR). This consists of two pathways: non-homologous end-joining (NHEJ) and microhomology-mediated recombination (MMEJ). NHEJ is the main pathway of T-DNA integration, and often results in small deletions or insertions of both the T-DNA and the genomic locus where it is inserted. During NHEJ, both ends of the break are processed and ligated. When integrated via NHEJ, the T-DNA is also subject to such processing (reviewed in Bleuyard *et al.*, 2006). In yeast, it was found that enzymes from the NHEJ pathway were essential for the integration of T-DNA lacking homology to the

yeast genome (Van Attikum *et al.*, 2001). When integration via NHEJ occurs, often random 'filler' DNA is added to the site of integration. Filler DNA may introduce microhomology, enabling integration via MMEJ. Filler DNA is also associated with preservation of the T-DNA ends (Windels *et al.*, 2003).

In contrast to the error-prone NHR events, homologous recombination (HR)-mediated events can lead to very precise site-specific alteration of the genome by exchanging the genetic locus or gene of interest, or a part thereof, for a modified version that is introduced into the cell, a process also called gene targeting (GT). Homologous recombination can only occur when the incoming DNA has homology to the genome of the recipient. The break is processed to produce a single stranded overhang, which invades the homologous region in another stretch of DNA. This region is used to bridge the gap and serve as a template for repair. Eventually, this creates a Holliday structure of four strands of DNA. When this structure has been resolved, the remaining nick in the DNA is repaired (Bleuyard *et al.*, 2006).

When integration events occur via NHEJ, such an exchange does not take place; the gene of interest will integrate elsewhere and thus co-exist with the original locus. Moreover, sequences at random positions in the genome will be disturbed. Obviously, HR is the desired pathway for accomplishing site-directed mutagenesis. The field of GT is of high interest for both plant and medical science; fundamentally, to enable the study of the function of genes in their context, and applied, for the development of gene therapy. Much effort is devoted to develop an efficient way to achieve GT in a variety of species.

The chance of integration at a desired position is increased by the presence of a DSB at this site in the genome. This was first established in mammalian cells (Jasin *et al.*, 1985). In plants, this was shown by Puchta *et al.*, who showed in protoplasts that expression of homing endonuclease I-SceI (see below) will increase the chance of homologous recombination (HR) of a substrate carrying an I-SceI recognition site (Puchta *et al.*, 1993). Integration of the T-DNA into the DSB site via NHR also happens (Salomon and Puchta, 1998).

1.8 Tools for targeted integration: zinc fingers nucleases and homing endonucleases

As mentioned above, it has been known for a rather long time that the presence of DSBs in the genome is stimulatory to GT in higher eukaryotes (Jasin *et al.*, 1985). To create DSBs in a genome, different DNA binding proteins have been designed and discovered. The two

most used enzymes for the creation of a DSB are zinc finger nucleases (ZFNs) and homing endonucleases (HEs).

Zinc fingers (ZFs) are DNA-binding motifs; a specific sequence of 30 amino acids can be predicted to bind a certain three basepair sequence (triplet). The "zinc finger code" has been described for many of the GNN, ANN, and CNN triplets (Segal *et al.*, 1999, Dreier *et al.*, 2001, Dreier *et al.*, 2005). For TNN triplets, the available code is still limited. By combining different ZFs, polydactyl ZF domains (PZFs) can be made, which can target a longer, specific sequence of DNA. These PZFs can in principle be combined with other functional groups to create novel functional proteins, e.g. with activator or repressor domains, to create artificial transcription factors, or with a nuclease domain, to create a ZFN.

To create a ZF nuclease, which has DNA cleaving activity, a nuclease domain thus needs to be fused to the PZF. The nuclease domain of choice is the nuclease domain of the restriction enzyme *Fok*I. In *Fok*I, the DNA binding and cleaving activities are located in separate domains of the protein. Therefore, the *Fok*I nuclease domain can be made available for fusion to ZFs, creating a novel restriction enzyme (the zinc finger nuclease, ZFN) with a custom-made recognition sequence, which was first achieved by Kim *et al.*, in 1996. The ZFNs have been eagerly used by groups from a variety of research fields to create genomic lesions in a large number of organisms (reviewed in Carroll, 2008) to study DNA repair mechanisms or explore the possibilities for therapeutic genetic manipulation. With ZFN technology on the rise, hopes are high for the development of novel therapeutic gene targeting strategies for human disease and crop improvement in plants, as well as fundamental scientific questions in the area of DNA repair.

Apart from zinc fingers, homing endonucleases (HEs) can also be used for the introduction of highly specific DSBs in a genome (e.g. Chiurazzi *et al.*, 1996, Pacher *et al.*, 2007). HEs have been discovered in different organisms. Some are encoded by open reading frames present in introns of otherwise non-related genes, while others, designated inteins, are intervening sequences which are spliced out post-translationally (Stoddard, 2005). They function as selfish genetic elements; there is no obvious benefit for the host organism. The HE copies the homing endonuclease ORF into an allele that does not contain it, or to unrelated sites that happen to contain the recognition site. The HE recognizes a long (14-40 bp) DNA sequence, which is present in the homologous allele that does not contain the coding sequence for the HE. Sequence recognition is somewhat relaxed, indicating that the homing endonuclease will not be lost when small mutations in the homologous allele will

occur (Jurica and Stoddard, 1999).

Five families of homing endonucleases have been identified: the LAGLIDADG, HNH, His-Cys box, and GIY-YIG families, and a largely uncharacterized family in the introns of cyanobacteria (Jurica and Stoddard, 1999). The LAGLIDADG class is most extensively characterized, and consists of ORFs in introns mainly found in the mitochondria and chloroplasts of single-celled eukaryotes. The well-studied homing endonucleases I-SceI, I-CreI and HO belong to this group. Both I-SceI and HO contain two LAGLIDADG motifs, indicating that they function as monomers, while homing endonucleases that contain only one LAGLIDADG motif function act as homodimers. I-CreI is an example of the latter class (Stoddard, 2006, Chevalier *et al.*, 2004).

Homing endonucleases of the LAGLIDADG family create DSBs with 4 nucleotide 3'-OH overhangs. The ends are processed and the homologous allele is used as a template for HR (Jurica and Stoddard, 1999). The homing endonuclease is only involved in the initial creation of the DSB and not in the subsequent repair.

I-SceI was the first homing endonuclease to be identified. It was discovered that the yeast mitochondrial marker ω was inherited in a non-Mendelian way, with ω^- converting into ω^+ at a very high rate (reviewed in Dujon, 1989). Only later it was found that ω was an optional intron in the *LSU* mitochondrial gene and the non-Mendelian inheritance pattern was caused by the activity of I-SceI, encoded by an ORF in the optional intron, causing the optional intron to be copied into the other allele (Dujon, 1980).

The yeast mating-type switching enzyme HO is an example of a domesticated homing endonuclease; it is encoded by a free-standing gene, and contains only remnants of the sequences required for intein splicing. Exactly how HO became domesticated is unknown. HO is responsible for changing the mating type of a yeast cell from **a** to α or vice versa. Inactive loci for **a** or α flank the locus for the active mating type (MAT locus). HO creates a DSB whereupon either one of the inactive loci are copied into the active locus (Pietrokovski, 1994, Strathern *et al*, 1982, Kostriken *et al.*, 1983). This ensures that there is a heterogeneous population, capable of mating to create diploid cells.

Although the application of ZFNs for site-directed mutagenesis can be accompanied with problems (of a mostly technical nature, rather than fundamental) (Durai et al., 2005), recent progress in this field is very impressive (reviewed in Wu *et al.*, 2009). Therefore it can be expected that problems such as to how to design the optimal PZF domains for a particular genetic locus without any undesired effects at other loci can in principle be solved. Nevertheless, alternatives to PZF-based technology will be most welcome. The development

of HEs with modified recognition sequences is now coming of age, which will enable the application of homing endonucleases at a broader scale in the future (Pâques and Duchateau, 2007, Galetto *et al*., 2009).

1.9 Outline of this thesis

The focus of the research described in this thesis is on VirD2. This protein is at the center of AMT as it is not only essential for T-strand processing, but also accompanies the T-strand to its final destination in the recipient genome. At the start of the research leading to this thesis, it was already known that only the N-terminal half of VirD2 is required for T-strand processing. The function of the C-terminal half was basically unknown; only the very C-terminal end had been implied as important for translocation and/or integration of the T-DNA. Chapter 2 describes experiments in which we defined a minimal version of VirD2, stripped of all but the relaxase domain. When supplied with a T4SS translocation signal derived from the virulence protein VirF, this minimal VirD2 restores virulence of a VirD2 mutant *Agrobacterium* strain. This indicates that the relaxase domain is the only domain absolutely required for virulence. The other domains of VirD2, the domain of unknown function (DUF) and the C-terminus, were dispensable and could be replaced by the VirF translocation signal. DUF and the C-terminus are possibly involved in the localization of the T-complex to the T4SS, a function that apparently can be taken over by the very small T4SS translocation signal of VirF (F). It also gives support for the prevailing model of the T4SS as a protein – rather than a DNA - translocation machine. In Chapter 3 we describe studies to elucidate the exact role of DUF in localization. We determined, by tagging GFP to VirD2 truncations, that the N-terminal 60 amino acids of DUF, DUF-60, are crucial for polar localization of VirD2. In VirD2-204, fluorescent signal is spread throughout the cell, while in VirD2-264 polar localization is restored to near wild type levels. The translocation signal of VirF, which consists of only 37 amino acids, also acts by restoring polar localization to VirD2 truncations. The phenotype of polar localization correlates with the ability of the corresponding *Agrobacterium* strain to transform plants.

Protein translocation from *Agrobacterium* to any recipient is of interest both for fundamental research and for biotechnological applications. By creating fusions to Cre and demonstrating Cre recombinase activity in the recipient, Vergunst *et al.* proved that several Vir proteins were translocated. Moreover, it was shown that these Vir proteins contained

C-terminal translocation signals, as these protein parts could lead to translocation of Cre fusion proteins to recipient cells via the T4SS of *Agrobacterium* (Vergunst *et al*., 2000 and 2005). Our finding that a T4SS translocation-deficient VirD2 mutant could be made T4SS translocation-competent again by addition of the translocation signal of VirF (designated F) led us to develop VirD2-fusion proteins that can actually be translocated to plants. While N-terminal fusion of protein domains to VirD2 easily lead to impairment of the relaxase function, C-terminal fusions are likely to block the native T4SS translocation signal of VirD2. However, by placing the fusion protein in between VirD2 and F, the relaxase function remained intact and translocation ability was restored. Detection of VirD2-mediated translocation events is easily done by assessing the expression of marker genes encoded on the T-DNA. Furthermore, VirD2 is covalently attached to a T-strand, which results in a possibility to combine the delivery of DSB-inducing protein domains with the concomitant delivery of a DNA molecule that could potentially be used as repair template.

In Chapter 4, we describe the translocation efficiency of several VirD2 fusion proteins. The proteins studied are of great interest for the development of site-directed mutagenesis methods, as they contained zinc fingers or the homing endonucleases I-SceI and HO. All tested protein domains could be translocated, albeit with very different efficiencies. Especially the I-SceI domain was found to be translocated at a relatively high level in a translational fusion with VirD2 and F. Recovery of I-SceI induced damage after *in planta* expression showed that fusion with VirD2 and F did not abrogate the nuclease activity of the fusion protein. We can therefore conclude that protein domain delivery into plant cells via *Agrobacterium*-mediated transformation using VirD2 fusion proteins is feasible.

In Chapter 5, we show that the recombinant relaxase/effector proteins VirD2-I-SceI-F and VirD2-204-I-SceI-F retain cleavage activity on an I-SceI site after delivery to plant cells by *Agrobacterium*-mediated transformation. Several I-SceI-induced mutations were recovered from *Arabidopsis* transformants treated with such recombinant relaxase effector protein, but no GT events were discovered.

The results presented highlight the biotechnological possibilities of functional protein domain delivery via VirD2, and present a lead for further investigation into the development of site-directed mutagenesis strategies in plants via *Agrobacterium*-mediated transformation.

1.10 References

Alt-Morbe, J., Stryker, J. L., Fuqua, C., Li, P. L., Farrand, S. K., and Winans, S. C. 1996. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. J. Bacteriol. 178(14): 4248-57.

Alvarex-Martinez, C.E., and Christie, P.J. 2009. Biological diversity of prokaryotic Type IV Secretion Systems, Microbiol. and Mol. Biol. Rev. 73(4): 775-808.

Atmakuri, K., Cascales, E., Burton, O. T., Banta, L., and Christie, P. J. 2007. *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. EMBO (Eur. Mol. Biol. Organ.) J. 26(10): 2540-2551.

Backert, S., and Selbach, M. 2008. Role of type IV secretion in *Helicobacter pylori* pathogenesis. Cell. Microbiol. 10(8): 1573-1581.

Bako, L., Umeda, M., Tiburcio, A.F., Schell, J., and Koncz, C. 2003. The VirD2 protein of *Agrobacterium*-transferred DNA interacts with the TATA box binding–protein and a nuclear protein kinase in plants. Proc. Natl. Acad. Sci. U.S.A. 100(17): 10108-10113.

Ballas, N, and Citovsky, V. 1997. Nuclear localization binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. Proc. Acad. Natl. Sci. U.S.A. 94(20): 10723-10728.

Bleuyard, J.Y., Gallego, M.E., and White, C.I. 2006. Recent advances in understanding of double-strand break repair machinery of plants. DNA repair. 5: 1-12.

Bravo-Angel, A. M., Gloeckler, V., Hohn, B., and Tinland, B. 1999. Bacterial conjugation protein MobA mediates integration of complex DNA structures into plant cells. J. Bacteriol. 181(18): 5758-5765.

Bundock, P., den Dulk-Ras, A., Beijersbergen, A., and Hooykaas, P. J. 1995. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. EMBO (Eur. Mol. Biol. Organ.) J. 14(13): 3206-14.

Byrd, D.R., and Matson, S.W. 1997. Nicking by transfesterification: the reaction catalyzed by a relaxase. Mol. Microbiol. 25(6): 1011-1022.

Caplan, A.B., Van Montagu, M., and Schell, J. 1985. Genetic analysis of integration mediated by single T-DNA borders. J. Bacteriol. 161(2): 655-664.

Carroll, D. 1008. Zinc finger nucleases as gene therapy agents. Gene Ther. 15(22): 1463-1468.

Cascales, E., and Christie, P.J. 2004. Definition of a bacterial type IV secretion pathway for a DNA substrate. Science. 304(5674): 1170-1173.

Chen, I., Christie, P.J., and Dubnau, D. 2005. The ins and outs of DNA transfer in bacteria. Science. 310(5753):1456-1460.

Chevalier, B, Sussman, D., Otis, C., Noel, A. J., Turmel, M., Lemieux, C., Stephens, K., Monnat,

R. J., Jr., and Stoddard, B. L. 2004. Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. Biochemistry. 43(44): 14015-14026.

Chiurazzi, M, Ray, A., Viret, J.F., Perera, R., Wang, X.H., Lloyd, A.M., and Signer, E.R. 1996. Enhancement of somatic introchromosomal homologous recombination in Arabidopsis by the HO endonuclease. The Plant Cell. 8: 2057-2066.

Christie, P.J. 2004. Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. Biochem. Biophys. Acta. 1693: 219-234.

Christie, P.J., and Vogel, J.P. 2000. Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. Trends in Microbiol. 8(8): 354-360.

Christie, P.J., Ward, J.E., Winans, S.C, and Nester, E.W. 1998. The *Agrobacterium tumefaciens virE2* gene product is a single-stranded-DNA-binding protein that associates with T-DNA. J. Bacteriol. 170(6): 2659-2667.

Citovsky, V., Wong, M. L., and Zambryski, P. 1989. Cooperative interaction of *Agrobacterium* VirE2 protein with single-stranded DNA: implications for the T-DNA transfer process. Proc. Natl. Acad. Sci. U.S.A. 86: 1193-1197.

Dehio, C. 2008. Infection-associated type IV secretion systems of *Bartonella* and their diverse roles in host cell interaction. Cell. Microbiol. 10(8):1591-1598.

Deng, Chen, L., Peng, W. T., Liang, X., Sekiguchi, S., Gordon, M. P., Comai, L., and Nester, E. W. 1999. VirE1 is a specific molecular chaperone for the exported single-stranded-DNA-binding protein VirE2 in *Agrobacterium*. Mol. Microbiol. 31(6): 1795-1807.

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.

Djamei, A., Pitzschke, A., Nakagami, H., Rajh, I and Hirt, H. 2007. Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. Science. 318(5849): 453-456.

Draper, O., Cesar, C.E., Machon, C., de la Cruz, F., and Llosa, M. 2005. Site-specific recombase and integrase activities of a conjugal relaxase in recipient cells. Proc. Natl. Acad. Sci. U.S.A 102(45): 16385-16390.

Dreier, B, Beerli, R.R, Segal, D.J., Flippin, J.D, and Barbas, C.F. 2001. Development of zinc finger domains for recognition of the 5'-ANN-5' family of DNA sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 276: 29466-29478.

Dreier, B, Fuller, R.P., Segal, D.J., Lund, C.V., Blancefort, P., Huber, A., Koksch, B., and Barbas, C.F. 2005. Development of zinc finger domains for recognition of the 5'-CNN-5' family of DNA sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 280: 35588-35597.

Dujon, B. 1980. Sequence of the intron and flanking exons of the mitochondrial 21S rRNA gene of yeast strains having different alleles at the omega and rib-1 loci. Cell 20(1): 185-97

Dujon, B. 1989. Group I introns as mobile genetic elements: facts and mechanistic speculations – a review. Gene. 82: 91-114.

Durai, S., Mani, M., Kandavelou, K., Wu, J., Porteus, M.H., and Chandrasegaran, S. 2005. Zinc finger nucleases: custom-designed scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res. 33(18): 5978-5990.

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. U.S.A 86(23): 9154-9158.

Francia M.V., Varsaki, A., Garcillán-Bárcia, M.P., Lahorre, A., Drainas, C., and de la Cruz, F. 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. 28(1): 79-100.

Francis, K.E., and Spiker, S. 2005. Identification of *Arabidopsis thaliana* transformants without selection reveals a high occurrence of silenced T-DNA integrations. The Plant Journal. 41: 464-477.

Frost, L.S., Ippen-Ihler, K., and Skurray, R.A. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol. Rev. 58(2): 162-210.

Frost, L.S., Leplae, R., Summers, A.O., and Toussaint, A. 2005. Mobile genetic elements: the agents of open source evolution. Nat. Rev. Microbiol. 3(9):722-732.

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.

Garcillan-Barcia, M. P.,Jurado, P., González-Pérez, B., Moncalián, G., Fernández, L.A., and de la Cruz, F. 2007. Conjugative transfer can be inhibited by blocking relaxase activity within recipient cells by intrabodies. Mol. Microbiol. 63(2): 404-416.

Garcillan-Barcia, M. P., Francia, M.V., and De la Cruz, F. 2009. The diversity of conjugative relaxases and its application in plasmid classification. FEMS Microbiol Rev 33(3): 657-87.

Gomis-Ruth, F.X., and Coll, M. 2006. Cut and move: protein machinery for DNA processing in bacterial conjugation. Curr. Opin. Struct. Biol. 16(6): 744-752.

de Groot, M. J., Bundock, P., Hooykaas, P. J., and Beijersbergen, A. G. 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nat Biotechnol. 16(9): 839-42.

van der Graaff, E., den Dulk-Ras, Am, and Hooykaas, P.J.J. 1996. Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants. Plant Mol. Biol. 31: 677-681.

Guo, M, Hou, Q., Hew, C. L., and Pan, S. Q. 2007a. *Agrobacterium* VirD2-binding protein is involved in tumorigenesis and redundantly encoded in conjugative transfer gene clusters. Mol. Plant Microb. Interact. 20(10): 1201-1212.

Guo, M. Jin, S., Sun, D., Hew, C. L., and Pan, S. Q. 2007b. Recruitment of conjugative DNA transfer substrate to *Agrobacterium* type IV secretion apparatus. Proc. Natl. Acad. Sci. U.S.A. 104(50): 20019-20024.

Hamilton, H.L., Dominguez, N.M., Schwartz, K.J., Hackett, K.T., and Dillard, J.P. 2005. *Neisseria gonorrhoeae* secretes chromosomal DNA via a novel Type IV Secretion System. Mol. Microbiol. 55(6): 1704-1721.

Hawkey P.M., and Jones A.M. The changing epidemiology of resistance. 2009. J Antimicrob.

Chemother. (64) Suppl 1:i3-10.
Hoekema, A., Hirsch, P.R., Hooykaas, P.J.J., and Schilperoort, R.A. 1983. A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. Nature. 303: 179–180.
Hofreuter, D, Odenbreit, S., and Haas, R. 2001. Natural transformation competence in *Helicobacter pylori* is mediated by the basic components of a type IV secretion system. Mol. Microbiol. 41(2): 379-391.
Hooykaas-van Slogteren, G. M., Hooykaas, P. J., and Schilperoort, R. A. 1992. Expression of Ti plasmid genes in monocotyledonous plants infected with *Agrobacterium tumefaciens*. Biotechnology 24: 382-383.
Jasin, M. 1985. High frequency of homologous recombination in mammalian cells between endogenous and introduced SV40 genomes. Cell. 43(3): 695-703.
Juhas, M., Crook, D.W., and Wood, D.W. 2008. Type IV secretion systems: tools of bacterial horizontal gene transfer and virulence. Cell. Microbiol. 10(12): 2377-2386.
Jurica, M.S., and Stoddard, B.L. Homing endonucleases: structure, function and evolution. 1999. CMLS Cell. Mol. Life Sci. 1304-1326.
Khan, S.A. 2005. Plasmid rolling-circle replication: highlights of two decades of research. Plasmid. 53: 126-136.
Kim, S.I., Veena, and Gelvin, S.B. 2007. Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. The Plant J. 51: 779-791.
Kim, Y.G., Cha, J., and Chandrasegaran, S. 1996. Hybrid restriction enzymes: zinc finger fusions to *Fok*I cleavage domain. Proc. Natl. Acad. Sci. U.S.A. 93: 1156-1160.
Kononov, M.E., Bassuner, B., and Gelvin, S.B. 1997. Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. The Plant J. 11(5): 945-957.
Kostriken, R., Strathern, J.N., Klar, A.J., , Hicks, J.B., and Heffron, F. 1983. A site-specific endonuclease essential for mating-type switching in *Saccharomyces cerevisiae*. Cell. 35(1): 167-74.
Lacroix, B., Loyter, A., and Citovsky, V. 2008. Association of the *Agrobacterium* T-DNA-protein complex with plant nucleosomes. Proc. Natl. Acad. Sci. U.S.A. 104(40): 15429-15434.
Lacroix, B., and Citovsky, V. 2009. *Agrobacterium* aiming for the host chromatin. Comm. and Integrative Biol. 2(1): 42-45.
Lanka, E., and Wilkins, B. M. DNA processing reactions in bacterial conjugation. 1995. Annu. Rev. Biochem. 64: 141-169.
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.
Lee, Y. W., Jin, S., Sim, W. S., and Nester, E. W. 1996. The sensing of plant signal molecules by *Agrobacterium*: genetic evidence for direct recognition of phenolic inducers by the VirA protein. Gene. 179(1): 83-8.

Lee, Y. W., Jin, S., Sim, W. S., and Nester, E. W. 1995. Genetic evidence for direct sensing of phenolic compounds by the VirA protein of *Agrobacterium tumefaciens*. Proc. Natl. Acad. Sci. U.S.A. 92(26): 12245-12249.

Llosa, M., Gomis-Ruth, F.X., Coll, M., and de la Cruz, F. 2002. Bacterial conjugation: a two-step mechanism for DNA transport. Mol. Microbiol. 45(1): 1-8.

Lu, J., dan Dulk-Ras, A., Hooykaas, P. J. J., and Glover, M. J. *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. 2009. Proc. Natl. Acad. Sci. U.S.A. 106(24): 9643-9648.

Miranda, A., Janssen, G., Hodges, L., Peralta, E.G., and Ream, W. 1992. *Agrobacterium tumefaciens* transfers extremely long T-DNAs by a unidirectional mechanism. J. Bacteriol. 174(7): 2288-2297.

Otten, L., Piotrowiak, G., Hooykaas, P., Dubois, M., Szegedi, E., and Schell., J. 1984. Restoration of virulence of *vir* region mutants of *Agrobacterium tumefaciens* strain B6S3 by coinfection with normal and mutants *Agrobacterium* strains. Mol Gen. Genet. 195: 159-163.

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

Pacher, M., Schmidt-Puchta, W. and Puchta, H. 2007. Two unlinked double-strand breaks can induce reciprocal exchanges in plant genomes via homologous recombination and non-homologous end joining. Genetics. 174: 21-29.

Pansegrau, W., and Lanka, E. 1996. Mechanisms of initiation and termination reactions in conjugative DNA processing. J. Biol. Chem. 271(22): 13068-13076.

Peralta, E.G., Helmiss, R., and Ream, W. 1986. Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. The EMBO (Eur. Mol. Biol. Organ.) J. 5(6): 1137-1142.

Pietrokovski, S. 1994. Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins. Protein Sci. 3: 2340-2350.

Pitzschke, A, and Hirt, H. 2010. New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. EMBO (Eur. Mol. Biol. Organ.) J. e-pub ahead of print.

Podevin, N., De Buck, S., De Wilde, C., and Depicker, A. 2006. Insights into recognition of the T-DNA border repeats as termination sites for T-strand synthesis by *Agrobacterium tumefaciens*. Transgenic Res. 15: 557-571.

Puchta, H. 2002. Gene replacement by homologous recombination in plants. Plant Mol. Biol. 48: 173-182.

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.

Ramanathan, V., and Veluthambi, K. 1995. Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of T_L-DNA. Plant Mol. Biol. 28: 1149-1154.

Scheiffele, P, Pansegrau, W., and Lanka, E. 1995. Initiation of *Agrobacterium tumefaciens* T-DNA processing. Purified proteins VirD1 and VirD2 catalyze site- and strand-specific cleavage of

superhelical T-border DNA *in vitro*. J. Biol. Chem. 270(3): 1269-1276.
Schroder, G. and Lanka, E. 2005. The mating pair formation of conjugative plasmids – A versatile secretion machinery for transfer of proteins and DNA. Plasmid. 54(1): 1-25.
Segal, D.J., Breier, B, Beerli, R.R., and Barbas, C.F. 1999. Toward controlling gene expression at will: selection and design of zinc finger domains recognizing each of the 5'-CNN-3' DNA target sequences. Proc. Natl. Acad. Sci. 96: 2758-2763.
Shurvinton, C.E., Hodges, L., and Ream, W. 1992. A nuclear localization signal and the C-terminal omega sequence in the *Agrobacterium tumefaciens* VirD2 endonuclease are important for tumor formation. Proc. Natl. Acad. Sci. USA. 89(24):11837-11841.
Stoddard, B. L. 2005. Homing endonuclease structure and function. Q. Rev. Biophys. 38(1): 49-95.
Strathern, J. N., Klar, A. J., Abraham, J. A., Ivy, J. M., Nasmyth, K. A., and McGill, C. 1982. Homothallic switching of yeast mating type cassettes is initiated by a double-stranded cut in the MAT locus. Cell. 31(1): 183-192.
Stingl, K., Müller, S., Scheidgen-Kleyboldt, G., Clausen, M., and Maier, B. 2010. Composite system mediates two-step DNA uptake into *Helicobacter pylori*. Proc. Natl. Acad. Sci. U.S.A. 107(3): 1184-1189.
Steck, T., Close, T.J., and Kado, C.I. 1989. High levels of double-stranded transferred DNA (T-DNA) processing from a nopaline Ti plasmid. Proc. Natl. Acad. Sci. U.S.A. 86: 2133-2137.
Swanson, M.S., and Hammer, B.K. 2000. *Legionella pneumophilia* pathogenesis: a fateful journey from amoebae to macrophages. Annu. Rev. Microbiol. 54:567-613.
Toro, N., Datta, A., Yanovsky, M., and Nester, E. 1988. Role of the overdrive sequence in T-DNA border cleavage in *Agrobacterium*. Proc. Natl. Acad. Sci. U.S.A. 85(22): 8558-8562.
Toro, N., Datta, A., Carmi, O.A., Young, C., Prusti, R.K., and Nester, E.W. 1989. The *Agrobacterium tumefaciens virC1* gene product binds to overdrive, a T-DNA transfer enhancer. J. Bacteriol. 171(12): 6845-6849.
Tzfira, T., Frankman, L.R., Vaidya, M., and Citovsky, V. 2003. Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. Plant Physiol. 133: 1011-1023.
Tzfira, T., Li., J., Lacroix, B., and Citovsky, V. 2004. *Agrobacterium* T-DNA integration: molecules and models. TRENDS in genetics. 20(8): 375-383.
Tzfira, T., Valdya, M., and Citovsky, V. 2004. Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. Nature. 431: 87-92.
Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M., and Chilton, M. D. 1982. Short direct repeats flank the T-DNA on a nopaline Ti plasmid. Proc. Natl. Acad. Sci. U.S.A. 79(20): 6322-6326.
Van Attikum, H., Bundock, P., and Hooykaas, P.J.J. 2001. Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. EMBO (Eur. Mol. Biol. Organ.) J. 20(22): 6550-6558.
Van der Graaff, E., Den Dulk-Ras., A., and Hooykaas, P.J.J. 1996. Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants. Plant Mol. Biol. 31(3): 677-681.

Van Haaren, M.J.J., Sedee, N.J.A., Schilperoort, R.A., and Hooykaas, P.J.J. 1987. Overdrive is a T-region transfer enhancer which stimulates T-strand production in *Agrobacterium tumefaciens*. Nucleic Acids Res. 15(21): 8983-8997.

Vergunst, A. C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C. M., Regensburg-Tuïnk, T. J., and Hooykaas, P. J. 2000. VirB/D4-dependent protein translocation from *Agrobacterium tumefaciens* into plant cells. 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. 2005. Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. Proc. Natl. Acad. Sci. U.S.A. 102(3): 832-837.

White, C.E., and Winans, S.C. 2007. Cell-cell communication in the plant pathogen *Agrobacterium tumefaciens*. Phil. Trans. R. Soc. Lond. B. Biol. Sci. 362(1483): 1135-1148.

Windels, P., De Buck, S., Van Bockstaele, E., De Loose, M., and Depicker, A. 2003. T-DNA integration in *Arabidopsis* chromosomes. Presence and origin of filler DNA sequences. Plant Physiol. 133: 2061-2068.

Zaltsman, A., Krichevsky, A., Loyter, A., and Citovsky, V. 2010. *Agrobacterium* induces expression of a host F-box protein required for tumorigenicity. Cell Host & Microbe. 7(3): 197-209.

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

Ziemienowicz, A., Tinland, B., Bryant, J, Gloeckler, V., and Hohn, B. 2000. Plant enzymes but not *Agrobacterium tumefaciens* VirD2 mediate T-DNA ligation *in vitro*. Mol. Cell Biol. 20(17): 6317-6322.