Unexpected chromosomal DNA transfer from *Agrobacterium tumefaciens* to plant cells: mechanisms and solutions

Dissertation

zur Erlangung des Doktorgrades (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Tobias Berson

aus Engelskirchen

Bonn, Juni 2014

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn.

Erstgutachter:	Prof. Dr. Diedrik Menzel
Zweitgutachter:	Prof. Dr. Lukas Schreiber
Tag der Promotion:	10.10.2014
Erscheinungsjahr:	2014

Eidesstattliche Erklärung

Hiermit versichere ich, dass diese Dissertation von mir selbst und ohne unerlaubte Hilfe angefertigt wurde. Es wurden keine anderen als die angegebenen Hilfsmittel benutzt. Ferner erkläre ich, dass die vorliegende Arbeit an keiner anderen Universität als Dissertation eingereicht wurde.

Bonn, Juni 2014

Tobias Berson

Summary

The soil bacterium and plant pathogen *Agrobacterium tumefaciens* has the ability to transfer a defined part of its genome, the so called transferred DNA (T-DNA), to plant cells. This transfer has been successfully exploited in modern plant biotechnology and today researchers utilize the *Agrobacterium*-mediated plant transformation to gather important information about all aspects of plant biology. Furthermore, crop modification by *Agrobacterium*-mediated genome engineering is the fastest growing crop technology in the world and transgenic crops are cultivated in several countries with applications in food, feed and other industries.

It was recently discovered that, besides genes located on T-DNA, sometimes other large bacterial chromosomal DNA fragments (*A*chrDNAs) are unintentionally transferred from bacteria to plants by an unknown mechanism. This additional DNA transfer to plant cells added a new aspect to our understanding of horizontal gene transfer and genome evolution but also major complications to the generation and analysis of transgenic plants. Furthermore, the unnoticed transfer of large *A*chrDNA sequences to transgenic crops implicates important biosafety risks, when releasing transgenic plants to nature.

In this work, the underlying mechanisms behind the undesired transfer of *A*chrDNA to plants were investigated with the eventual goal to find ways to prevent them and thereby to improve the safety and reliability of *Agrobacterium*-mediated plant transformation.

In the course of this study, the *A. tumefaciens* transposable element IS*426* was found to integrate itself repeatedly into T-DNA vectors. IS*426* is one of the most frequently found *A*chrDNAs in plant cells but the way it is transferred is unknown. In this work it was shown that IS*426* cannot transfer to plant cells without the simultaneous transfer of a T-DNA. Additionally, its ability to control neighbouring gene expression was described.

The two chromosomal copies of IS*426* were sequentially deleted by homologous recombination in *Agrobacterium* strain A136. Thus, the first step towards a safer

plant transformation strain was made.

The transfer of *A*chrDNA to plant cells was successfully visualized by inserting the gene encoding the green fluorescent protein (*gfp*) into different regions of the bacterial genome. The tagged bacteria were used for plant transformation. GFP expressing plant cells could be observed, indicating that the *gfp* gene was transferred from the bacterial chromosomes to plants.

Using an *A. tumefaciens* mutant strain led to the finding that VirD2, one of the most important proteins involved in T-DNA transfer, is also involved in the transfer of *A*chrDNA to plant cells. Thus, it was hypothesized that VirD2 can bind to bacterial chromosomal regions and from there mediate the transfer of *A*chrDNA to plant cells.

By using bioinformatics, cloning and transient tobacco transformation assays, it was shown that a cryptic origin of transfer-like sequence (oriT-like), as well as a cryptic T-DNA border-like sequence (RB-like), were responsible for transferring *A*chrDNA to plant cells in a VirD2 dependent manner.

The deletion of one of these sequences from the genome of *A. tumefaciens* drastically reduced this transfer. Hence, it was shown for the first time that, apart from the Ti-plasmid borders, the *A. tumefaciens* genome contains additional sequences from which a transfer of DNA can be initiated.

The generated knowledge and material can now be used to locate the remaining chromosomal DNA transfer origins. Their collective deletion from the genome should allow the generation of safer *Agrobacterium* plant transformation strains.

Furthermore, this work should help to investigate whether these unexpected chromosomal DNA transfers are involved in the interaction/infection of *Agrobacterium* with other microorganisms or plants.

Contents

Abbreviations

*A*chrDNA	*Agrobacterium* chromosomal DNA	NOS	nopaline synthase
AmpR	ampicillin resistance gene	nt	nucleotide
bp	base pair	o/n	over night
CDS	coding sequence	orf	open reading frame
ChlR	chloramphenicol resistance gene	ori/oriR	origin of replication
DSB	double stranded break	oriT	origin of transfer
FST	flanking sequence tag	prom	promoter
GFP	green fluorescent protein	QS	quorum sensing
GOI	gene of interest	RB	right border
HR	homologous recombination	SpecR	spectinomycin resistance gene
IRL	inverted repeat left	sta	stability region
IRR	inverted repeat right	T4SS	type IV secretion system
kbp	kilo base pair	T4CP	type IV coupling protein
LB	left border	T-DNA	transfer DNA
Mb	mega base pair	term	terminator
MCS	multiple cloning site	T-strand	transfer strand
NHEJ	non homologous end joining	Ti	tumour inducing
NLS	nuclear localization signal	*vir*-genes	virulence genes
		Vir-proteins	virulence Proteins

1 Introduction

1.1 *Agrobacterium tumefaciens*, a plant bioengineer

Members of the genus *Agrobacterium* are gram negative soil bacteria with the unusual ability of interkingdom DNA transfer. By this transfer *Agrobacterium* is able to induce uncontrolled cell proliferation in plant cells. Thus, it is one of the main causes for the formation of plant tumours, the so called crown galls, which can be frequently observed in nature (Fig. 1.1 A). The transferred DNA (T-DNA) does not only induce tumour growth, but at the same time the plant cell is driven to produce certain metabolites, the so called opines. These serve as a nutrient source for the bacterium. The best known member of the *Agrobacterium* genus is *Agrobacterium tumefaciens* (*A. tumefaciens*) (Fig. 1.1 B), which causes the crown gall disease. *A. tumefaciens* contains a linear and a circular chromosome as well as two plasmids (pTi and pAt). Other members are *A. rhizogenes* which is the causative agent of the hairy root disease (Conn, 1942), *A. vitis* which stimulates tumour growth in grape vine (Ophel and Kerr, 1990), *A. rubi* which induces the cane gall disease (Starr and Weiss, 1943), *A. radiobacter* which is a non-pathogenic strain and *A. larrymoorei* isolated from *Ficus benjamina* (Bouzar and Jones, 2001; Farrand et al., 2003). However, the nomenclature is still debated since sometimes the only difference between the *Agrobacterium* species is the tumour inducing (Ti) plasmid, containing the T-DNA and encoding the proteins necessary for its transfer. By plasmid curing and re-transformation it is thus possible to convert one *Agrobacterium* species into another one (Costantino et al., 1980). Thus, the term "Biovar" is probably more appropriate (Gelvin, 2003). The bacterium's unique interkingdom DNA transfer ability was very successfully exploited by scientists, making *Agrobacterium* mediated plant transformation today the method of choice for genetic plant modification and thereby revolutionizing modern plant biology.

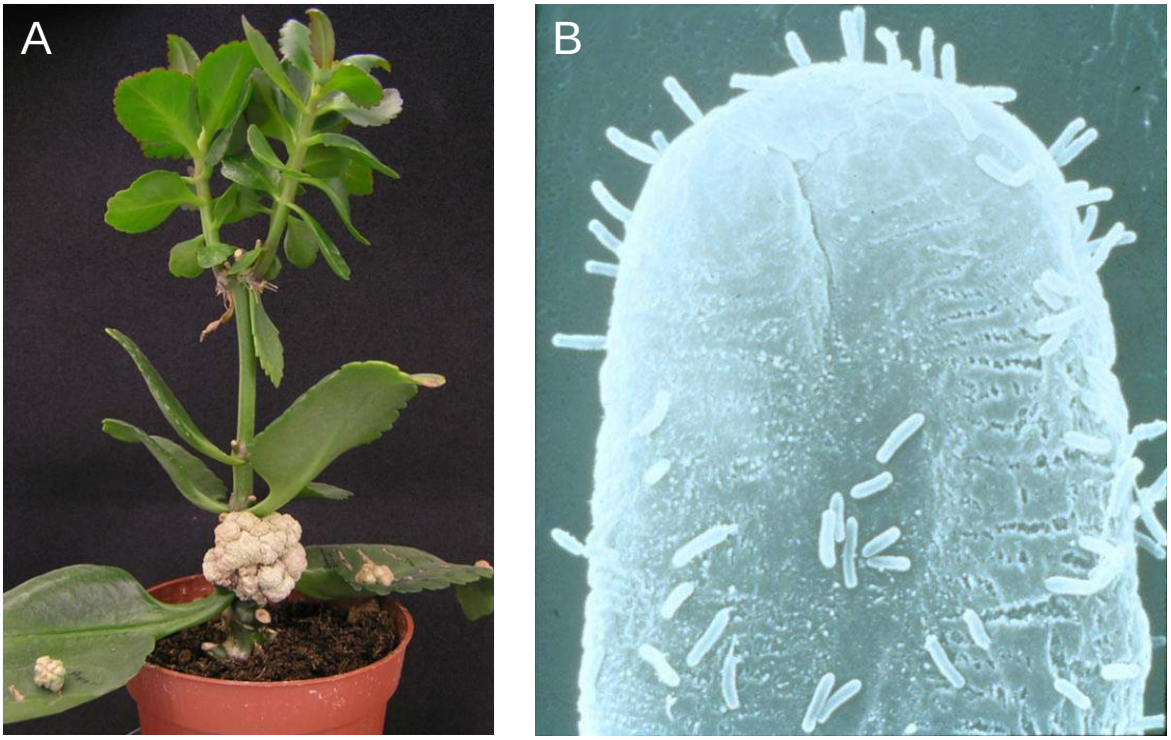


Figure 1.1: *A. tumefaciens* infects plant cells and induces tumour growth.
A: Tumour growth induced by *A. tumefaciens* on a *Kalanchoe* plant. **B**: Electron microscope picture of *A. tumefaciens* on a plant surface (Image B by Martha Hawes, University of Arizona. National Science Foundation).

1.1.1 History of *Agrobacterium*-mediated plant transformation

The first time that crown galls appeared in scientific literature was in the mid 19th century when scientists reported their formation in grapevine (*Vitis vinifera*). In 1907 Smith and Townsend showed for the first time that *A. tumefaciens* was the causative agent of the crown gall disease as they were able to reproduce the gall formation by inoculating marguerite daisies (*Bellis perennis*) with bacteria prior isolated from these plants. The authors proposed to name this bacterium "*Bacterium tumefaciens*" (Smith and Townsend, 1907). The name "*Agrobacterium*" was first proposed 35 years later (Conn, 1942).

In the following years scientists continued to investigate the tumour inducing ability of the bacterium in the hope to better understand the mechanism behind oncogenesis in mammals. In this time some remarkable observations were made. They found that even after removal of the bacteria, the plant cells still continued to proliferate and thus did not seem to need the continuous presence of bacteria to form galls (White

and Braun, 1942; Braun, 1958). Likewise, transformed plant tissue cultures were able to grow without addition of external growth hormones, which is normally necessary for persistent growth. It was also observed that transformed plant tissue did not only form tumours but also produced opines, uncommon amino acid - sugar compounds. Intriguingly, the type of opine produced by the plant cell depended on the *A. tumefaciens* strain used for inoculation (Petit et al., 1970). It was later proposed that the bacterium permanently activated the production of growth hormones in the plant cell by transferring either a metabolic product, a virus or a chemical agent like DNA (Braun, 1947). Braun termed the putative agent responsible for the tumorous growth the "Tumour inducing principle" (TIP).

In the 1970s many groups from all over the world tried to identify the nature of Braun's TIP. In these years the scientists observed that pathogenic Agrobacteria carried a large plasmid while non-pathogenic bacteria did not (van Larebeke et al., 1974). Furthermore, it was possible to switch the phenotype from non-pathogenic to pathogenic by introducing the large plasmid, making it obvious that this plasmid contains whatever is necessary for tumour induction (Kerr, 1971; Chilton et al., 1974). In 1977 it was the group of Eugene Nester which published that the by Braun proposed TIP is in fact DNA, which is transferred and stably incorporated into the plant cell genome. They also showed that not the complete plasmid is transferred to plants, but only a part of it. This part would later be called the T-DNA (Chilton et al., 1977).

After the discovery that *Agrobacterium* is able to transform plant cells, scientists tried to use the bacterium's unique ability to introduce new genes into plant cells. This was successfully achieved in 1983 by three independent groups (Bevan et al., 1983; Fraley et al., 1983; Herrera-Estrella et al., 1983).

1.1.2 The transfer of T-DNA to the plant cell

The transfer of T-DNA to plant cells is a complicated process with several bacterial and plant proteins involved (Fig. 1.2). The T-DNA transfer is best understood in *A. tumefaciens* but the fundamental mechanism is true for the other pathogenic bacteria of the *Agrobacterium* genus. They all posses a large plasmid which, depending on the species, is called the Ti- (tumour inducing) or Ri-(root inducing) plasmid. A smaller part of this plasmid, delimited by two "border" sequences, is called the transferred DNA (T-DNA). The T-DNA is mobilized and gets transferred as a single strand (T-strand) together with a subset of proteins through a pilus-structure to the plant cell. There, the T-strand, as a complex with bacterial and plant proteins, travels through the

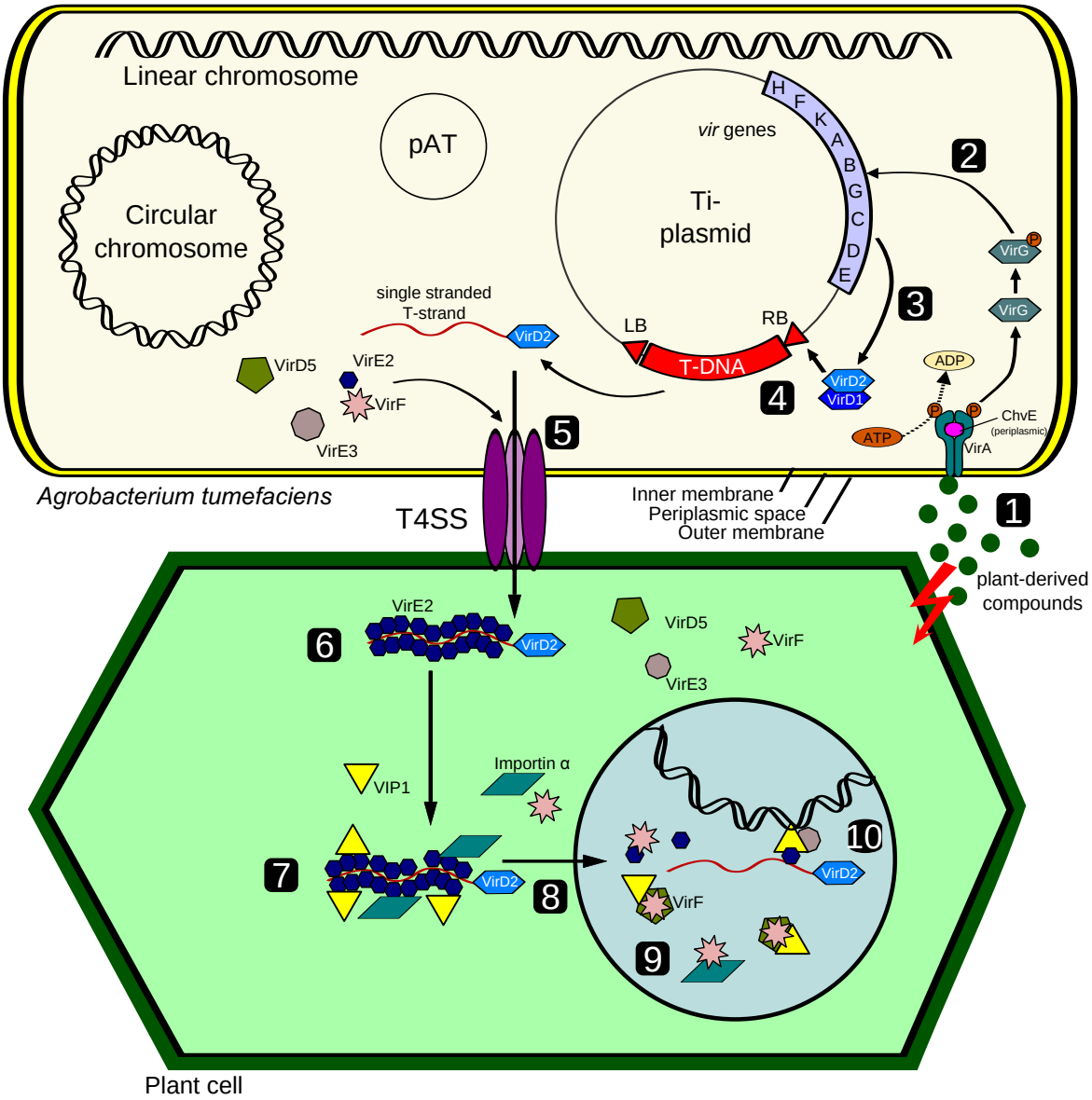


Figure 1.2: *A. tumefaciens* plant transformation process. **1**: Upon wounding, plant cell releases compounds which activate the VirA homodimer. ChvE binds to sugars released by the plant and also activates VirA (binding in the periplasm; not depicted in figure). Activated VirA phosphorylates VirG. **2**: Activated VirG binds to the *vir* boxes and triggers transcription of *vir*-genes. **4**: VirD1/VirD2 complex binds to RB and LB of the T-DNA and mediates mobilization of the single stranded T-strand. **5**: VirD2 and T-strand, together with VirE2, VirF, VirD5 and VirE3 are transferred to the plant via the T4SS. **6**: In the plant cell, VirE2 coats and protects the T-strand (formation of the T-complex). **7**: VirE2 interacts with VIP1 and Importin α and forms the T-super complex. VirE3 might substitute for a missing VIP1. **8**: T-strand is imported into the plant cell nucleus. **9**: T-strand is stripped of its coating proteins by VirF. VirD5 stabilizes VirF. **10**: T-strand is integrated into the plant genome. Genes located on the T-DNA are transcribed.

plant cytoplasm and translocates into the nucleus where the T-strand gets eventually integrated into the plant genome (reviewed in Gelvin, 2003; McCullen and Binns, 2006; Gelvin, 2012).

Initiation of T-DNA transfer

T-DNA transfer starts with a wounded plan cell. Upon wounding, plant cells release several compounds, among them are phenolic compounds such as acetosyringone and sugar molecules (Stachel et al., 1985, 1986; Ankenbauer and Nester, 1990). These molecules, together with an acidic pH, trigger a positive chemotaxis in the bacterium (Shaw, 1991) as well as the initiation of the T-DNA transfer. Before the T-DNA transfer can take place, bacteria attach to the plant cell surface, resulting in a biofilm formation (Danhorn and Fuqua, 2007). Chromosomal virulence genes such as *chvA*, *chvB* and *pscA* mediate the attachment (Douglas et al., 1982; Cangelosi et al., 1989; Marks et al., 1987; Altabe et al., 1990; Rudder et al., 2014).

Located on the Ti-plasmid is a region termed the virulence region (*vir*-region). With a few exceptions this region encodes all the proteins necessary for the transfer of T-DNA. The on the *vir*-region encoded genes are called *vir*-genes. The respective proteins are called Vir-proteins.

Among them is the constitutively expressed transmembrane protein VirA (Melchers et al., 1989; Jin et al., 1990). Together with the also constitutively expressed VirG, VirA forms a two-component sensory system in which VirA has the role of a periplasmic antenna (Melchers et al., 1989) while VirG is located in the cytoplasm and serves as a transcription factor able to trigger the transcription of the other *vir*-genes (Krishnamohan et al., 2001). VirA extends through the plasmamembrane of the bacterium with one end reaching into the periplasm and the other end in the cytoplasm. It is capable of sensing the phenolic compounds released by wounded plant cells (Lee et al., 1995). VirA is also able to interact with ChvE, one of the few chromosomally encoded proteins involved in T-DNA transfer. It binds to (plant released) sugars in the bacterium's periplasm (Gao et al., 2006). Sugar binding recruits the protein to VirA, thereby further activating it (Chang and Winans, 1992; Shimoda et al., 1993; Peng et al., 1998; He et al., 2009). Upon sensing, the VirA homodimer autophosphorylates (Brencic et al., 2004b) and in a second step the phosphate is transferred to VirG, the cytoplasmic response regulator (Jin et al., 1990; Winans, 1991). Phosphorylation activates VirG and the active VirG binds to a certain region in the promoters of the remaining *vir*-genes, the so called *vir*-boxes, thereby triggering their expres-

sion and thus activating the bacterium's virulence machinery (Pazour and Das, 1990; Scheeren-Groot et al., 1994; Krishnamohan et al., 2001; Gao et al., 2006).

The phenolic compounds released by the plant cell do not only activate the T-DNA transfer but are in fact toxic for the bacterium. The proteins encoded on the *virH* operon are activated by the compounds and are probably involved in their degradation (Brencic et al., 2004a; Joubert et al., 2004).

Release of the T-DNA

In *A. tumefaciens* strain C58 the T-DNA is a $\approx$ 25 kbp region located on the Ti-plasmid. It harbours all the genes necessary for tumour growth and opine production. These genes are controlled by promoters which are active in plants. The T-DNA is flanked by two 25 bp imperfect direct repeats, the so called right and left border (RB and LB) (Yadav et al., 1982; Wang et al., 1984). The two borders serve as recognition sequences for the proteins involved in release of the single stranded T-strand. They are the only *cis*-acting sequences that define the T-DNA. The main protein responsible for release of the T-strand is VirD2. VirD2 has a key role in transfer of the T-strand to the plant nucleus, since after release it stays covalently attached to the T-strand, until it is integrated into the plant genome (Herrera-Estrella et al., 1988; Ward and Barnes, 1988). Because of its important role it is also often referred to as the "pilot protein" of T-DNA transfer. Together with VirD1, VirD2 forms an endonuclease capable of introducing nicks in the 25 bp border repeats of one DNA strand, leading to the release of the single stranded T-strand (Jasper et al., 1994; Scheiffele et al., 1995). The VirD1/VirD2 endonuclease introduces a nick at the "lower strand" of the T-DNA's RB between nucleotides 3 and 4, accompanied by the covalently attachment of VirD2 to the 5' end of the T-strand by tyrosine[29] (Vogel and Das, 1992; Filichkin and Gelvin, 1993). The T-DNA's RB seems to have a more important role than the LB since VirD2 initiates the transfer at the RB and stays covalently attached to it. T-strand mobilization starts at the RB and is terminated at the LB, where the VirD1/VirD2 endonuclease introduces a second nick between nucleotides 3 and 4. Thus, a single stranded T-strand is completely released from the Ti-plasmid (Albright et al., 1987; Podevin et al., 2006). Because RB and LB are direct repeats with only a few nucleotides difference and in both sequences the nick is introduced at the same position, only 3 nucleotides of the RB and 22 nucleotides of the LB remain attached to the T-strand (Fig. 1.3).

Special sequences, called overdrive sequences, are often found close to many RBs

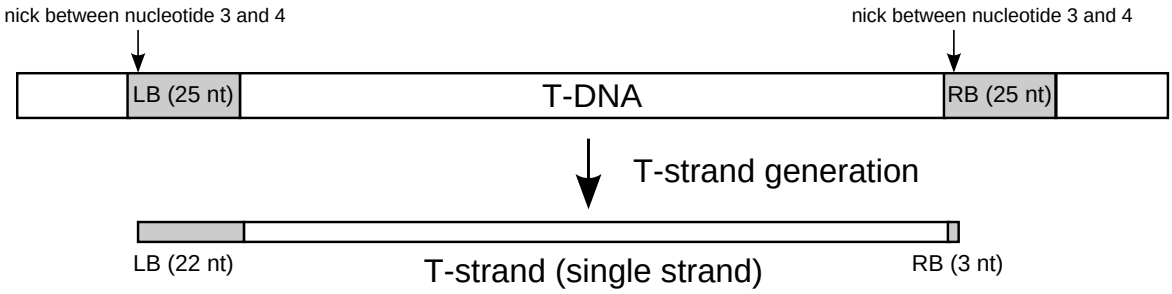


Figure 1.3: Generation of the T-strand. The VirD1/VirD2 complex introduces nicks between nucleotide 3 and 4 of each border. 22 nt of the LB and 3 nt of the RB remain attached to the T-strand.

but are absent from left borders (Toro et al., 1988; van Haaren et al., 1987). These sequences somehow enhance the T-DNA transfer to plant cells but the mechanism behind this enhanced transfer is not clear (Peralta et al., 1986). Likely the virulence proteins VirC1 and VirC2 are capable of binding to overdrive sequences and thus might facilitate the binding of the VirD2/VirD1 endonuclease complex at the RB (Toro et al., 1989; Lu et al., 2009).

Transport through the T4SS

The T-DNA strand with the attached VirD2 protein is transported to the plant cell via a so called type IV secretion systems (T4SS). Generally, T4SS systems are bacterial translocation channels which are able to transfer DNA and proteins across the cell envelope of bacteria. T4SS are involved in three different processes: The conjugal transfer of DNA to other bacteria, the delivery of effector proteins to eukaryotic cells and the release or uptake of DNA to or from an extracellular milieu (Alvarez-Martinez and Christie, 2009). The T4SS involved in conjugation constitute the largest group. The T-DNA transfer T4SS of *A. tumefaciens* consists of 11 different Vir proteins (VirB1 - VirB11) also called the Mating pair formation (Mpf) components and VirD4 (Fig. 1.4). In *A. tumefaciens*, VirD4 has the role of the type IV coupling protein (T4CP). T4CPs are conserved ATPases, interacting with both the substrates (T-DNA/VirD2 complex, VirE2) and the members of the T4SS, thereby recruiting the DNA/protein complex to the channel while energizing the process (Simone et al., 2001; Atmakuri et al., 2003; Vergunst et al., 2005).

In addition to VirD4, VirB4 and VirB11 are two additional ATPases energizing transfer of the substrate and assembly of the pilus (Atmakuri et al., 2004). VirB4 is a large conserved transmembrane protein with loops extending in both the cyto- and

the periplasm (Dang and Christie, 1997; Rabel et al., 2003) while VirB11 is located on the cytoplasmic side of the inner membrane (Rashkova et al., 1997).

VirB3, VirB6, VirB8 and VirB10 are located in the inner membrane, each having transmembrane domains and segments extending into the cyto- or periplasm. These subunits are forming the plasma membrane/periplasmic part of the T4SS having structural and scaffold functions. The precise function of VirB3 is still not known, but it is essential for substrate translocation (Berger and Christie, 1994). It also has been suggested to stabilize VirB4 and interact with VirB2 (Jones et al., 1994; Yuan et al., 2005).

VirB6 is a polytopic transmembrane protein with a large periplasmic domain that is probably involved in guiding the T-DNA complex through the pilus (Jakubowski et al., 2004). It has been shown to interact with the T-DNA substrate as well as VirB8, VirB10 and the VirB7/9 complex (Cascales and Christie, 2004; Jakubowski et al., 2004).

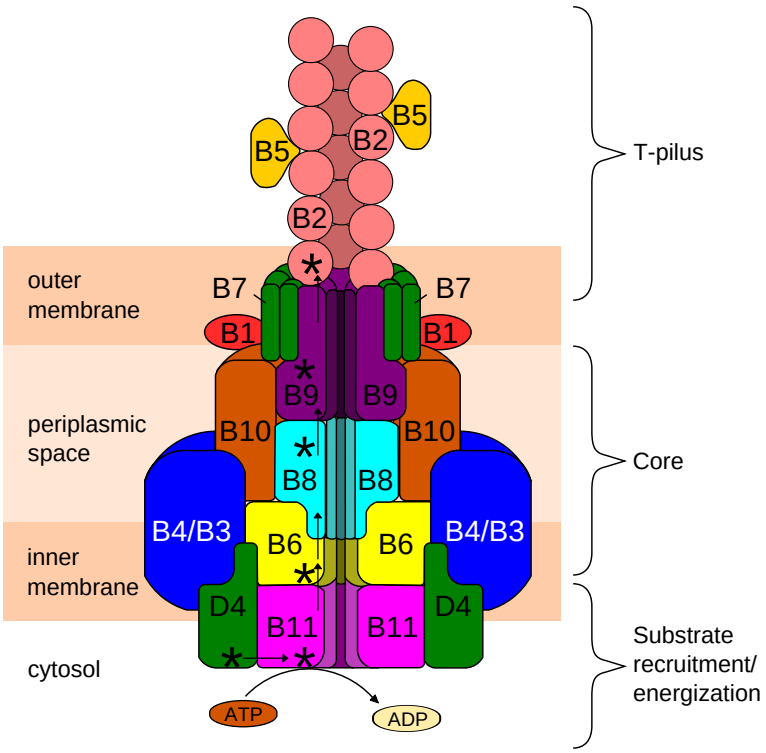


Figure 1.4: Generalized Type IV secretion system. Simplified locations and interactions between members of the T4SS are shown (see 1.1.2 for more details). Asterisk labelled proteins interact with the transported DNA substrate and arrows indicate the substrate's way through the complex. Image adapted from McCullen and Binns (2006).

VirB8 and VirB10 are bitopic proteins with a large periplasmic domain. Both

proteins interact with multiple other VirBs, probably giving these proteins a key role in nucleation, assembly and scaffold function (Kumar et al., 2000; Judd et al., 2005).

The proteins of the periplasm and outer membrane are VirB1, VirB2, VirB5, VirB7 and VirB9. VirB1 belongs to the family of transglycosidases and is probably involved in local degradation of the periplasmic peptidoglycan layer, typical for gram negative bacteria. Thus, it probably facilitates the assembly (Zahrl et al., 2005). They do not posses a transmembrane domain but are transported across the inner membrane into the periplasm. Although debated, VirB1 seems to be an essential factor for the assembly of the T-pilus (Höppner et al., 2004; Zupan et al., 2007).

VirB7 and VirB9 form a heterodimer complex in which VirB7 is a stabilizer for VirB9. The VirB7/9 complex is located in the outer bacterial membrane and forms a channel for the pilus (Anderson et al., 1996; Spudich et al., 1996; Baron et al., 1997). Together with VirB10, the VirB7/9 complex spans the entire cell envelope of the bacterium (Fronzes et al., 2009).

VirB2 is the main protein involved in formation of the pilus (Lai and Kado, 1998). These small proteins polymerize with each other, thus leading to the formation of the mature pilus (Lai et al., 2002). Interestingly, *A. thaliana* proteins have been identified which interact with VirB2. These BTIs (VirB2 interactors) termed plant proteins have been shown to be upregulated upon *A. tumefaciens* infection (Hwang and Gelvin, 2004).

VirB5 is exported from the bacterium and localizes extracellularly at the tip of the pilus (Aly and Baron, 2007). It was suggested that, besides VirB2, VirB5 is involved in attachment of the bacterium to the plant cell (Backert et al., 2008).

The DNA/protein substrate is guided through the channel by interacting with VirD4, VirB11, VirB6, VirB8, VirB9 and VirB2 (McCullen and Binns, 2006).

Cytoplasmic travelling and nuclear import

After crossing the T4SS the T-strand has to travel through the cytoplasm and enter the nucleus. It is assumed that once in the plant cell, the T-strand gets coated with hundreds of VirE2 proteins which were also transferred to the plant cell via the T4SS (Gelvin, 1998). These proteins likely protect the T-strand from degradation by plant nucleases. The VirE2 coated T-strand with VirD2 covalently bound to its 5' end is referred to as the T-complex.

The transfer of the T-complex through the cytoplasm to the nucleus is still not understood but some limited data suggest that the actin/myosin system is involved in

the transport (Zhu et al., 2003). However, microtubules might also play a role (Salman et al., 2005).

VirD2 as well as the VirE2 proteins carry a nuclear localization signal (NLS) which is important for nuclear targeting of the T-DNA. Interestingly, deletion of the C-terminal bipartite NLS of VirD2 does not significantly influence the transformation ability of the bacteria, speaking for a strong redundancy of the nuclear import system (Shurvinton et al., 1992; Mysore et al., 1998).

A. tumefaciens also "hijacks" several plant proteins to facilitate the nuclear import of the T-strand. Among the first ones identified was IMPa-1 (old: AtKAPα) which belongs to the importin-α family (9 members in *A. thaliana*). It was reported that IMPa-1 interacts with VirD2 and mediates its nuclear import and thus also the import of the T-strand (Ballas and Citovsky, 1997). Later, this was also confirmed for the remaining 9 members (Bakó et al., 2003; Bhattacharjee et al., 2008; Lee et al., 2008). Furthermore, it was shown that these proteins also interact with VirE2 (Citovsky et al., 2006; Bhattacharjee et al., 2008; Lee et al., 2008), speaking again for the redundancy of the system. IMPa-4 seems to have a special role in the nuclear import since a mutation of this protein led to a decreased *A. tumefaciens* transformation efficiency while mutations in the other importin-α proteins did not (Bhattacharjee et al., 2008). The to the plant cell exported bacterial protein VirE3 seems to interact with two members of this family. In addition to the Importin-α proteins, there seems to be at least one member of the Importin-β family which is involved in T-strand nuclear uptake (Zhu et al., 2003).

In a yeast two-hybrid screening assay for plant proteins interacting with VirE2, two additional proteins were identified and termed VIP1 and VIP2 (VirE2 interacting protein 1 & 2). These proteins did not show any interaction with VirD2 (Tzfira et al., 2001). In tobacco plants where VIP1 was upregulated, the transformation efficiency by *A. tumefaciens* increased, indicating an important role for the plant protein in nuclear T-DNA import (Tzfira et al., 2002).

Interestingly, the transferred protein VirE3 interacts with VirE2 in plant cells and seems to be able to account for a missing VIP1, since expression of VirE3 in *vip1* tobacco plants could restore transformation efficiency (Lacroix et al., 2005). Thus, *A. tumefaciens* seems to have developed a "backup" system in case of limited VIP1 in certain plant species. Taken together the data indicate that the T-strand import takes place in a complex with VirD2, VirE2, VIP1 and Importin-α proteins.

Phosphorylation of the T-complex proteins also seems to be important. It was

shown that VirD2 can be phosphorylated by cyclin-dependant kinase-activating kinases (CAK2Ms) (Bakó et al., 2003) and be dephosphorylated by a protein phosphatase 2C (PP2C) (Tao et al., 2004). Overexpression of PP2C led to a cytoplasmic localization of a VirD2/GUS fusion construct. The serine^{397} is likely the phosphate acceptor, since an exchange to alanine prevented nuclear import. Plants in which PP2C was mutated are hypersensitive to *A. tumefaciens* transformation (Tao et al., 2004). Phosphorylation of VIP1 also seems to be important for the nuclear T-DNA import. The MAP-kinase MPK3 is able to phosphorylate VIP1, thereby triggering the nuclear import of the protein (Djamei et al., 2007).

Another class of proteins which might play a role in T-complex nuclear targeting are the cyclophilins (Deng et al., 1998). However, the role of these proteins is not understood since deletion of the interacting domain of VirD2 does not reduce transformation efficiency (van Kregten et al., 2009).

Integration into the plant genome

Once in the nucleus, the T-strand targets the plant chromatin, a step which is likely not happening in transient transformation (Gelvin, 2010). Again, *A. tumefaciens* probably "hijacks" several plant proteins to conduct this step.

For a long time it was assumed that T-DNA integrates preferentially into transcriptionally active regions (Brunaud et al., 2002; Szabados et al., 2002; Schneeberger et al., 2005; Li et al., 2006). However, it was later shown that this observation probably resulted from a biased plant selection in these experiments (Francis and Spiker, 2005; Kim et al., 2007). The authors reported a completely random T-DNA distribution when considering plants without a prior selection for a transgene (Francis and Spiker, 2005). Thus, it is assumed that the T-complex is not directed to actively transcribed genes but more general to histones. Indeed, it was shown that VIP1 also interacts with plant histones and thus might be a key protein for directing the T-complex to the chromatin (Li et al., 2005a; Loyter et al., 2005; Lacroix et al., 2008).

Before integration, the T-strand has to be stripped of its bound proteins. The bacterial encoded protein VirF likely plays an important role during this process. VirF is an F-box protein and probably involved in tagging the proteins of the T-complex for proteolysis by the 26S proteasome (Regensburg-Tuïnk and Hooykaas, 1993; Tzfira et al., 2004; Lacroix et al., 2005). Similar to VIP1 and VirE3, it might be the case that *A. tumefaciens* uses a plant protein for T-complex protein degradation and, in case the host plant species does not encode such a protein, delivers its own version

as a backup. In the case of VirF, the plant protein which it substitutes might be VBF (VIP1-binding F-box protein) (Zaltsman et al., 2010). By expressing VBF with a T4SS secretion signal in a *virF A. tumefaciens* mutant strain, the authors were able to restore transformation efficiency in tomato. It was recently demonstrated that the bacterial virulence protein VirD5, which is also transferred to the plant cell via the T4SS, might serve as a stabilizer for VirF and prevents its degradation (Magori and Citovsky, 2011).

How the T-strand is integrated into the plant genome is the least understood part of the transformation process. Currently there are two main models how this integration could take place: the strand invasion and the double strand break repair model (Tzfira et al., 2004).

The strand invasion model postulates that the T-strand (with VirD2 attached) screens the plant genome for microhomology regions between T-strand and plant DNA. Once found, VirD2 introduces a nick in one strand of the plant DNA where the T-strand invades. After replication the second T-DNA strand is synthesized and the integration complete (Tinland and Hohn, 1995).

According to the double strand break repair model the T-strand is in a first step converted to a double strand and in a second step ligated into a double-strand break (DSB) in the plant DNA by non-homologous end joining (NHEJ). The observation that artificially induced DSBs in the plant genome are hotspots for T-strand integration supports this model (Salomon and Puchta, 1998; Chilton and Que, 2003; Tzfira et al., 2003).

It is not known which enzyme mediates the ligation of plant DNA with T-DNA. Initially the reported ligase activity of VirD2 was suspected to accomplish this (Pansegrau et al., 1993). However, further studies suggested that also a plant ligase is needed, but which one is still elusive (Ziemienowicz et al., 2000).

Ku80 is a protein involved in DSB repair. Li et al. (2005a) showed that a *ku80 A. thaliana* mutant shows a drastically decreased T-DNA integration rate while the transient transformation efficiency is not altered. Overexpression of Ku80 led to an increased stable transformation rate, pointing towards an important role of this protein in the last step of T-DNA integration (Li et al., 2005a).

VIP1 is not only involved in nuclear import but also might play a role in T-DNA integration. In an *A. thaliana vip1* mutant line where VIP1 was truncated but still able to mediate nuclear import, the stable transformation efficiency was decreased (Li et al., 2005b). Furthermore, the VirE2 interacting protein 2 (VIP2) seems to be also

involved in T-DNA integration, since *vip2* tobacco plants showed a reduced stable transformation efficiency (Anand et al., 2007). Since VIP2 is a transcription factor involved in the transcriptional regulation of histones, it is plausible that the for efficient T-DNA integration necessary histones are down-regulated in this mutant line.

Histones in general have an important role in T-DNA integration. Mutations in various *A. tumefaciens* histones (H2A, H2B, H3, H4) result in the so called rat (resistant to Agrobacterium transformation) phenotype (Zhu et al., 2003).

1.1.3 Quorum sensing in *A. tumefaciens*

A. tumefaciens posses a complex quorum-sensing (QS) cell to cell communication system which is directly involved in pathogenesis. QS allows bacteria to monitor the cell density of their population (Fuqua et al., 1994). *A. tumefaciens* exceeding a certain density triggers an increased Ti-plasmid replication by activating the plasmid's *repABC* operon and its conjugal transfer by triggering expression of the plasmid encoded *tra/trb* genes. Thus, in an environment where *A. tumefaciens* is able to induce crown gall formation, it "passes around" the Ti plasmid and hence gives the population a selective advantage. Populations where QS was impaired showed also less dramatic crown gall symptoms (Haudecoeur et al., 2009).

For quorum sensing the bacterium uses the TraR/TraI system which is similar to the well studied LuxI/LuxR system triggering bioluminescence in *Vibrio fischeri* (Nealson et al., 1970; Fuqua et al., 1994; Gray and Garey, 2001; Hagen et al., 2010). In the *Agrobacterium* system TraI mediates the synthesis of diffusible 8-oxo-octanoyl-homoserine lactone (OC8HSL) molecules (Zhang et al., 1993). OC8HSL accumulates in *Agrobacterium* populations in a density-dependent manner. If a certain threshold is exceeded, it can bind and activate TraR homo dimers (Luo and Farrand, 1999; Zhu and Winans, 1999). Activated TraR transcription factors bind to the tra box in promoter regions and thus activate genes responsible for Ti plasmid replication and conjugation (Zhang et al., 2002). Additionally, activated TraR also induces the transcription of more TraI, thus amplifying the process (Fuqua et al., 1994; Hwang et al., 1994). The negative feedback regulator TraM can bind to TraR and block its function (Hwang et al., 1994; Costa et al., 2012). This negative regulation prevents activation of the QS response before a certain cell density is reached. If enough OC8HSL is present, the negative regulation by TraM is not sufficient any more and the positive feedback loop is started.

Interestingly, the plant cell is also involved in the bacterium's QS system. Plant cells

transformed by *A. tumefaciens* produce opines as a nutrient source for the bacterium. Some special opines trigger the transcription of TraR, thereby activating the QS response (Habeeb et al., 1991). This regulation allows the bacteria to only disseminate the Ti-plasmid in an environment, where it is beneficial for the bacterium to have one because keeping a Ti plasmid is cost intensive (Platt et al., 2012).

The plant itself produces gamma-amino butyric acid (GABA). GABA can be transferred to *A. tumefaciens* by an ABC transporter encoded on the circular chromosome (Planamente et al., 2010). In the bacterium it triggers the enzymatic cleavage of the *N*-acyl-homoserine lactone (Chevrot et al., 2006). Thus, the plant is actively repressing the bacterium's QS response and decreases the number of pathogenic bacteria harbouring a Ti-plasmid.

1.2 *Agrobacterium* and plant biotechnology

As described in 1.1.1 the discovery that *A. tumefaciens* is able to incorporate DNA into the plant genome quickly led to its exploitation for molecular plant engineering (Bevan et al., 1983; Fraley et al., 1983; Herrera-Estrella et al., 1983).

The only *cis*-acting elements necessary for T-DNA transfer are the direct border repeats. Thus, the bacterium integrates virtually any DNA sequence placed between these borders into the plant genome. However, for easy and routine plant transformation, modifications in the system had to be made.

1.2.1 The binary vector system for plant transformation

In today's commonly used plant transformation vectors the T-DNA containing the genes necessary for tumour growth and opine production are deleted and often replaced by a multiple cloning site (MCS), making it easy to introduce a gene of interest (GOI) between the borders. These plasmids are referred to as "disarmed". Another obstacle which had to be overcome is the giant size of the Ti-plasmid. The wild type Ti-plasmid of *A. tumefaciens* strain C58 has a size of about 200 kbp. Even after dramatic reduction and limitation to only the sequences necessary for plant transformation, the plasmid still has a size of about 50 kbp (pCambia5105, personal communication Dr. Bekir Ülker). Additionally, the Ti-plasmid does only replicate in a low copy number in *E. coli*, making cloning of new genes in these plasmids very labour intensive and impractical (Zambryski et al., 1983; Fraley et al., 1985).

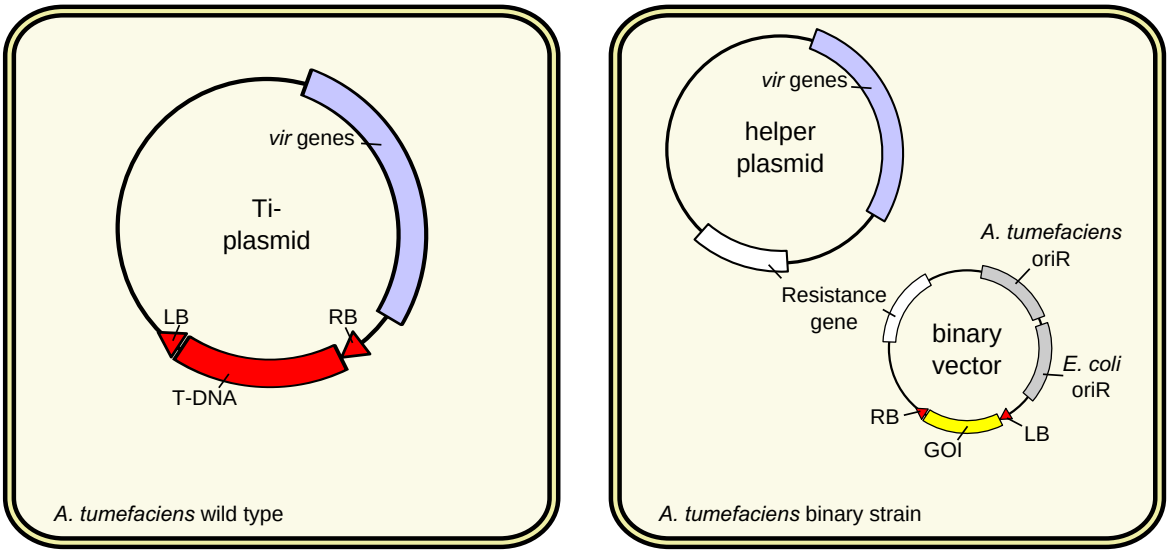


Figure 1.5: The binary vector system. Comparison between a wild type *Agrobacterium* and a binary strain used for plant transformation. In the binary strain, the Ti plasmid is split into two replicons, one containing the *vir*-genes and one the borders of the T-DNA flanking a gene of interest (GOI). RB, LB: left and right border; oriR: origin of replication.

This problem was overcome by the binary vector system (Hoekema et al., 1983). The fact that all the *vir*-genes can act in *trans* while the only necessary *cis*-acting elements are the border repeats led to the split of the *vir*-genes and the T-DNA on two separate replicons. A large one, the so called helper (Ti) plasmid, contains all the *vir*-genes, an origin of replication (oriR) for *A. tumefaciens* and a selectable marker. This plasmid is usually kept in the bacteria and not modified. A second plasmid, the so called binary vector, contains RB and LB often flanking a MCS, an oriR for *A. tumefaciens*, an oriR which allows high copy replication in *E. coli* and a bacterial selectable marker (Fig. 1.5). They also often contain a plant selectable marker gene for direct selection of transformed plants. In newer binary vectors the plant selectable marker is always next to the LB while the GOI is cloned next to the RB, since the 3' side of the T-DNA is often exposed to nucleolytic degradation in the plant, while the 5' side is protected by VirD2. This way no false positives, where the plant selectable marker is still intact but the GOI got degraded, can be obtained (Lee and Gelvin, 2008). This system made it easy for plant scientist to clone their GOI into the MCS of a small binary vector in an *E. coli* system. Successfully cloned vectors are isolated from *E. coli* in high copy numbers and used to transform *A. tumefaciens* harbouring the helper plasmid. Alternately, the plasmid can be mobilized into *Agrobacterium* by conjugation.

Over the years, many specialized binary vector systems were developed which offer the possibility for example for direct GUS/GFP fusions or gateway cloning (Hellens et al., 2000; Tzfira et al., 2005; Komori et al., 2007; Lee and Gelvin, 2008). Most of the common *A. tumefaciens* laboratory strains carrying a helper plasmid are based on the completely sequenced nopalin strain C58 (Wood et al., 2001). Care must be taken when choosing the appropriate strain/helper plasmid. Depending on the used binary vector, a strain must be chosen whose helper plasmid is not selected with the same selectable marker or is incompatible with its origin of replication. For example the strain GV3101:pMP90RK carries a kanamycin and gentamycin resistance gene on its helper plasmid and thus cannot be used together with binary vectors with a kanamycin resistance gene like for example pGreen or pBIN derivates (Bevan, 1984; Hellens et al., 2000).

1.2.2 *Agrobacterium* host range and transformation methods

A. tumefaciens has the broadest host range of any studied plant pathogenic bacteria (Pacurar et al., 2011) and has been listed as no. 3 in the top ten list of plant pathogenic bacteria (Mansfield et al., 2012). The ability of *A. tumefaciens* to infect certain species while others are resistant is not yet understood completely. A major factor determining the host range might be the Ti-plasmid (Loper and Kado, 1979; Thomashow et al., 1980). It was shown that VirC (Yanofsky et al., 1985; Yanofsky and Nester, 1986) and VirF (Melchers et al., 1990; Regensburg-Tuïnk and Hooykaas, 1993) proteins have an influence on the host range and VirH seems to be important for *A. tumefaciens* to transform maize (Jarchow et al., 1991). With the development of modified *A. tumefaciens* strains and new transformation protocols, scientists are now able to transform many dicot and monocot angiosperms, gymnosperms (Gelvin, 2003), fungi including yeast (Bundock et al., 1995; Bundock and Hooykaas, 1996) and even human cells (Kunik et al., 2001). Of particular interest for crop biotechnology is the possibility to transform maize (Ishida et al., 1996), rice (Hiei et al., 1994; Shri et al., 2013), barley (Tingay et al., 1997) and wheat (Cheng et al., 1997). New protocols and strains are published on a monthly basis and continuously expand the bacterial host range.

There are two main ways how *A. tumefaciens* mediated plant transformation in biotechnology is achieved: the regeneration dependent or independent way. In the regeneration dependent methods tissue cultures, leaf-discs, callus cultures, protoplasts are co-cultivated with Agrobacteria. This is followed by plant regeneration from the

transformed tissue. For many species this method is the only way to obtain transgenic plants. However, plant regeneration protocols do not exist for many plants.

The regeneration independent protocols are generally easier to conduct, since no contamination prone tissue culture is involved. The easiest regeneration independent method is probably the floral-dip transformation of *A. thaliana* (Clough and Bent, 1998; Logemann et al., 2006). In this method the inflorescence of *A. thaliana* plants is simply dipped in an *A. tumefaciens* cell suspension. Sometimes *Agrobacterium* transforms the megaspore (Ye et al., 1999; Bechtold et al., 2000; Desfeux et al., 2000). Seeds derived from the transformed ovaries result in stably transformed plants, which can be selected for. This method is frequently used in many laboratories for obtaining transformed *A. thaliana* plants. However, this method can lead to multiple and complex T-DNA insertion patterns which in turn can lead to silencing of the transferred genes (Hobbs et al., 1993; Jorgensen et al., 1996; Meyer and Saedler, 1996).

An earlier regeneration independent method is the vacuum infiltration (Bechtold et al., 1993). In this method complete plants are submerged in a bacteria solution and a vacuum is applied, leading to the uptake of the bacteria by the plant. Similar to the floral dip method, seeds derived from transformed ovaries can be harvested and used to grow transgenic plants.

The leaf infiltration method is usually not used to obtain stably transformed plants but for transient assays where the T-DNA is not integrated into the genome. The method is mostly used for expressing reporter proteins such as GFP, for a limited time. To achieve this, an *A. tumefaciens* solution is infiltrated into plant leaves (for example of *N. benthamiana*) with the help of a syringe. The leaf cells are efficiently transformed by the bacteria. However, the T-DNA is not always integrated into the genome of the plant (stable transformation) but probably kept in the nucleus without integration (transient transformation) (Gelvin, 2010).

1.2.3 The role of *Agrobacterium* in today's green biotechnology

In the year 1983 *A. tumefaciens* mediated plant transformation with a GOI was successfully applied for the first time (Bevan et al., 1983; Fraley et al., 1983; Herrera-Estrella et al., 1983). Since then it became the method of choice for plant transformation and today it is used routinely in countless laboratories all over the world and has tremendously increased our knowledge of the function of plant genes/proteins and plant biology. The applications are numerous. For example it can be used for overexpression of plant genes for reverse genetic approaches, protein fusions to reporter genes

such as GUS, GFP or luciferase for localization studies, plant promoter to reporter gene fusions for expression studies or for plant promoter trapping experiments with promoter-less reporter genes (Springer, 2000; Alonso and Ecker, 2006).

Another important application are the T-DNA insertion lines. Scientists exploit the fact that the T-DNA integrates randomly in the plant genome and thereby disrupts genes upon integration. This leads to a knock out mutant for this particular gene (Krysan, 1999). By transforming a large number of plants with *A. tumefaciens*, theoretically every gene of the plant genome is expected at some point to be disrupted by a T-DNA insertion (*de facto* this is not completely true because for reasons not yet understood, T-DNA does not integrate in some regions of the plant genome). To find out where the T-DNA integrated and which plant gene got disrupted, flanking sequence tags (FSTs) are generated by sequencing the flanking regions to the left or right of the inserted T-DNA (Ortega et al., 2002; Strizhov et al., 2003). These FSTs are stored in databases. The large scale T-DNA insertional muatagenesis made it possible that scientists can access *A. thaliana* mutants of nearly every gene for reverse genetic approaches to understand gene function. (McElver et al., 2001; Samson et al., 2002; Sessions et al., 2002; Alonso et al., 2003; Rosso et al., 2003).

Furthermore, the *A. tumefaciens* mediated plant transformation can be used to introduce new genes into the plant genome and thus equip the plant with new traits. This is used for the generation of genetically modified crop plants. The introduced traits can for example be resistance against certain herbicides for monocropping, increased resistance against herbivores, increasing the biomass (Petersen et al., 2012), decreased sensitivity to environmental conditions, delaying fruit ripening or drive the plant to produce additional chemical compounds such as vitamins (golden rice)(Ye et al., 2000; Paine et al., 2005).

The method has also been applied for medical purposes by introducing the ability to produce medically relevant compounds such as vaccines or antibodies in plants (Rodgers et al., 1999; Arntzen et al., 2005). However, so far no products are commercially available yet but attempts to produce them continue (Ma et al., 2005; Thomas et al., 2011).

1.2.4 Unintended DNA transfer from *Agrobacterium* to plants

Even though today *Agrobacterium*-mediated plant transformation is one of the most used techniques for generation of stably transformed plants, it has a few drawbacks. As mentioned above, it frequently happens that more than one T-DNA copy is inserted in

the plant genome. This phenomenon is often observed when transforming *A. thaliana* plants by the floral-dip method and the insertion of multiple copies can cause silencing of the transcripts (Tang et al., 2007). Also the truncation of one or more of the T-DNA copies is a frequently observed problem, since it might lead to the expression of truncated proteins and to misinterpretation of phenotypes. It was reported that the chance of inserting multiple copies can be decreased by launching the T-DNA from the bacterial chromosome instead of a plasmid (Oltmanns et al., 2010).

Furthermore, caution should be exercised when designing the T-DNA region. It was shown that the use of strong promoters on the T-DNA, for example to drive expression of a selectable marker gene, might result in the unintended transcription of flanking plant genes after integration of the T-DNA into the plant genome. This could lead to misleading phenotypes (Ülker et al., 2008b).

Another frequently observed phenomenon is the transfer of plasmid backbone (Martineau et al., 1994). This can be caused by an event called border-skipping (Kononov et al., 1997). Thereby the T-DNA is normally cleaved by the VirD1/VirD2 nuclease at the RB but for some reason the LB is skipped. This leads to generation of a long T-strand (maximum the length of the binary vector) with the vector backbone sequence at the 3' end. This might be troublesome, since for instance the unintended expression of antibiotic resistance genes in plants might lead to false phenotypes. This problem can also be overcome by launching the T-DNA from the chromosome. However, integrating the T-DNA into the chromosome in the first place is laboursome.

Sometimes the LB can be recognized as an RB from the T-strand initiating proteins (Ramanathan and Veluthambi, 1995). This leads to a transfer of the vector backbone instead of the T-DNA. Another important problem when transforming plants with the *A. tumefaciens* mediated transformation was observed by Ülker et al. (2008a). The authors made the observation that not only vector backbone is an unwanted DNA which sometimes ends up in plant cells but *A. tumefaciens* can also transfer parts of its chromosome to plant cells. The origin of the *Agrobacterium* chromosomal DNA (*A*chrDNA) transfer is somewhat random, making it difficult to explain this phenomenon.

1.3 *Agrobacterium* transfers chromosomal DNA to plants

When Ülker and colleagues characterized a T-DNA insertion locus in the genome of *A. thaliana*, they discovered a short DNA fragment of bacterial origin in the recovered FST. It turned out that the fragment originated from the linear chromosome of *A. tumefaciens*. This observation led them to further investigate, if the transfer of *Agrobacterium* chromosomal DNA (*A*chrDNA) is a common trait of *A. tumefaciens* mediated plant transformation. This was confirmed in 2008 in their publication "T-DNA-mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants" (Ülker et al., 2008a).

1.3.1 *A*chrDNA was found in T-DNA insertion libraries

In a large scale approach the authors analysed FSTs obtained from over 375,000 *A. thaliana* T-DNA insertion lines from different databases (FLAGdb, SAIL, GABI-Kat, SIGnAL). They found that all of these databases contained examples where *A*chrDNA was inserted into the plant genome next to the T-DNA. The transfer was irrespective of the *A. tumefaciens* strain used for transformation or the binary vector system harbouring the T-DNA. Additionally, by analysing a rice FST database it was concluded that the unwanted *A*chrDNA transfer is not limited to *A. thaliana* plants, but does also happen in rice. The transferred chromosomal fragments could be large (> 18 kbp) and contained up to 18 complete bacterial open reading frames. Interestingly, in the plant genome the *A*chrDNA was more often associated with the T-DNA's RB than with the LB (Fig. 1.6). Taken together, Ülker et al. (2008a) concluded that approximately 1 out of 250 ($\approx 0.4\,\%$) via *Agrobacterium* transformed transgenic *A. thaliana* plants contains *A*chrDNA.

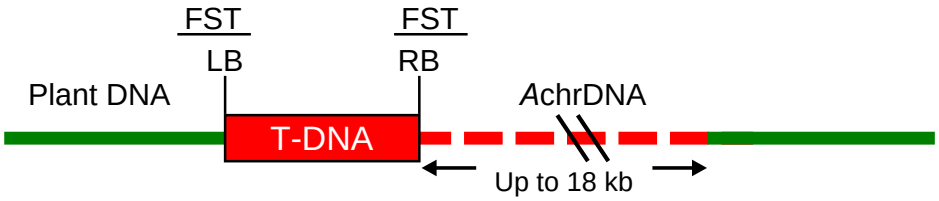


Figure 1.6: *A*chrDNA was found next to the T-DNA in plant cells. Representative scheme of plant genome (green) containing an inserted T-DNA flanked by *A*chrDNA (red). FSTs were used to screen for *A*chrDNA fragments.

1.3.2 *A*chrDNA hotspots were frequently detected in plant cells

When mapping the sequence of the *A*chrDNA fragments to the linear chromosome of *A. tumefaciens*, it became obvious that the transfer is not entirely random. The authors were able to locate certain hotspots on the linear chromosome which were more often transferred to the plant cell then other sequences. These hotspots were similarly distributed in all four FST collections (Ülker et al., 2008a). On the first look, an area of several thousand bps approximately around 1.4 Mb of the linear chromosome seems to be the most frequently transferred region (Fig. 1.7). Additionally, a transposable element called IS*426* is often transferred to plants together with T-DNA. The first quarter of the linear chromosome does not contain any hotspots of *A*chrDNA transfer.

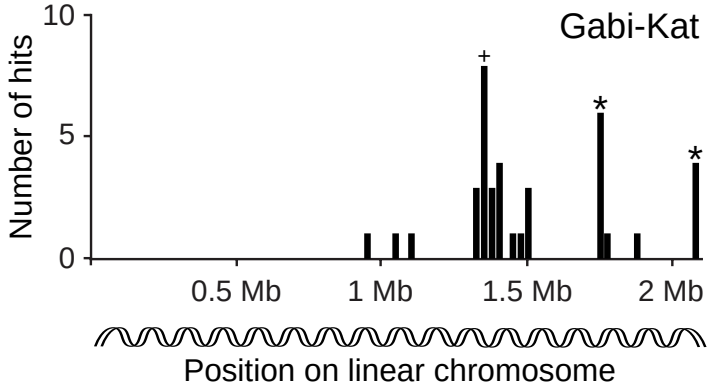


Figure 1.7: Hotspots of *A*chrDNA on the linear chromosome in GABI-Kat lines. The most frequently transferred region is around 1.4 Mb (+). Both copies of IS*426* were also often found in GABI-Kat lines (*). The first quarter of the linear chromosome does not contain any hotspots of *A*chrDNA transfer. Taken from Ülker et al. (2008a).

The right border-like hotspot

In the frequently transferred 1.4 Mb hotspot region, the authors detected a sequence which shows homology to a T-DNA right border sequence. Therefore they hypothesized that this sequence might be involved in the transfer of *A*chrDNA. In this scenario the VirD1/VirD2 complex would bind to this sequence on the chromosome and mediate *A*chrDNA transfer to plant cells. It even contains a sequence which resembles an overdrive (overdrive sequences are frequently associated with RBs; see 1.1.2). Because of the homology they called it the RB-like sequence (in this study referred to as RB-like1).

Bacterial transposable element IS*426* was frequently found in plants

The bacterial transposable element IS*426* (formally known as IS*136*) was frequently found in the genome of *A. thaliana* T-DNA insertion lines. Ülker et al. (2008a) showed that in three independent lines an incomplete IS*426* was always flanked by two T-DNAs. Interestingly, the right terminal inverted repeat (IRR) was always next to the truncated RB of a T-DNA. The other flanking T-DNA still contained a complete RB.

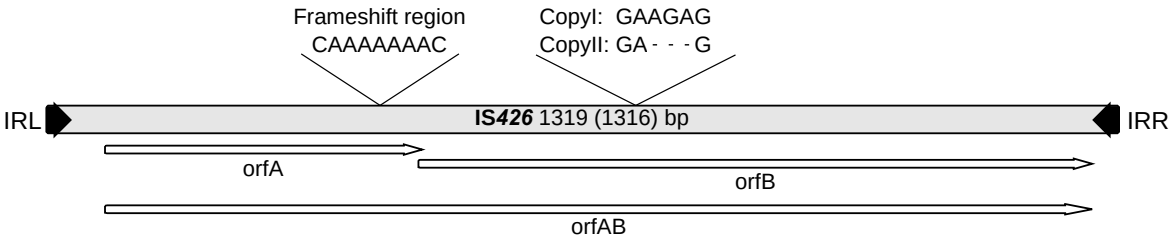


Figure 1.8: Schematic drawing of IS*426*. The insertion sequence is delimited by two inverted repeats (IRL and IRR) and contains two open reading frames, orfA and orfB. Because of a frameshift region, the combined open reading frame orfAB can be translated. The linear chromosome contains two IS*426* copies, which differ in an additional base pair triplet in copy-I.

IS*426* is a member of the large IS*3* subfamily and is closely related to the well characterized IS*911* and IS*2* transposable elements (Mahillon and Chandler, 1998; Siguier et al., 2006). Two copies of IS*426* are found on the linear chromosome of *A. tumefaciens* strain C58 with the only difference of three additional nucleotides in copy-I (see Fig. 1.8). A third copy is located on the At plasmid. However, this copy contains many mutations and lacks large parts of its 3' end, including the IRR. Like IS*2*, IS*426* is delimited by two imperfect inverted repeats, the IRR (inverted repeat right) and IRL (inverted repeat left) with a weak promoter located next to the IRL (Lewis et al., 2004). The repeats are flanking two -1 frameshifted overlapping open reading frames (orfA and orfB) (Lewis et al., 2011) which together encode the transposase. OrfA of IS*426* contains an A_7 motif which, like the A_6G motif in IS*2*, constitutes a frameshift region which can lead to the synthesis of the complete functional transposase (OrfAB) by translational frameshifting (Lewis and Grindley, 1997).

IS*2*, IS*426* and other members of the IS*3* as well as other IS families replicate via a two step copy and paste mechanism (Duval-Valentin et al., 2004). In the first step the weak promoter at the IRL drives the basal expression of the transposase. The transposase binds to IRL and IRR and mediates via an intermediate branched figure-of-eight structure (Lewis and Grindley, 1997; Polard and Chandler, 1995) the formation of a closed double-stranded minicircle, while leaving an original copy in the

chromosome (copy and paste). The minicircle formation of IS*426* was reported by Alexandra Stirnberg in her master's thesis in the PME-group (Stirnberg, 2011). In the second step, the minicircle integrates into another position on the chromosome. In the minicircle, IRL and IRR are adjacent and together form the strong promoter P_{junc}, which is able to drive high expression of the transposase, necessary for the insertion of the minicircle into the target region (Ton-Hoang et al., 1997; Lewis et al., 2004).

IS*426* constitutes an active insertion sequence in *A. tumefaciens*, which was reported to frequently insert into bacterial chromosomes, plasmids and binary vectors. Thereby it is disrupting genes or is unintentionally transferred to plant cells (Vanderleyden et al., 1986; Luo and Farrand, 1999; Ülker et al., 2008a; Llop et al., 2009; Rawat et al., 2009).

1.3.3 Possible explanations for the transfer of *A*chrDNA to plant cells

In their study, Ülker et al. (2008a) only observed *A*chrDNA flanking the T-DNA. It is tempting to speculate that the *A*chrDNA transfer process is linked to the T-DNA transfer. However, due to the fact that they analysed FSTs, this observation might be misleading, since this way it is not possible to detect *A*chrDNA fragments, which integrated spatially separated from the T-DNA (unlinked). Ülker et al. (2008a) and the author of the corresponding commentary article (Gelvin, 2008) speculated several hypotheses about the mechanism behind the transfer of *A*chrDNA to plants.

Insertion of T-strands into bacterial chromosomal DNA and re-launching together with flanking *A*chrDNA

Ülker et al. (2008a) hypothesized that the *A*chrDNA transfer might be a two step process. In a first step the mobilized T-strand, instead of being transported through the T4SS to plant cells, is first integrated into the bacterial chromosome. In a second step, the T-DNA, together with adjacent bacterial chromosomal DNA, is re-launched from the bacterial chromosome and transported to the plant nucleus.

As explained in 1.1.2, during T-DNA processing only 3 nucleotides of the RB remain attached to the T-DNA while 22 remain of the LB. This means the RB is almost completely lost during procession (Wang et al., 1984). If a T-DNA integrates into the bacterial chromosome, the second remobilization has to be mediated by other means, since no RB is available for the VirD1/VirD2 complex to bind to.

The authors hypothesized that chromosomal sequences, such as the RB-like sequence mentioned above (1.3.2) could act as substitute borders for the missing RB. It has been shown that the sequence of the RB can be degenerated to a certain degree and still serves as a recognition sequence for the VirD1/VirD2 complex to initiate DNA transfer (Rommens et al., 2005). Thus, it would be possible that chromosomal sequences might also serve as "substitute borders". In such a case the transfer would start from the RB-like sequence in the bacterial chromosome, continue with adjacent bacterial chromosomal DNA, followed by the T-DNA sequence and be terminated at the T-DNA's almost intact LB. The new chimeric T-strand would then be transferred to the plant cell and integrate into the genome. In this scenario, T-DNA and *A*chrDNA would always be linked in the plant genome. If this is the way *A*chrDNA gets transferred to plant cells, it has to be explained why T-strands should get integrated into the bacterial chromosome.

The integration could potentially be mediated by either illegitimate or homologous recombination. An illegitimate recombination could be facilitated by nicks or DSBs in the chromosomal DNA, introduced for example by transposon jumping, stress induced DNA damage or replication.

Alternatively, it might be the case that homologous regions in the processed T-strand and the chromosome cause integration of the strand into the bacterial chromosome. It would also be imaginable that instead of the T-strand being completely incorporated, its 3' end might get ligated to a free 5' end of chromosomal DNA which was prior induced by a nick or a DSB. The T-DNA would be normally transferred to the plant cell and "pull" *A*chrDNA with it.

In a different scenario, homologous regions in the binary vector and the bacterial chromosome might lead to homologous recombination. Such a recombination would lead to an integration of the complete vector into the bacterial chromosome. In this case the relaunch could be mediated by the T-DNA's RB. By LB border skipping *A*chrDNA could become part of the T-strand.

Insertion of bacterial chromosomal DNA into the T-DNA vector

In another hypothesis parts of chromosomal DNA could be mobilized and integrate themselves into the binary vector. The mobilization of chromosomal DNA could for example be mediated through a transposase, a cryptic origin of transfer (oriT) or an RB-like sequence. The so mobilized chromosomal sequences could then get integrated into the binary vector. Any sequence inserting between the RB and LB would become

part of the T-DNA and would get transferred to the plant cell nucleus in a "normal" T-DNA transfer process. For example, the insertion of a transposable element into the T-DNA sequence would constitute a simple explanation for an *A*chrDNA transfer to plant cells. In *A. tumefaciens* this was shown for IS*426* (Rawat et al., 2009) and for the transposon Tn*5393* of *A. tumefaciens* strain LBA4404 (Kim and An, 2012).

The insertion of a transposable element into the backbone sequence of the binary vector in turn could generate large homology regions, leading to the integration of the complete vector into the chromosome. As described above, T-DNA launching from such an inserted vector might result in the transfer of *A*chrDNA by skipping of the LB.

Linkage of independently processed T-strands and *A*chrDNAs

In another scenario the linkage of T-DNA and *A*chrDNA might not take place in the bacterium but in the plant cell. This would imply that *A*chrDNA can get mobilized and transferred to plant cells independently. Such an independent mobilization of *A*chrDNA could be mediated by chromosomal border-like sequences, which are recognized by the VirD1/VirD2 complex.

The two DNA strands (T-strand and *A*chrDNA) would then be transported to the plant nucleus independently from each other. Both strands could integrate into the same locus in the plant genome or spatially separated from each other. It has been shown that if a plant is transformed with two independent *Agrobacterium* strains carrying individual T-DNAs, both are frequently found next to each other in the plant genome (De Neve et al., 1997). If this is the way that *A*chrDNA gets transferred, it would mean that Ülker et al. only detected the cases where *A*chrDNA and T-DNA integrated next to each other because analysis of FSTs would not allow detection of an unlinked *A*chrDNA somewhere else in the genome. It would also mean that the estimation that one in 250 transgenic plants contains *A*chrDNA is an underestimation and the transfer occurs more often. If *A*chrDNA can be independently mobilized and transferred to the plant cell nucleus by the VirD1/VirD2 complex, it could also get transferred to the plant cell without any binary vector/T-DNA present in the bacterium.

Conjugation mediated transfer of *A*chrDNA to plant cells

Apart from the virulence associated T4SS, the VirD4/VirB-system, *A. tumefaciens* carries two more type IV related secretion systems. The Trb system is also located

on the Ti-plasmid but, in contrast to the VirD4/VirB-system, is involved in the conjugative transfer of the Ti plasmid between two bacteria (Li et al., 1998). The AvhB system is located on the AT-plasmid and mediates its conjugal transfer between two bacteria (Chen et al., 2002). Similar to the VirB-system, both involve proteins for substrate processing (Relaxases, Dtr proteins), substrate recruitment (coupling protein, T4CP) and for transmembrane channel formation (Mpf subunits). In addition to these systems, *A. tumefaciens* might carry a cryptic 4th secretion channel on its linear chromosome. However, so far it is not clear if it is an active system (Leloup et al., 2002).

All T4SS are ancestrally related and it was shown in an early study that the *mob* and oriT sequences from the broad host range plasmid RSF1010 are able to mediate the transfer of plasmids from *Agrobacterium* into plant cells (Buchanan-Wollaston et al., 1987). Thus, instead of the VirD1/VirD2 complex mediating the mobilization starting from RB-like sequence, it might be components of the Trb or AvhB system which mediate the mobilization of chromosomal DNA. In this scenario, not RB-like sequences would act as starting sequences but oriT-like sequences and the transfer could be mediated through any of the type IV secretion channels. In the plant cell the virulence proteins (VirE2/VirD2) and host cell factors might associate with the "conjugated" strand and mediate its integration into the nucleus.

Because of the close relationship between virulence and conjugation systems it might be possible that the VirD1/VirD2 complex binds to (cryptic) oriT sequences (oriT-like) and starting from there, mediate mobilization of chromosomal DNA to plant cells. The cleavage of oriT sequences by VirD2 has already been demonstrated (Pansegrau et al., 1993; Grove et al., 2013). Furthermore, Dube et al. (2004) showed that an oriT sequence on a plasmid was responsible for its VirD2 dependent transfer to plant cells. Thus, not only chromosomal border-like sequences might serve as a starting point for the VirD1/VirD2 mediate transfer of *A*chrDNA but also chromosomal oriT-like sequence might be recognized.

1.3.4 Implications of the *A*chrDNA transfer to plant cells

The unwanted transfer of additional bacterial chromosomal DNA during T-DNA transformation constitutes a severe problem. The origin of the transferred *A*chrDNA is somewhat random and can contain many bacterial open reading frames. Therefore it is difficult to make assumptions if the transferred genes are transcribed in plant cells and what risks they might generate. If bacterial genes are expressed in plant cells,

the plant might show a misleading phenotype. For example, if an *A. thaliana* T-DNA insertion line expresses the proteins encoded by the bacterial genes, it might lead to false assumptions about the respective knocked out gene.

Since the T-DNA insertion locus in the plant genome during routine scientific research is often not determined, it can be expected that at least one in 250 transgenic plants can potentially lead to false results. If the integration of *A*chrDNA and T-DNA happens unlinked from each other, this number could be even higher. Furthermore, the integration of *A*chrDNA unlinked to T-DNA could result in the unintended and unnoticed knockout of a plant gene by *A*chrDNA, again resulting in misleading phenotypes. Another problematic scenario is that genetically modified plants (GM plants) released to nature might contain *A*chrDNA. Expression of the genes located on the *A*chrDNA might influence the environment or even lead to health risks in animals and humans, for example by causing allergies.

In general, the unintended transfer of *A*chrDNA during plant transformation constitutes an obstacle in the generation of clean transformants. Because the mechanism behind this transfer is elusive, at the moment cost intensive extensive sequencing of the transformed plants is the only way to rule out a transfer of *A*chrDNA.

1.4 Aims of this study

The transfer of DNA from chromosomes of *Agrobacterium* (*A*chrDNA) to plant cells constitutes a new and so far uncharacterised aspect of *A. tumefaciens* mediated plant transformation. The unintentional transfer of large bacterial chromosomal DNA fragments is worrisome and adds complications to the generation and analysis of transgenic plants. Furthermore, the unnoticed transfer of large *A*chrDNA sequences to transgenic crops constitutes a potential biosafety risk when plants are released to nature.

In this thesis, the mechanism behind the transfer of *A*chrDNA to plant cells should be explored with the final goal to generate an *A. tumefaciens* transformation strain with a minimized chance of co-transferred *A*chrDNA fragments to plants.

The *A*chrDNA transfer could be a consequence of a hypothetical alternative procession of the T-DNA. Instead of being directly transported to the plant cell, it gets occasionally integrated into the bacterium's own chromosome. A re-mobilization, together with flanking chromosomal DNA, could result in the transfer of *A*chrDNA to plant cells. Initial experiments should clarify, if a T-strand integration into the bacterium's own chromosome can actually occur. If this is the case, in the next step it should be determined if the re-launch of such an integrated T-DNA from the bacterial chromosome can lead to the transfer of *A*chrDNA to plant cells.

Furthermore, the possibility of a T-DNA independent transfer of *A*chrDNA should be explored. For this, reporter genes should be inserted into specific regions of the *A. tumefaciens* genome by homologous recombination. Strains carrying these insertions, but no T-DNA, should be subsequently used to transform plants. Expression of reporter proteins would indicate that a DNA transfer from the bacterial chromosome to the plant cell took place. If this is the case, the proteins involved in the transfer process should be determined by using *A. tumefaciens* mutant strains.

Additionally, it should be assessed if bacterial chromosomal sequences can act as starting points for the transfer of DNA to plant cells. For this reason, candidate sequences will be PCR amplified and inserted into an *A. tumefaciens* vector. This vector should harbour a reporter gene but not any sequences, able to mediate DNA transfer to plants (no T-DNA borders). The insertion of a sequence with this ability should result in transfer of the vector to plant cells and in expression of the reporter.

If such chromosomal sequences can be identified, an *A. tumefaciens* strain in which these sequences are deleted from the genome should be generated, resulting in a plant transformation strain with a decreased *A*chrDNA transfer potential.

2 Material

2.1 Bacterial strains

Molecular cloning was performed in *E. coli* using either 5-alpha or Turbo competent cells. Different *A. tumefaciens* strains were used for plant transformation.

Bacterial strain	**Genotype**	**Reference**
A. tumefaciens C58	Wildetype (stock no. 5172)	Wood et al. (2001)
A. tumefaciens GV3101(pMP90)	based on C58 Rif^R, Gent^R; ΔT-DNA	Koncz and Schell (1986)
A. tumefaciens A136	based on C58; Rif^R ΔpTiC58	Watson et al. (1975)
A. tumefaciens AT*ΔvirD2*	Rif^R, Carb^R based on pMP6000(ΔT-DNA, Δ*virD2*)	Bravo-Angel et al. (1998)
NEB Turbo competent *E. coli*	F' *proA*$^+$*B*$^+$ *lacI*^{q} Δ*lacZM15/fhuA2* Δ(*lac-proAB*) *glnV galK16* *galE15 R(zgb-210::Tn10)*Tet^S *endA1 thi-1* Δ*(hsdS-mcrB)5*	New England Biolabs
NEB 5-alpha competent *E. coli*	*fhuA2*Δ(*argF-lacZ)U169 phoA* *glnV44* Φ80 Δ*lacZM15 gyrA96* *recA1 relA1 endA1 thi-1* *hsdR17*	New England Biolabs

Table 2.1: Bacterial strains used in this study. R: resistance

2.2 Plant material

Nicothiana benthamiana plants were used for transient transformation assays. *Arabidopsis thaliana* (ecotype Columbia-0) plants were used for stable transformations.

2.3 DNA

2.3.1 Vectors

The following table contains the underlying vectors which were used to generate all vectors in this study. Desired sequences were either excised by restriction digestion or amplified by PCR reactions.

Name	Sequences used	Reference
pBucc120	SpecR, ChlR, *pat*, pVS1-ori/sta	Dr. Ülker, University of Bonn
pCATGFP	*gfp*, Amp-promoter	Group Prof. Menzel, University of Bonn
pCR™-Blunt II-TOPO®	*nptII*, *bla*, pUC-ori	Invitrogen
pCR®2.1-TOPO	pUC-ori, MCS	Invitrogen

Table 2.2: Vectors used in this study.

2.3.2 Primers

Table 2.3 shows the primers used in this study. If primers were used to amplify fragments designated for molecular cloning, primer sequence contains respective restriction sites.

Name	Purpose	Sequence (5'→3')
116	*nptII*, *bla* and pUC ori for pIPTmKanR (for)	NNNAAGCTTGGATAACGCAGGAAAGAAC
134	*nptII*, *bla* and pUC ori for pIPTmKanR (rev)	NNNAAGCTTTCGCATGATTGAACAAGATGGA
174	Screening for IS*426*-Ko (for)	ACCGGACAAGTCGGCTAGATT
175	ChlR for pIPTmKanR (for)	GCAACTAGTTTCTCGAGATTAGGCACCCCAGGCTTTACACTT
176	ChlR for pIPTmKanR (rev)	GCTCCCGGGAATCTAGAACTGTCACTGTAATACGCTGCTTCA
272	pUC-ori and MCS for pBasic (for)	GTCACGCGTTCCTTTTTGATAATCTCATGACCA

273	pUC-ori and MCS for pBasic (rev)	GTCCCATGGGTTTTCCCAGTCACGACGTT
274	SpecR for pBasic (for)	GTCGTCGACATGTTATGGAGCAGCAACGA
275	SpecR for pBasic (rev)	GTCAGATCTACGCTCAGTGGAACGAAAAC
276	RB for pBasic (for)	GTCGAGCTCGGCATGCACATACAAATGGA
277	RB for pBasic (rev)	GTCACTAGTATCGCCCTTCCCAACAGT
286	Screening for IS*426* in pIPTmKanR (rev)	TGGCATCGTAAAGAACATTTTG
287	Screening for recombination in pIPTmKanR (rev)	CATATTGGCCACGTTTAAATCA
340	Probe forIS*426*-KO DNA-blot (for)	CGACCGCCTCCAACATCATGTC
341	Probe forIS*426*-KO DNA-blot (rev)	TGAACTGCCCCCCATTTCGAC
316	LB for pBasic (for)	AGTCCCATGGGCATTAATGAATCGGCCAAC
317	LB for pBasic (rev)	AGTCGTCGACACTGATGGGCTGCCTGTATC
318	pVS1-ori for pBasicS1 (for)	GATCACGCGTGCGCCCTGGTAGATTG
319	pVS1-ori for pBasicS1 (rev)	GATCAGATCTGCACTTGAGCGCAGCGAGG
397	Screening for IS*426* in pIPTmKanR (for)	AGGTCGGTCTTGACAAAAAGAAC
402	Screening for IS*426* in pIPTmKanR (for)	ATGCTATCGAGCATCTCTCTGAC
450	RB-like1 for pBasicS1-RB-like1 (for)	GATCACTAGTATTACCGTCACCGATCTTGC
451	RB-like1 for pBasicS1-RB-like1 (rev)	TATCGAGCTCAAACTCCATGCCTTCATGCT
492	RB-like1 for pBasicS1-RB-like1 (for)	GATCGTCGACATTACCGTCACCGATCTTGC
493	RB-like1 for pBasicS1-RB-like1 (rev)	GATCCCATGGAAACTCCATGCCTTCATGCT
466	IRR for pBasicS1-GFP-IS (for)	TAATGAGCTCAGGAAGTCATGTGCCGAGATT

467	IRR for pBasicS1-GFP-IS (rev)	GATCACTAGTGATCGAGGACTACAACGAAATCC
468	IRL for pBasicS1-GFP-IS (for)	GATACCATGGGTCAGCTTTTGCTCGGTTGT
469	IRL for pBasicS1-GFP-IS (rev)	GATCGTCGACAATTCCACCAGCGTTTGTTC
468	IRL for pBasicS1-GFP-IS (for)	GATACCATGGGTCAGCTTTTGCTCGGTTGT
469	IRL for pBasicS1-GFP-IS (rev)	GATCGTCGACAATTCCACCAGCGTTTGTTC
518	*pat* for pBasicS1 (for)	TATCTCTAGACAAACCGAAGGCGGGAAACG
519	*pat* for pBasicS1 (rev)	GATCCCATGGTTAATAACACATTGCGGACG
522	dIRL for pBasicS1-GFP-IS-dIRL (for)	GATACCATGGAATTCCACCAGCGTTTGTTC
523	dIRL for pBasicS1-GFP-IS-dIRL (rev)	GATAGTCGACGTCAGCTTTTGCTCGGTTGT
566	IS*426* for pBasicS1GFP-cIS (for)	ATGTAAGCTTATTCCACCAGCGTTTGTTCG
567	IS*426* for pBasicS1GFP-cIS (rev)	TATAGAGCTCAAGTCATGTGCCGAGATTGG
573	$HS1_{LC}$ homology region (for)	GCTCGAGCTCAAAGCAGAATAGCGCATCT
574	$HS1_{LC}$ homology region (rev)	TATTGAGCTCCGACAGGAACTGCGCAG
581	$NE1_{LC}$ homology region (for)	ATATGAGCTCGTCATTCCATCTCTCCTATAGC
582	$NE1_{LC}$ homology region (rev)	ATATAAGCTTGAGGTGATACTCGCTGCT
591	Screening for IS*426* in pIPTmKanR (rev)	GTTTACCCGCCAATATATCCTGTCA
626	IS*426* for promoter test (for)	CGCGCACTAGTTCTCGTCATGTGAACTGCCC
627	IS*426* for promoter test (rev)	TATACCCGGGACCTCCTGGAGTGCACCCCATTTCAC
675	Amp promoter for	TATAACTAGTCCTTGCCATTGCCGGGATCG

	promoter test (for)	
676	Amp promoter for promoter test (rev)	GCGACCCGGGACTCTTCCTTTTTCAATATT
693	prom.less *nptII* for promoter test (for)	TATACCCGGGATGATTGAACAAGATGGATT
694	prom.less *nptII* for promoter test (rev)	GCTCCTCGAGGTTCTTCTGAATTGAAAAA
755	$NE2_{LC}$ homology region (for)	TATAGAGCTCCATTCACCACCAGTTCGTTG
756	$NE2_{LC}$ homology region (rev)	TATAGAGCTCATGTCGAAATCCGCAATCTC
794	5 kbp for p5kbHS1$_{LC}$GFP (for)	GATCACTAGTTCGATGAGATAACCACGCCG
795	5 kbp for p5kbHS1$_{LC}$GFP (rev)	TCTAGAGCTCTTCAGCGTGATCGGCTTCTT
816	ABC-pro (410 bp) (for)	GCGCACTAGTAGCGTAATTCACCCAGGACGT
817	ABC-pro (410 bp) (rev)	TATACCCGGGACAATTGATCCGGTTGATG
818	ABC-pro (1852 bp) (rev)	TATACCCGGGTGCAAAGTACCGGTGGGAAA
812	IS*426* for pBasicS1GFP complete-IS (for)	TATAGTCGACCGCCACTTTTGCCTTGAAAG
813	IS*426* for pBasicS1GFP complete-IS (rev)	GATACCATGGCCTAGATTGATTTAGCCCTGA
814	NOS term for pBasicS1 GFP-complete-IS (for)	GATCAAGCTTACTAGTTGACCCCTAGAGTCAAGCAG
815	NOS term for pBasicS1 GFP-complete-IS (rev)	GTCTTCTAGATCCCGGGGAATACTAACGTCTCTAC
869	$HS2_{LC}$ homology region (for)	TAATGAGCTCGATTTTTCGGGCTCCGGCAT
870	$HS2_{LC}$ homology region (rev)	GCGTGAGCTCGATTTTCACCTGGTTTCAAA
871	$HS1_{CC}$ homology region (for)	GAGTGAGCTCGTCTGTTCTCCTCCAGTTAG
872	$HS1_{CC}$ homology region (rev)	ATTAGAGCTCCATGATGATCCCTTTATCCC
886	*atsa*-promoter (for)	GCGCACTAGTCGGAACCTCATTTCCTGCCT

887	*atsa*-promoter (rev)	TATACCCGGGAGTCACCTCACGAATTGCGG
896b	5 kbp for p5kbHS2$_{LC}$GFP (for)	GCGCACTAGTAAAACCACGAGAACGACCAC
897	5 kbp for p5kbHS2$_{LC}$GFP (rev)	GCCAGAGCTCAAAATCCACCATCCTCAACG
898	5 kbp for p5kbNE2$_{LC}$GFP (for)	TCGAGAGCTCATTGCCTTGCCTATTCCACC
899	5 kbp for p5kbNE2$_{LC}$GFP (rev)	GGCGAAGCTTTTGACGATGAAGAAGGTGCG
900	5 kbp for p5kbHS1$_{CC}$GFP (for)	TTAAGAGCTCGACGCCTGTTTTGGTGTTTT
901	5 kbp for p5kbHS1$_{CC}$GFP (rev)	TGACAAGCTTACGACGGATTTAACCTCACG
919	IS*426* copyI-KO upstreamHR (for)	TACTGGTACCATGGACAGGACGATCAACGT
920	IS*426* copyI-KO upstreamHR (rev)	TATCGGATCCTCATGGTATCTCCTGTTCCC
921	IS*426* copyI-KO downstreamHR (for)	TATAGTCGACTTCACATGACGAGACGACCG
922	IS*426* copyI-KO downstreamHR (rev)	TTAAGGATCCAATCAGAAAACGTCCGATGG
949	IS*426* copyII-KO upstreamHR (for)	AGAGAGATCTTTGGCAACATTCTGCGGCGC
950	IS*426* copyII-KO upstreamHR (rev)	TATAGGCGCGCCATTTATGCTGCACACGATCT
957	oriT-like1(3,268 bp) (for)	TATAAAGCTTATCGCAAAGACAAGCGAACT
959	oriT-like1(200 bp) (rev)	GCATAAGCTTCATAGGGGGTGGAATCGTAA
962	oriT-like1(3,268 bp) (rev)	ATTAGAGCTCAATCTCTTCCCGAACGTC
963	oriT-like1(200 bp) (for)	TATAGAGCTCCTCCATCCCATCGCATCTT
974	IS*426* copyII-KO downstreamHR (for)	CATATCTAGAGGTCTGATTGGACAGCCAGT
975	IS*426* copyII-KO downstreamHR (rev)	CATAGTCGACTCCAGCCAAGAACATAACAT
976	oriT-like1-KO upstreamHR (for)	GCTCTCTAGATCATGAAGGATCATGGCTGC

977	oriT-like1-KO upstreamHR (rev)	TATTGTCGACTACCATTCGCCAGAATGACC
978	oriT-like1-KO downstreamHR (for)	CGCGCAGATCTTATTTATCGCCGCGCTTTCG
979	oriT-like1-KO downstreamHR (rev)	TAATGGCGCGCCTAGTCCACATCACTCGTTGC
985	oriT-like1-KO screening (for)	ATGGCATCCTGTGGATATGG
986	oriT-like1-KO screening (rev)	ATTGAACCAAGGTTACGGCG
989	oriT-like1(31 bp) (for)	AGCTTTCACTGATGACGACCTATGAATACATCCTGCTGATCAGAGCT
990	oriT-like1(31 bp) (rev)	CTGATCAGCAGGATGTATTCATAGGTCGTCATCAGTGAA
991	oriT-like1(61 bp) (for)	TGTCAAGCTTCATCACCATCTATCAGGCGCTTTCACTGATGACGACC
992	oriT-like1(61 bp) (rev)	GCGGAGCTCTGATCAGCAGGATGTATTCATAGGTCGTCATCAGTGAAAG
1027	oriT-like4(261 bp) (for)	NNNNAAGCTTATTGTCTTCGCCCAATTGTT
1028	oriT-like4(605 bp) (for)	NNNNAAGCTTGGGTGACTGGATTCTTCTTGA
1029	oriT-like4(605 bp) (rev)	NNNNGAGCTCGATGACCTCACCTTGACGGTA
1030	oriT-like4(261 bp) (rev)	NNNNGAGCTCAAGACATCACCCAGCGTGTTC
1033	RB-like2(221 bp) (for)	NNNNAAGCTTAAATCCTCACATGCATCCCTTA
1034	RB-like2(221 bp) (rev)	NNNNGAGCTCGCCCTTCTTATAGAGGGTGAGG
1035	LB-like1(198 bp) (for)	NNNNAAGCTTCGGCAGCTTTTCCAGATAATAA
1036	LB-like1(198 bp) (rev)	NNNNGAGCTCCTGCTCGAAGTGCTGTTCAT
1037	052H10-locus (2,199 bp) (for)	NNNNAAGCTTCAGTGGAAAAACGCCAAGAT
1038	052H10-locus (2,199 bp) (rev)	NNNNGAGCTCCTTTACGATCTGACGAAGATG
1042	RB-like2(31 bp) (for)	AGCTACCACTTCGACCGGCAATATATCCAGCCTCT
1043	RB-like2(31 bp) (rev)	CTAGAGAGGCTGGATATATTGCCGGTCGAAGTGGT
1044	230H11-locus (2,344 bp) (for)	NNNNAAGCTTTATACCTCGACCATCGTTTTCCT
1045	230H11-locus (2,344 bp) (rev)	NNNNGAGCTCTAAAGCAAGCCCCATTTGATATT
1048	oriT-like3(551 bp) (for)	NNNNAAGCTTTTCCTGATATCACCATGCTCACT
1049	oriT-like3(551 bp) (rev)	NNNNGAGCTCGCCAGTTGGCGACTACTACTTTT
1054	oriT-like2(318 bp) (for)	NNNNAAGCTTAGTCGGAACGTTTGAATTCCT

1055	oriT-like2(318 bp) (rev)	NNNNGAGCTCACCCTGTAACCCCCATGAATA

Table 2.3: Primers used in this study. Primers which were used in the same PCR reaction are not separated by a line.

2.3.3 DNA size markers

Markers used to determine the size of DNA fragments on agarose gels. A biotinylated version of the NEB 2-Log DNA size marker was used for DNA blot analysis.

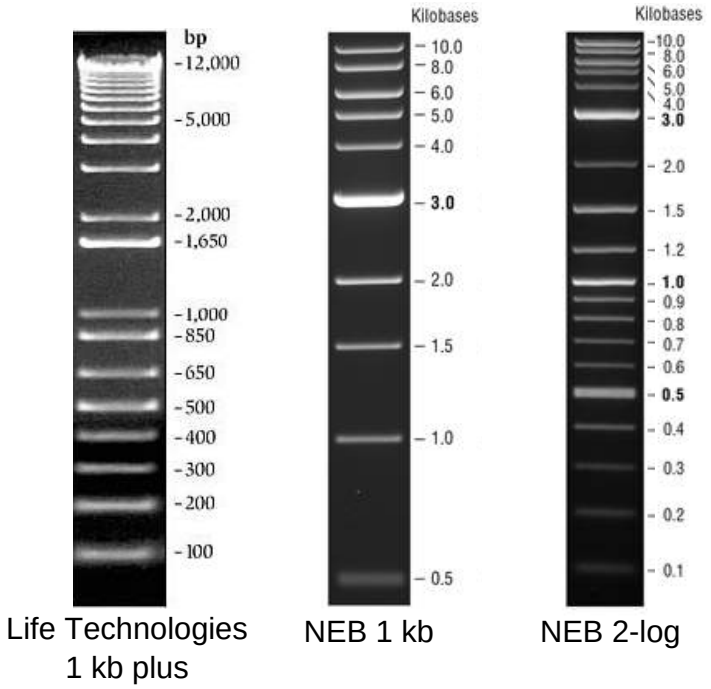


Figure 2.1: DNA size markers used in this study.

2.3.4 Probes for DNA blot analysis

- External probe for confirmation of correct vector integration in section 4.2.4 was generated with primers 753 and 754 using *A. tumefaciens* GV3101(pMP90) genomic DNA as template.
- Internal *gfp*-probe for confirmation of correct vector integration in section 4.2.4 was generated with primers 751 + 752 using pCATGFP as template.
- Probe used to confirm the deletion of IS*426* in section 4.1.1 was generated with primers 340 + 341 and *A. tumefaciens* A136 genomic DNA as template.

2.3.5 Sequencing

DNA sequencing was performed by GATC (Konstanz, Germany) and Source Bioscience (Berlin, Germany).

2.4 Kits

Kits were used to isolate plasmid DNA from *A. tumefaciens* and *E. coli*, DNA purification, probe generation/labelling for DNA blot analysis and detection of labelled probes.

Name	**Manufacturer**
NucleoSpin® Plasmid kit	Macherey Nagel GmbH Düren (Germany)
NucleoSpin® Gel/PCR Clean-up kit	Macherey Nagel GmbH Düren (Germany)
NEBlot® Phototope® Kit	New England Biolabs
Photostar® detection Kit	New England Biolabs

Table 2.4: Kits used in this study.

2.5 Solutions

All stock solutions were prepared using standard protocols. Additional solutions are shown below.

LB medium	10 g/l	tryptone
	10 g/l	NaCl
	5 g/l	yeast extract
		pH 7
		autoclave
		(for plates 15 g/l agar)
TAE-buffer (1x)	40 mM	Tris
	1 mM	EDTA
	0.1 %	Acetic acid
TBE-buffer (1x)	89 mM	Tris
	2.5 mM	EDTA
	89 mM %borate	
Agarose gel		1% agarose in
		1x TAE buffer
	0.5 μg/ml	ethidium bomide
Infiltration medium	20 mM	citric acid
	2 %	sucrose
		adjust pH 5.2 (with NaOH)
		autoclave
	100 μM	acetosyringone (dissolved in DMSO)

Table 2.5: Solutions used in this study.

2.6 Antibiotics

Antibiotic	Stock	Final	Dissolvent
Ampicillin (Amp)	100 mg/ml	20 mg/l	ddH_2O
Carbenicillin (Carb)	50 mg/ml	100 mg/l	ddH_2O
Chloramphenicol (Chl)	10 mg/ml	20 mg/l	50 % ethanol
Gentamycin (Gent)	30 mg/ml	40 mg/l	ddH_2O
Kanamycin (Kan)	100 mg/ml	50 mg/l	ddH_2O
Rifampicin (Rif)	50 mg/ml	100 mg/l	DMSO
Spectinomycin (Spec)	50 mg/ml	100 mg/l	ddH_2O

Table 2.6: Antibiotics used in this study and their stock- and final concentrations.

2.7 Equipment

Equipment	Manufacturer
electrophoretic chambers	Bio-Rad Laboratries, München Germany
cooling water bath	Colora
gel documentation	Biostep, Jahnsdorf, Germany
centrifuge, PCR cycler, photometer, cooling centrifuge	Eppendorf, Hamburg, Germany
incubation chamber, clean bench	Heraeus, Hanau, Germany
photometer add on	Implen, München, Germany
-80 °C freezer	Kendro, Hanau, Germany
PCR cycler	Life Technologies GmbH, Darmstadt, Germany
micro centrifuge	Scanspeed, Lynge, Denmark
Axioplan fluorescence microscope & camera	Carl Zeiss AG, Oberkochen, Germany

Table 2.7: Equipment used in this study.

2.8 Chemicals and Enzymes

Chemical/Enzyme	**Manufacturer**
Acetate, Potassium acetate, Tryptone	AppliChem GmbH, Darmstadt, Germany
Spectinomycin	EnZo Life Sciences
Potassium chloride, Sodium chloride	Fluka
Ampicillin, Carbenicillin, Ethidum bromide, Chloramphenicol, EDTA, Kanamycin, Rifampicin	GERBU Biotechnik GmbH
Agarose, RNAseA	Invitrogen, Darmstadt, Germany
Calcium chloride, Ethanol	Merck, Darmstadt, Germany
dNTPs, Q5®-DNA-Polymerase, restriction enzymes and buffers, SOC-medium, Taq-DNA-Polymerase, T4 DNA-ligase, DNA-ladders, T4 polymerase, Klenow-fragment, alkaline phosphatase	New England Biolabs Frankfurt a. M., Germany
Tris, glucose, glycerol, hydrochloric acid, isopropanol, magnesium chloride, yeast extract	Carl-Roth Karlsruhe, Germany
Sodium dodecyl sulfate (SDS)	SERVA Feinbiochemica, Heidelberg, Germany

Table 2.8: Chemicals and enzymes used in this study.

2.9 Software

DNA *in silico* work was done, using the DNA editing software Geneious version 6.1.2 by Biomatters. The Nikon ACT-1 software was used for procession of microscopic images.

3 Methods

3.1 Molecular biological methods

3.1.1 Molecular cloning

All vectors were generated using standard cloning techniques. Sequences of interest were PCR amplified using either Phusion or Q5 polymerase with 3' → 5' exonuclease activity. Amplified fragments were purified by the NucleoSpin®Gel and PCR Clean-up kit.

Fragments and vectors were cut by type II restriction endonucleases. Agarose gel electrophoresis was used to separate fragments of different sizes. Desired fragments were excised from the gel and purified by the NucleoSpin® Gel and PCR Clean-up kit. If necessary, DNA overhangs were blunted by treatment with Klenow fragment or T4 polymerase. Vectors were dephosphorylated by alkaline phosphatase. Prior to ligation, fragments were purified using the NucleoSpin® Gel and PCR Clean-up kit. DNA ligation was performed at either 16 ℃ over night or at room temperature (RT) for 3 h using T4 ligase. Ligated DNA was used to transform competent *E. coli* cells by heat shock transformation.

Transformed cells were selected on LB agar containing respective antibiotics. Grown colonies were first analysed by colony PCR. Positive colonies were cultured in liquid LB medium over night, followed by plasmid isolation using the NucleoSpin® Plasmid kit. Isolated plasmids were tested by restriction digestion and if correct, send for sequencing. DNA was quantified by spectrophotometric measurement. All enzymatic treatments were conducted according to the manufacturer's instructions.

3.1.2 Agarose gel electrophoresis

DNA fragments were separated using agarose gel electrophoresis. 1 % agarose gels were prepared by mixing 1 g of agarose with 100 ml of 1 x TAE buffer followed by boiling until agarose is dissolved. After cooling to 50 ℃, ethidium bromide was added to a

final concentration of 20 μl/L. The gel was transferred to a gelelectrophoretic chamber. Prior to loading, DNA was mixed with DNA loading buffer. A voltage of 80 to 100 V was applied. The gel was documented using a UV-table. Gel pictures were taken using either the digital camera C-770 Ultra-zoom (Olympus, Hamburg Germany) or E5000 coolpix (Nikon, Düsseldorf, Germany) and an adapter for filtering ethidium bromide fluorescence.

3.1.3 Genomic DNA isolation

Genomic DNA was isolated by the CTAB genomic DNA extraction protocol (Ausubel et al., 1987). Isolated DNA was treated with Ribonuclease A (RNase A) and purified by phenol chloroform extraction followed by isopropanol/ethanol precipitation. DNA was resuspended in TE-buffer and stored at -20 ℃.

3.1.4 Integration of plasmids into the chromosome

In order to introduce reporter genes into bacterial chromosomal DNA, complete vectors were integrated into different loci by single homologous recombination. To achieve this, suicide recombination vectors carrying an antibiotic resistance gene, the pUC ori for replication in *E. coli*, reporter genes and a homology region were generated. The vectors cannot replicate in *A. tumefaciens* and therefore have to recombine with chromosomal DNA to survive antibiotic selection. The homology regions thereby direct the recombination with the desired locus.

50 μl electrocompetent cells of different *A. tumefaciens* strains were mixed with 200 ng plasmid DNA in a pre-cooled electroporation cuvette (10 mm). Cells were electroplated at 1700 V followed by a 4 h incubation at 28 ℃ without any antibiotic selection. Cell suspension was plated on spectinomycin containing LB agar plates and incubated for 3 d. Correct integration of the plasmid was confirmed by DNA blot analysis.

3.1.5 Replacing of chromosomal sequences with reporter genes

In order to replace a chromosomal sequence with a reporter gene, suicide recombination vectors carrying an antibiotic resistance gene, the pUC ori for replication in *E. coli*, reporter genes, an upstream homology region and a downstream homology region were generated. *A. tumefaciens* cells were transformed with the deletion plasmids using the

same conditions as in 3.1.4. Grown colonies were screened by colony PCR for events where two homologous recombinations took place and thus the chromosomal sequence was replaced with the antibiotic resistance gene. Genomic DNA of positive colonies was isolated and deletion was either confirmed by DNA blot (IS*426* deletion) or PCR (oriT-like1 deletion).

3.1.6 DNA blot analysis

To confirm vector integrations into the bacterial chromosome and deletions of IS*426*, DNA blot analysis was used. Genomic DNA of bacteria was isolated and 10 μg were digested by the respective endonuclease for 5 to 12 h. Digested DNA was separated by agarose gel electrophoresis in TBE buffer. Cut DNA was transferred to a Hybond-N+-membrane by capillary blotting and immobilized by UV cross-linking. Probes were labelled using the NEBlot® Phototope® Kit according to the manufacturer's manual. Membrane bound DNA was hybridized with the probe. Detection was performed using the Photostar® detection Kit (NEB) according to the manufacturer's manual.

3.1.7 IPTmKanR assay

A. tumefaciens GV3101(pMP90) cells with and without pIPTmKanR were grown in 5 ml LB-medium containing rifampicin and gentamicin (without pIPTmKanR) or rifampicin, gentamicin and spectinomycin (with pIPTmKanR). OD_{600} was measured and bacteria were brought to the same value by dilution with liquid LB medium. New 5 ml LB cultures containing the respective antibiotics were inoculated with 10 μl of the levelled bacteria culture and incubated over night. Bacteria were again brought to the same OD_{600} by dilution with liquid LB medium. 4 ml bacteria were collected by centrifugation and resuspended in 300 μl liquid LB medium. Two times 150 μl were plated on LB agar plates containing rifampicin, gentamicin and kanamycin (for concentrations see 2.6). The plates were incubated at 28 °C for 4 d. The experiment was repeated with different OD_{600} values.

3.1.8 Inverse PCR

The inverse PCR method (Ochman et al., 1988) was used to confirm that a T-strand integrated into the bacterial chromosome and to determine the locus of its integration.

Surviving colonies from the IPTmKanR assay were used to inoculate a liquid LB

culture and genomic DNA was isolated. 2.5 μg genomic DNA was cut with the restriction endonuclease NcoI and purified using the NucleoSpin® Gel and PCR Clean-up kit. Purified cut DNA was ligated in a total volume of 100 μl to encourage self ligation and impede ligation of the fragments with each other. Ligated DNA was again kit-purified. Purified DNA was used as a template in an inverse PCR reaction with primers 237 and 271. These primers anneal in inverse orientation to the *npt* CDS of pIPTmKanR. After circularization of the fragment the primers face each other, allowing amplification of the unknown sequence upstream of *nptII*. Sequencing of the resulting fragment gives information about the locus where the T-DNA integrated.

3.2 Bacterial work

3.2.1 Bacterial cultivation

Bacteria were grown under aerobe conditions in liquid LB medium or on solid LB agar plates containing the appropriate antibiotics. *E. coli* was grown at 37 °C and *A. tumefaciens* at 28 °C. Liquid cultures were shaken at 180 rpm. Bacterial cultures were stored as glycerol stocks with 25 % glycerol at -80 °C.

3.2.2 Preparation of bacterial competent cells

Preparation of chemocompetent *E. coli* cells

NEB Turbo Competent *E. coli* were used to inoculate a 5 ml liquid LB pre-culture. The cells were incubated over night at 37 °C. 500 μl of the pre culture was used to inoculate a 200 ml main culture. The main culture was incubated until an OD_{600} of 0.5 was reached. The cells were collected by centrifugation and washed with ice cold 0.1 M $CaCl_2$. Cells were collected by centrifugation and resuspended in 4 ml of 0.1 M $CaCl_2$ containing 15 % glycerol. Cells were stored at -80 °C in 25 μl aliquots.

Preparation of chemocompetent *A. tumefaciens* cells

A. tumefaciens cells were used to inoculate a 5 ml liquid LB pre culture. The cells were incubated for 2 d at 28 °C. 500 μl of the pre culture was used to inoculate a 200 ml main culture. The main culture was incubated until an OD_{600} of 0.6 was reached. The cells were collected by centrifugation and resuspended in 4 ml of 20 mM $CaCl_2$. Cells were stored at -80 °C in 100 μl aliquots.

Preparation of electrocompetent *A. tumefaciens* cells

A. tumefaciens cells were used to inoculate a 5 ml liquid LB pre-culture. The cells were incubated for 2 d at 28 °C. 500 µl of the pre culture were used to inoculate a 200 ml main culture. The main culture was incubated until an OD_{600} of 0.6 was reached. Cells were collected by centrifugation and washed three times with sterile water. Cells were collected again and resuspended in 4 ml of 10 % glycerol. Cells were stored at -80 °C in 50 µl aliquots.

3.2.3 Transformation of bacterial cells

Transformation of chemocompetent *E. coli* cells by heat shock

To transform *E. coli* cells by heat shock, plasmid DNA or ligation mix was added to a 25 µl aliquot of chemocompetent cells. Cells and DNA were incubated for 30 min on ice followed by the heat shock at 42 °C for 30 s. 250 µl LB medium was added and the cells were incubated at 37 °C. After 1 h, 100 µl of the cell suspension was plated on LB agar plates containing the respective antibiotics. The plates were incubated over night at 37 °C.

Transformation of chemocompetent *A*chrDNA cells by heat shock

To transform *A*chrDNA cells by heat shock, plasmid DNA was added to a 100 µl aliquot of chemocompetent cells. Cells and DNA were subsequently incubated for 5 min on ice, for 5 min in liquid nitrogen and for 5 min at 37 °C. 1 ml LB medium was added and the cells were incubated at 28 °C. After 2 h, 100 µl of the cell suspension was plated on LB agar plates containing the respective antibiotics. The plates were incubated for 2 to 3 d at 28 °C.

Transformation of electrocompetent *A*chrDNA cells by electroporation

To transform *A*chrDNA cells by electroporation, plasmid DNA was added to a 50 µl aliquot of electrocompetent cells. Cells and DNA were pipetted into a pre-cooled electroporation cuvette. The cells were electroporated at 1,800 V. Afterwards, 1 ml LB medium was added and the cells were incubated at 28 °C. After 2 h 100 µl of the cell suspension was plated on LB agar plates containing the respective antibiotics. The plates were incubated for 2 to 3 d at 28 °C.

3.3 Plant transformation

In this work the DNA transfer from *A. tumefaciens* to plant cells was investigated. To visualize this transfer, bacterial DNA was either labelled with the *gfp* or *pat* gene. Bacteria carrying *gfp* were used to transform leafs of *N. benthamiana* by leaf infiltration followed by a screening for GFP expressing leaf cells. Bacteria carrying the *pat* gene were used to stably transform *A. thaliana* plants by floral dip and subsequent BASTA® selection.

3.3.1 *N. benthamiana* leaf infiltration

Nicotiana benthamiana (*N. benthamiana*) plants were transformed by *A. tumefaciens* leaf infiltration. Bacterial strains were grown at 28 °C over night in 5 ml liquid LB containing the respective antibiotics. If more then one strain was infiltrated, the OD_{600} of the different strains was measured. Strains were brought to the same optical density by dilution with LB medium. 500 µl of each strain was used to inoculate a new 5 ml liquid LB culture. Bacteria were grown for 4 h at 28 °C. 2 ml cells were collected by centrifugation and resuspended in 1 ml infiltration medium. The optical density was measured and cells were diluted to an OD_{600} of 0,2 in 2 ml infiltration medium. Cells were grown for 1 h at 28 °C before the leaf infiltration was performed.

A 1 ml syringe was used to infiltrate the leafs of *N. benthamiana* with the bacterial solution. If two constructs should be compared to each other, both constructs were infiltrated into each half of a single leaf. Transformed plants were incubated for 3 d before the leafs were screened for protein expression under the microscope.

3.3.2 *A. thaliana* floral dip transformation

Arabidopsis thaliana (*A. thaliana*) plants were transformed by the floral dip method (Clough and Bent, 1998). A modified version of the protocol from Logemann et al. (2006) was used. A 5 ml LB medium pre-culture containing the respective antibiotics was inoculated with bacteria from a glycerol stock or from an LB-agar plate and grown for 2 d at 28 °C. 500 µl of the densely grown pre-culture was used to inoculate a 30 ml main culture. The main culture was incubated at 28 °C until an OD_{600} of 2.0 was reached. 120 ml of 5 % sucrose solution containing 0.003 % of the surfactant Silwet L77 was prepared. The 30 ml bacteria main culture was added to the sucrose/Silwet solution in a plastic bag.

The inflorescence of *A. thaliana* plants were submerged for 10 s in the bacteria solution. Dipped plants were incubated horizontally under a lid to maintain a high level of humidity. After 24 h the plants were placed in upward position and watered for 2 - 3 weeks until seeds became mature.

Before siliques became dry and started to release seeds, paper bags were placed over the plants to collect the seeds. After 2 months completely dry plants were cut of and seeds were collected. To select for transformed plants carrying the *pat* gene in their genome, seeds were sown on wet soil until germination. Seedlings were treated with BASTA® solution (0.005 % BASTA® in tap water). Treatment was repeated after 4 and after 7 days.

3.4 Microscopy

The Carl Zeiss Axioplan fluorescence microscope was used for all microscopic analyses. Images were documented using the Nikon digital camera DMX1200 and the corresponding program Nikon ACT-1 Version 2.70.

3.5 Genome analysis

The genome of *A. tumefaciens* was screened for sequences which show homology to either the border sequences of pTiC58 or oriT consensus sequences. The alignments were either performed by the Vector NTI®software or the MUSCLE alignment algorithm. The sequences used for alignments were:

pTiC58 RB	`GTTTACCCGCCAATATATCCTGTCA`
pTiC58 LB	`GTTTACACCACAATATATCCTGCCA`
IncP orit consensus	`DCAGGATRDS`
IncQ orit consensus	`TAADTGCGCCCT`

4 Results

In this work several experiments were conducted with the goal to understand the mechanisms behind the transfer of bacterial chromosomal DNA to plant cells.

4.1 T-DNA dependent transfer of *A*chrDNA

As described in section 1.3, Ülker et al. (2008a) analysed over 375,000 transgenic *A. thaliana* T-DNA insertion mutants for the presence of *A. tumefaciens* bacterial chromosomal DNA (*A*chrDNA). They concluded that approximately one in every 250 transgenic plants contains *A*chrDNA fragments next to the T-DNA. From this observation it was hypothesized that T-DNA and *A*chrDNA might get linked to each other inside the bacterium and are then, together as a "hybrid T-DNA", transferred to the plant cell (see introduction 1.3.3).

If this was the way how *A*chrDNA is transferred to the plant cell, it has to be explained how T-DNA and chromosomal DNA get next to each other inside the bacterium. It could be explained by a hypothetical alternative procession of the T-strand after its release from the T-DNA vector. Instead of being directly transferred to the plant cell through the type IV secretion system, it might first integrate into the bacterial chromosome. In a second step the T-strand could get mobilized again from the chromosome and, together with flanking *A*chrDNA transported to the plant cell. As mentioned in section 1.3.3, the RB is almost entirely left behind in the T-DNA vector and thus could not be involved in a second mobilization. The second mobilization step could be explained by an unprocessed RB or a chromosomal sequence which is similar to the Ti-plasmid borders (border-like sequences).

4.1.1 Insertional Promoter Trapping mediated Kanamycin Resistance (IPTmKanR) assay

The prerequisite for this hypothesis is that the T-strand, after its release from the T-DNA vector, can integrate into the bacterium's own chromosome. To test if such integrations do happen, an experiment termed "insertional promoter trapping mediated kanamycin resistance" (IPTmKanR) assay was designed. In brief, single *A. tumefaciens* cells in which a T-DNA integration into the bacterial chromosome took place survive an antibiotic selection, while bacteria where such an insertion did not take place are killed.

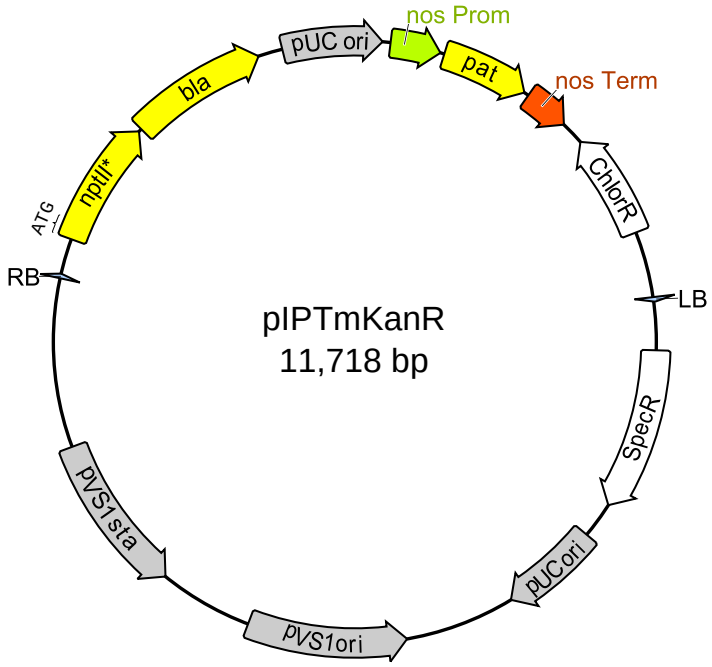


Figure 4.1: Map of pIPTmKanR. *nptII*: kanamycin resistance CDS without promoter; RB: right border; LB: left border; *bla*: carbenicillin resistance CDS; NOS Prom/Term: Promoter/Terminator from *nopaline synthase* gene; *pat*: BASTA® resistence gene (*phosphinothricin acetyl transferase*); ChlorR: chloramphenicol resistance gene; SpecR; spectinomycin resistance gene (*aadAI*); pUC ori: *E. coli* origin of replication, pVS1 ori/sta: broad host range origin of replication.

For this assay the IPTmKanR vector was generated (Fig. 4.1). It contains a T-DNA region designated by the LB and RB from *A. tumefaciens* C58 Ti-plasmid as well as the pVS1 origin of replication for plasmid replication in *A. tumefaciens*. On the T-DNA, next to the RB, an *nptII* CDS was cloned with the `ATG` start codon facing the RB. NptII allows the bacterium to grow on the antibiotic kanamycin, which normally interacts with the 30S subunit of prokaryotic ribosomes and blocks its function. However, since

the CDS of the gene was cloned next to the RB without any promoter sequence, a bacterium containing this plasmid should not be able to survive the presence of kanamycin.

Figure 4.2 shows the theoretical procedure of the assay. If the hypothesis that mobilized T-DNA occasionally integrates into the bacterium's own chromosome is correct, this should also be the case for T-strands mobilized from pIPTmKanR. Under the assumption that the T-DNA inserts randomly into the bacterial chromosomes, it would sometimes integrate next to an active chromosomal promoter. Since the CDS of *nptII* is located at the end of the T-DNA, it would then be under the control of the chromosomal promoter, leading to kanamycin resistance and colony formation on kanamycin containing LB agar medium. Hence, bacteria surviving kanamycin selection would constitute potential candidates where the integration of the T-strand into the bacterium's own chromosome took place. These candidate colonies would then be analysed for chromosomal T-strand integration by an inverse PCR strategy.

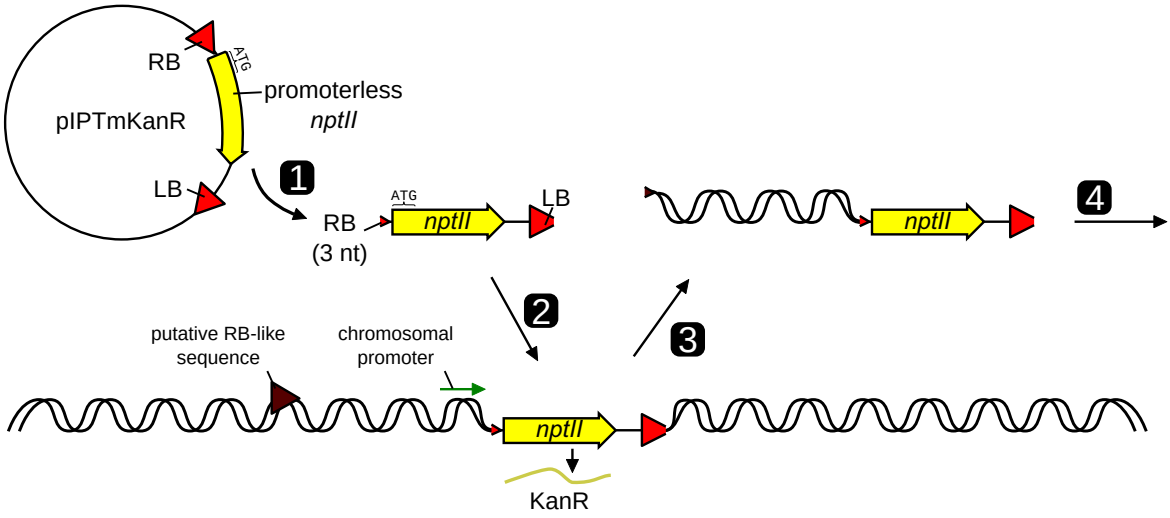


Figure 4.2: The IPTmKanR assay. **1**: T-strand mobilization of pIPTmKanR containing a promoterless *nptII* sequence. VirD2 processed T-strand contains an incomplete RB (3 nt) while the LB remains nearly complete (22 nt). **2**: T-strand integrates randomly into the bacterial chromosome and occasionally downstream of a promoter sequence. *nptII* is transcribed and the bacterium gains kanamycin resistance. **3**: T-strand and flanking *A*chrDNA are re-mobilized from the chromosome, possibly mediated by the VirD1/VirD2 complex binding to chromosomal RB-like sequences near the T-strand integration site. **4**: The T-strand/*A*chrDNA hybrid molecule is transferred to the plant cell where it integrates into the genome.

A. tumefaciens strain GV3101(pMP90) was transformed with pIPTmKanR. Transformed cells were selected using the antibiotics rifampicin (for *A. tumefaciens* strain),

gentamicin (for pMP90 helper plasmid) and spectinomycin (for pIPTmKanR). To perform the assay, confirmed transformants were plated on LB agar plates containing kanamycin. After four days the plates were screened for surviving bacteria (for details see 3.1.7).

It is known that bacteria can quickly adept to changing environments and develop spontaneous resistance against antibiotics (Jayaraman, 2011). Thus, the number of colonies, which developed resistance not because of a T-strand insertion but for example because of a mutation, should be assessed. For this reason, untransformed GV3101(pMP90) bacteria were plated under the same conditions as the pIPTmKanR containing ones.

Results of the IPTmKanR assay

Figure 4.3 shows representative colony formation after the IPTmKanR assay. Several colonies of various sizes grew, the first ones were visible after 3 days.

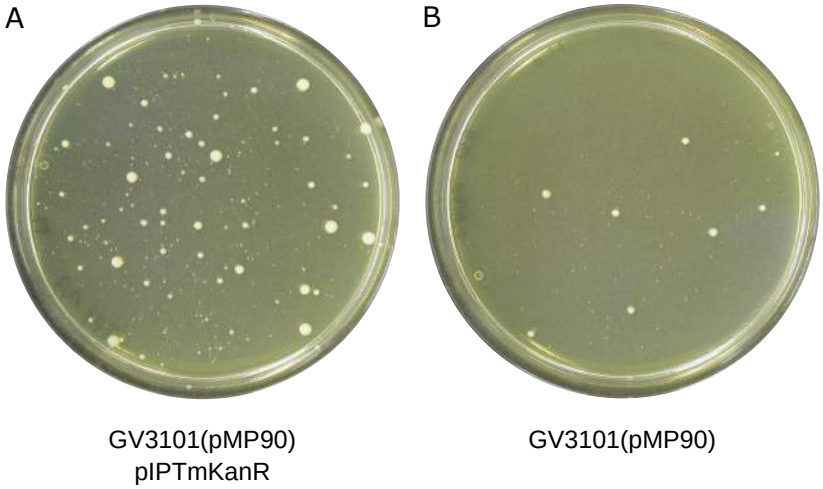


Figure 4.3: Kanamycin selection of bacteria with and without pIPTmKanR. A: Grown colonies after plating of GV3101(pMP90) cells containing pIPTmKanR on kanamycin containing medium after incubation for 4 days. **B**: Grown colonies after plating of untransformed GV3101(pMP90) cells on kanamycin containing medium after incubation for 4 days.

On kanamycin containing medium, bacteria containing pIPTmKanR were able to form much more colonies than the untransformed ones.

The number of grown colonies on the IPTmKanR plate minus the number of grown colonies on the GV3101(pMP90) plate are expected to be the true promoter trapping events linked to an integration of the T-strand. In order to estimate the amount of T-strand un-linked resistant colonies, transformed and untransformed bacteria were

grown under identical conditions and equal bacterial amounts were plated. By comparing the number of grown colonies it was estimated that ≈ 16 % of the GV3101(pMP90) pIPTmKanR colonies grown on the kanamycin containing medium might be derived from bacteria which gained spontaneous kanamycin resistance without any involvement of pIPTmKanR.

One kanamycin resistant colony was analysed by an inverse PCR approach. This assay can be used to identify the exact location on the chromosome where the insertion took place (see 3.1.8). It turned out that an insertion sequence called IS*426* was flanking the T-DNA. Thus, it seemed that the T-strand integrated next to IS*426* on the linear chromosome. Interestingly, IS*426* was also among the hotspots of frequently transferred *A*chrDNA to plant cells (Ülker et al. 2008a; Fig. 1.3.2). This strain was termed *At*KanR1 (*A. tumefaciens* kanamycin resistant colony 1).

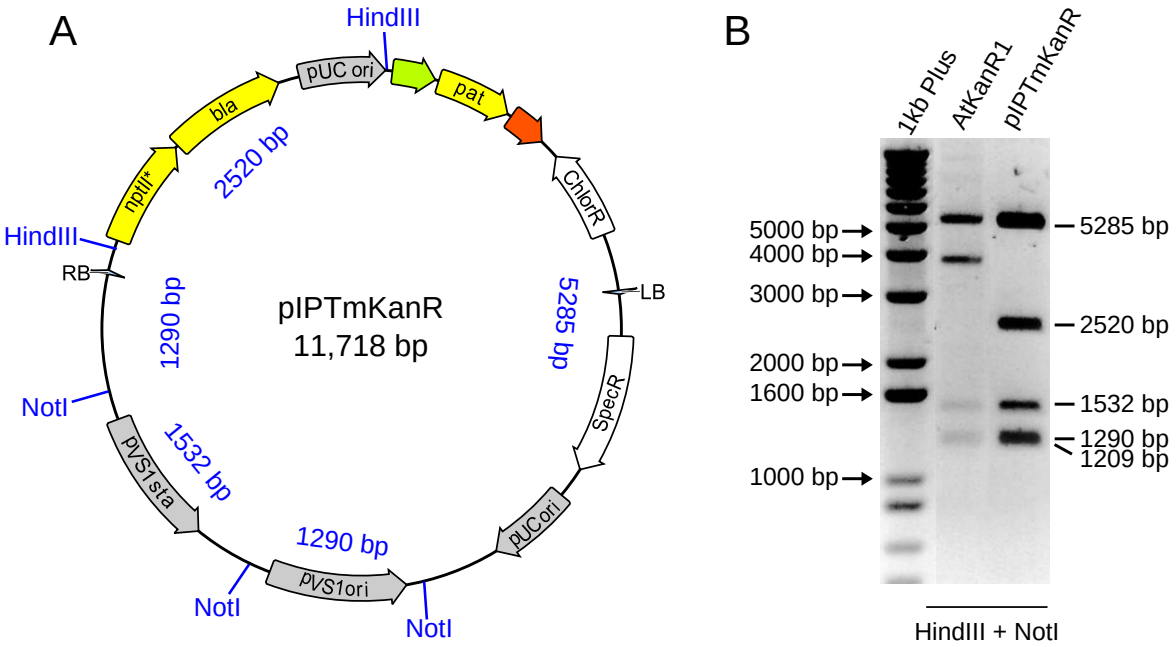


Figure 4.4: Digestion of pIPTmKanR isolated from *At*KanR1. **A**: Map of pIPTmKanR with HindIII and NotI restriction sites. Expected fragment sizes after digestion are indicated. **B**: Results of the digestion of pIPTmKanR isolated from *At*KanR1 and original pIPTmKanR as a control. The 2,520 bp fragment of the *At*KanR1 plasmid shifted to 3,800 bp.

Since IS*426* is a transposable element it might have been possible that, instead of the T-strand integrating next to IS*426* on the chromosome, IS*426* integrated itself into the plasmid. To see if this was the case, the plasmid of *At*KanR1 was isolated and analysed by restriction digestion with HindIII and NotI (Fig. 4.4 A). The resulting

band pattern was compared to the restriction pattern of the original pIPTmKanR plasmid. Figure 4.4 B shows that in *At*KanR1 the band representing the 2,520 bp fragment in pIPTmKanR shifted to a size of roughly 3,800 bp, indicating an insertion close to or inside *nptII*. Sequencing of the *At*KanR1 plasmid confirmed the presence of IS*426* inside the plasmid (Fig. 4.5 A).

In conclusion kanamycin resistance in *At*KanR1 strain was not induced due to a T-strand integration into the chromosome but because of a transposition of the chromosomally encoded IS*426* into the T-DNA of the pIPTmKanR plasmid.

PCR screening for IS*426* insertions in pIPTmKanR

IS*426* was one of the most frequently transferred *A*chrDNAs to plant cells. Because of its repeated detection, its transfer was further characterised. To see if the insertion of IS*426* was a unique event or if it frequently happened in the kanamycin resistant colonies, more colonies were analysed for the presence of IS*426* in pIPTmKanR.

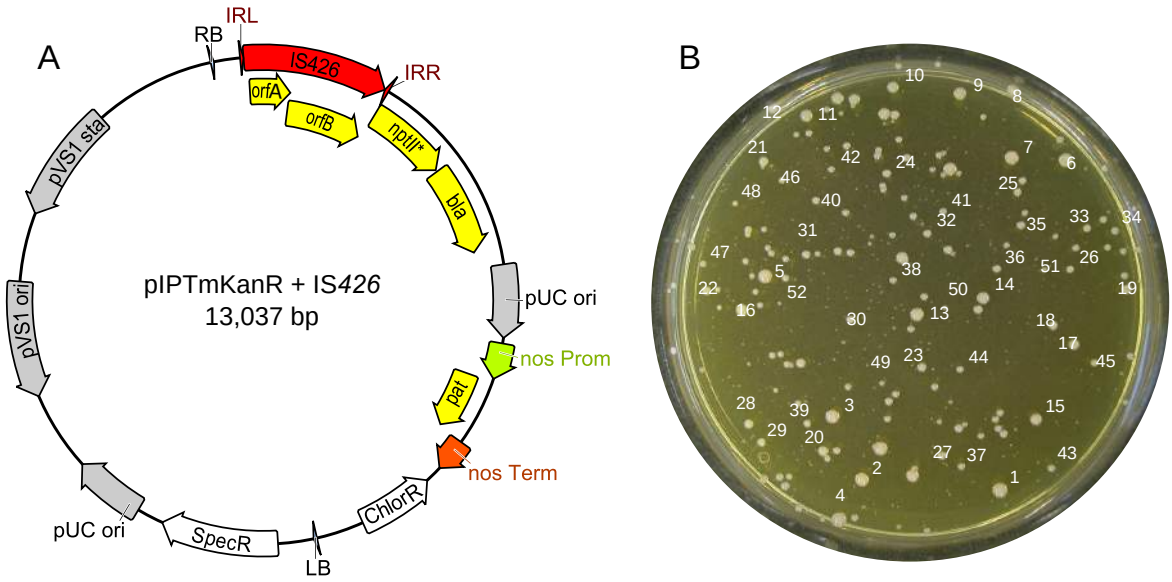


Figure 4.5: IS*426* inserted into pIPTmKanR. A: IS*426* inserted upstream of *nptII* in pIPTmKanR; IRR: inverted repeat right; IRL: inverted repeat left. **B**: For further analysis, 52 kanamycin resistant colonies were chosen according to their size from large to small.

52 colonies were chosen for a screening by colony PCR according to their size from large to small numbers (Fig. 4.5 B). Primers 591 and 397 annealing to RB and *nptII* were used to amplify the putative IS*426* insertion locus (Fig. 4.6 A). If the insertion is always directly upstream of *nptII*, colony PCR amplification of this area should allow a fast detection of IS*426* insertions.

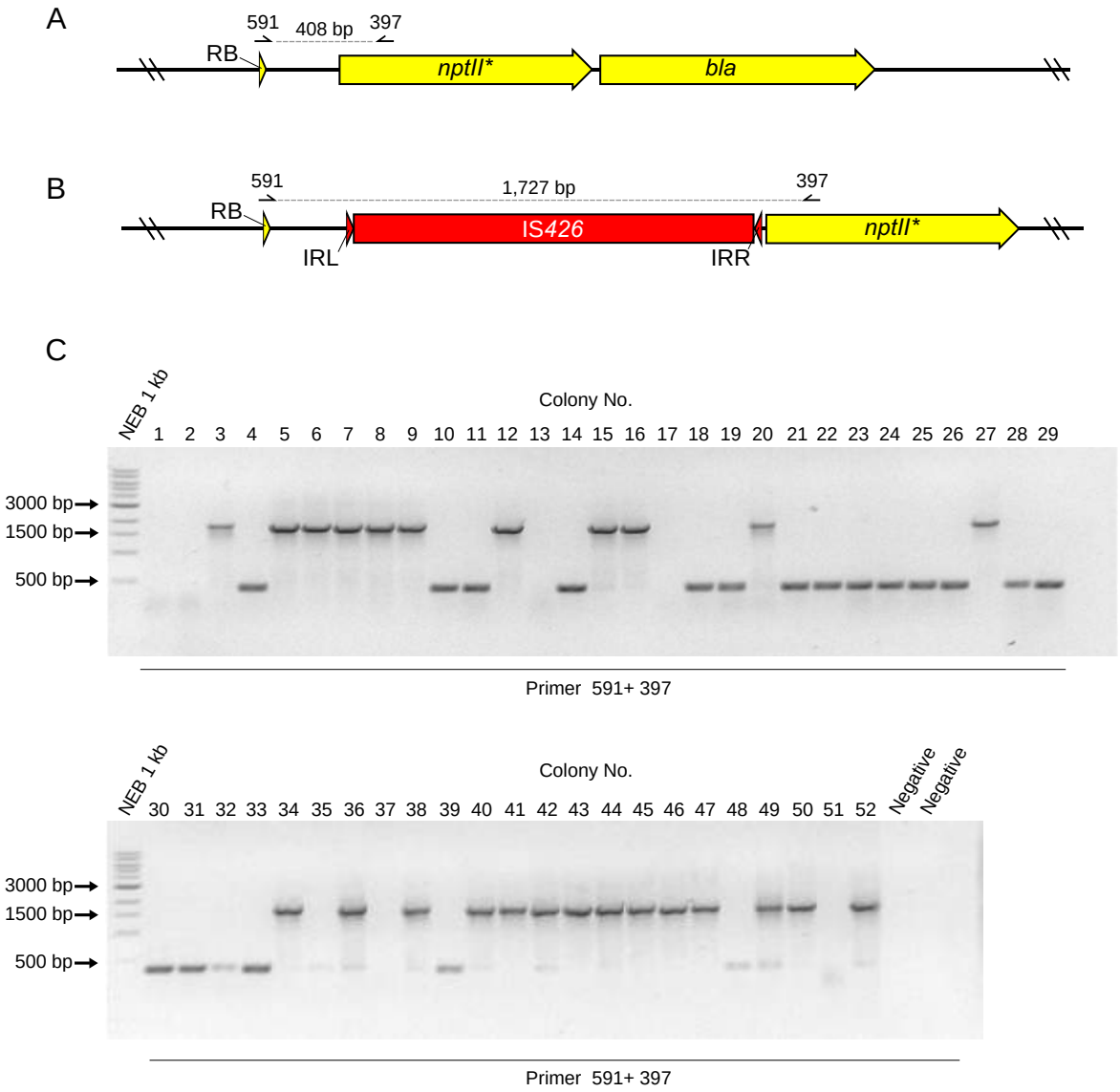


Figure 4.6: PCR screening for IS*426* insertions in pIPTmKanR. Arrows indicate primer binding sites. **A:** A 400 bp band represents the original plasmid without an insertion. **B:** A 1,700 bp bands represent IS*426* insertions into pIPTmKanR. **C:** Results from the colony PCR of 52 colonies using primers 591 and 397. Negative controls contain water instead of template DNA.

This PCR resulted in fragments of two different sizes. A 400 bp band representing the original pIPTmKanR plasmid without any insertion and a larger band of about 1,700 bp, representing an insertion of the 1.3 kbp sized IS*426* (Fig. 4.6 C). In some colonies, PCR resulted in no amplification at all. These cases were repeated and led to amplification of PCR products of either 400 bp or 1,700 bp size (results shown in table 4.1). In order to confirm the identity of the large band, several 1,700 bp PCR

products were sequenced. As expected, sequencing results showed that the 1,700 bp bands represent IS*426* insertions.

Two copies of IS*426* exist on the linear chromosome of strain C58 which only differ by a deletion of three nucleotides (see 1.3.2). Sequencing showed that both copies could integrated into the plasmid. The integration happened always upstream of *nptII* and always in the same orientation, with the inverted repeat right (IRR) of IS*426* facing towards the start codon of *nptII* (as in Fig. 4.5 A).

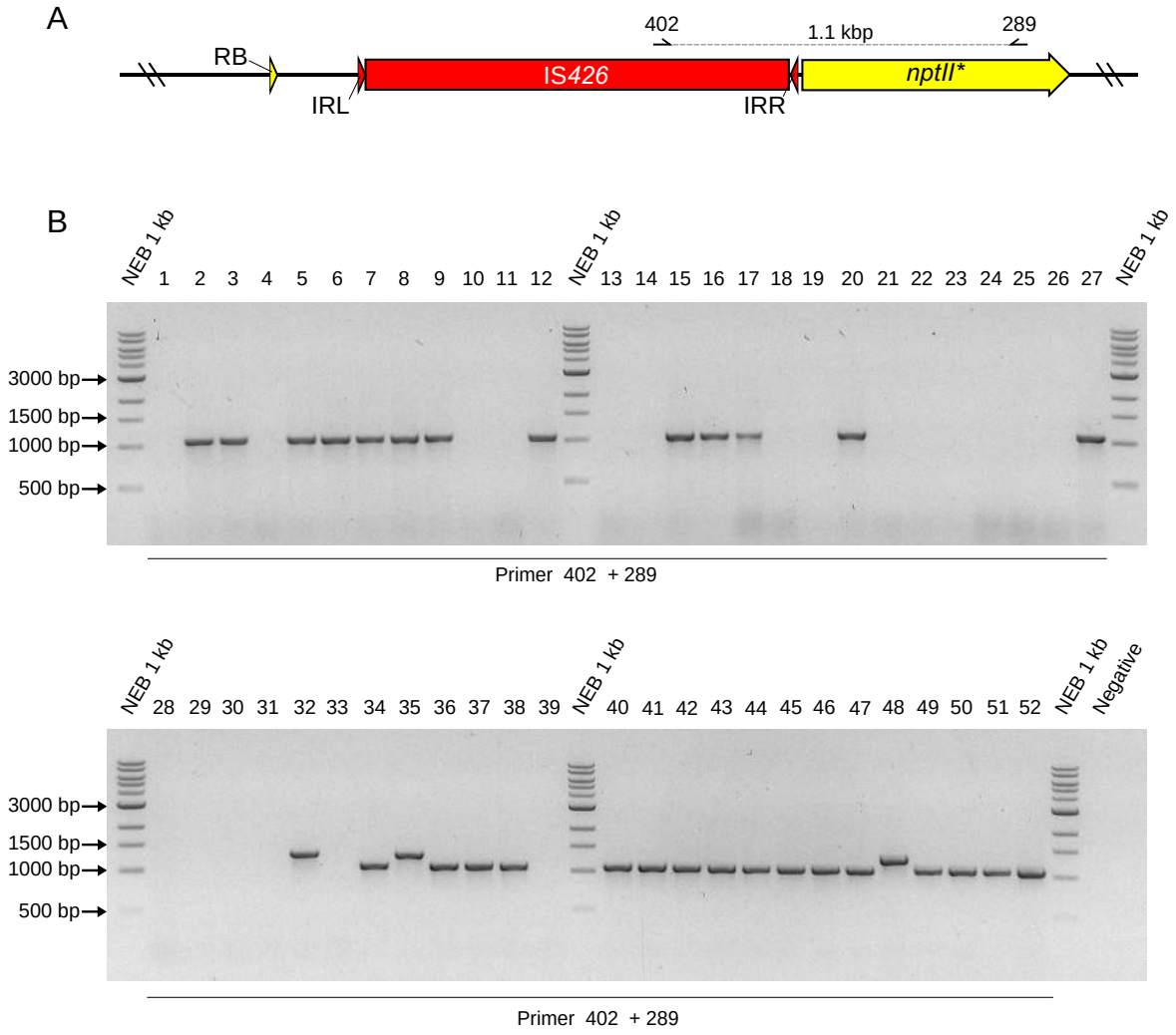


Figure 4.7: PCR screening for IS*426* insertions. Arrows indicate primer binding sites. **A:** The 1,100 bp bands represent IS*426* insertions into pIPTmKanR. **B:** Results of the PCR analysis of the 52 kanamycin resistant colonies using primers 402 and 289. In colonies 32, 35 and 48 IS*426* inserted further upstream.

To see if all 1,700 bp PCR fragments in figure 4.5 represented an insertion of IS*426*, a second colony PCR with primers 402 and 289, annealing within IS*426* and *nptII*

respectively, was performed. Again the same 52 colonies were used as template. Since one primer annealed to IS*426*, a PCR fragment would only amplify if IS*426* integrated into the plasmid. The results in figure 4.7 confirmed that indeed every 1,700 bp band of figure 4.6 represented an IS*426* insertion into the plasmid. The majority of the bands had the same size, indicating an insertion close to *nptII*. However, three bands were a bit larger and resulted in a 400 bp band in the previous PCR (591 and 397). Sequencing showed that the larger size represented an IS*426* insertion site further upstream and did not result in a 1,700 bp band with primers 591 and 397 because the insertion site was upstream of the primer 591 binding site. Taken together, 32 out of 52 colonies ($\approx 61\,\%$) showed an integration of IS*426* in pIPTmKanR.

Rearrangement of the plasmid mediated *nptII* transcription

An inverse PCR analysis (see 3.1.8) of a colony without an IS*426* insertion showed that a rearrangement of the plasmid led to a relocation of the promoter of the SpecR gene (*aadAI)*. The promoter was now in front of the *nptII* gene and transcribing it (data not shown). Because of the strong selective pressure, it might be possible that (non IS*426* mediated) resistant bacteria gained their resistance by such a rearrangement of the plasmid. To see if this relocation of the SpecR promoter happened in more than one colony, the remaining 20 colonies which did not have any IS*426* insertions (not highlighted colonies in table 4.1) were analysed by colony PCR with primers 287 and 397.

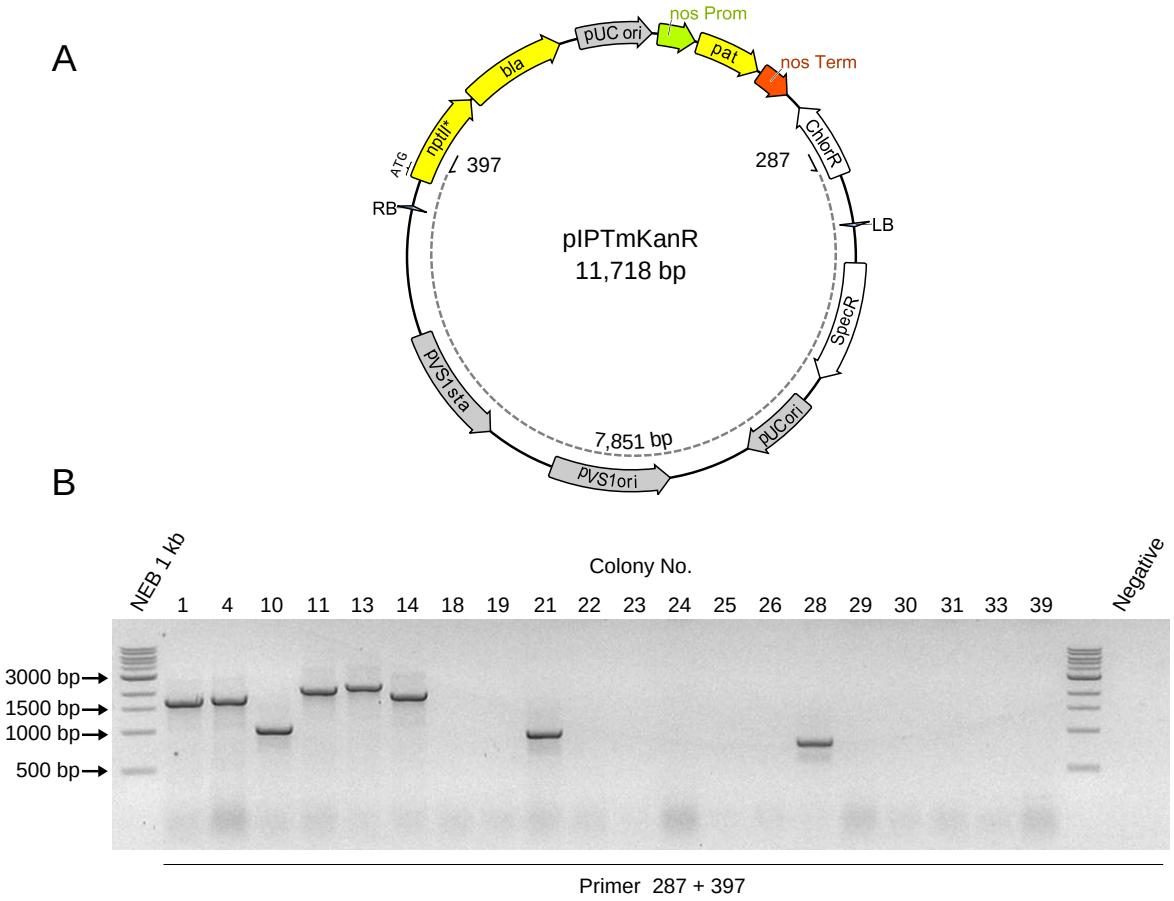


Figure 4.8: Colony PCR screening for plasmid recombinations. Arrows indicate primer binding sites. In 8 colonies fragments of various sizes were amplified indicating plasmid rearrangements in the respective colonies.

Primer 287 binds upstream of the SpecR promoter while primer 397 binds in the *nptII* CDS (Fig. 4.8 A). In the original pIPTmKanR this primer combination would result in a 7,851 bp fragment and is therefore too large to amplify. Only if a rearrangement of the plasmid took place and the promoter of the spectinomycin resistance gene is transcribing the *nptII*, a PCR product would be visible.

From the 20 tested colonies, 8 resulted in the amplification of PCR products of various sizes. This indicates that in these colonies not a T-DNA insertion into the bacterial chromosome, but a rearrangement of pIPTmKanR led to transcription of *nptII* (Fig. 4.8 B).

The remaining 12 colonies which did not gain resistance because of an insertion of IS*426* or a rearrangement of the plasmid were analysed by inverse PCR. However, no T-DNA integration into the chromosome could be detected. The results of the analysis of the 52 colonies is summarized in table 4.1.

Colony No.	591 + 397 (kbp)	402 + 289 (kbp)		Colony No.	591 + 397 (kbp)	402 + 289 (kbp)
1	0.4	/	*	**27**	1.8	1.1
2	1.8	1.1		* **28**	0.4	/
3	1.8	1.1		**29**	0.4	/
4	0.4	/	*	**30**	0.4	/
5	1.8	1.1		**31**	0.4	/
6	1.8	1.1		**32**	0.4	1.3
7	1.8	1.1		**33**	0.4	/
8	1.8	1.1		**34**	1.8	1.1
9	1.8	1.1		**35**	0.4	1.3
10	0.4	/	*	**36**	1.8	1.1
11	0.4	/	*	**37**	1.8	1.1
12	1.8	1.1		**38**	1.8	1.1
13	0.4	/	*	**39**	0.4	/
14	0.4	/	*	**40**	1.8	1.1
15	1.8	1.1		**41**	1.8	1.1
16	1.8	1.1		**42**	1.8	1.1
17	1.8	1.1		**43**	1.8	1.1
18	0.4	/		**44**	1.8	1.1
19	0.4	/		**45**	1.8	1.1
20	1.8	1.1		**46**	1.8	1.1
21	0.4	/	*	**47**	1.8	1.1
22	0.4	/		**48**	0.4	1.3
23	0.4	/		**49**	1.8	1.1
24	0.4	/		**50**	1.8	1.1
25	0.4	/		**51**	1.8	1.1
26	0.4	/		**52**	1.8	1.1

Table 4.1: Results of the screening for IS*426* insertions in pIPTmKanR. Grey: IS*426* insertion in front of *nptII*; Dark grey: IS*426* insertion farer upstream of *nptII*; White: No IS*426* insertion could be detected. Asterisk: Resistance probably mediated by plasmid rearrangement.

Using the IPTmKanR assay it was not possible to detect the insertion of a T-strand into the bacterial chromosome. Most of the colonies gained their resistance because of an integration of IS*426* or because of rearrangements of the plasmid. The remaining colonies developed their resistance probably independent of pIPTmKanR, as it was shown for untransformed bacteria. Since the integration of a T-strand into the own bacterial chromosome could not be shown, the studies were continued under the assumption that such an integration does not take place and that the *A*chrDNA transfer to plant cells is not dependent on it.

IS*426* can activate flanking genes

The fact that in kanamycin resistant colonies IS*426* always inserted in front of *nptII* in pIPTmKanR suggests that it has the ability to act as a promoter for flanking genes. The ability to control the expression of neighbouring genes is a common feature of many insertion sequences (Mahillon and Chandler, 1998). In order to confirm this ability for IS*426*, the insertion sequence was PCR amplified from *A. tumefaciens* genomic DNA and ligated in front of a promoterless *nptII* gene, which was prior integrated into vector pBasicS1 (Fig. 4.9 C; for description of the vector backbone see 4.2.1). The insertion sequence was integrated in the same orientation as it was found in the pIPTmKanR vectors. As a negative control, the same plasmid was used but without any sequence inserted (Fig. 4.9 A). As positive control, the promoter sequence from an ampicillin resistance gene was ligated in front of *nptII* (Fig. 4.9 B). Plasmids were isolated from *E. coli* cells and used to transform *A. tumefaciens* GV3101(pMP90) cells. Transformed cells were selected on kanamycin containing LB medium. Figure 4.9 shows that IS*426* and the amp-promoter were able to mediate the transcription of *nptII* and resulted in the formation of kanamycin resistant colonies. Thus, the ability of IS*426* to control expression of neighbouring genes was confirmed.

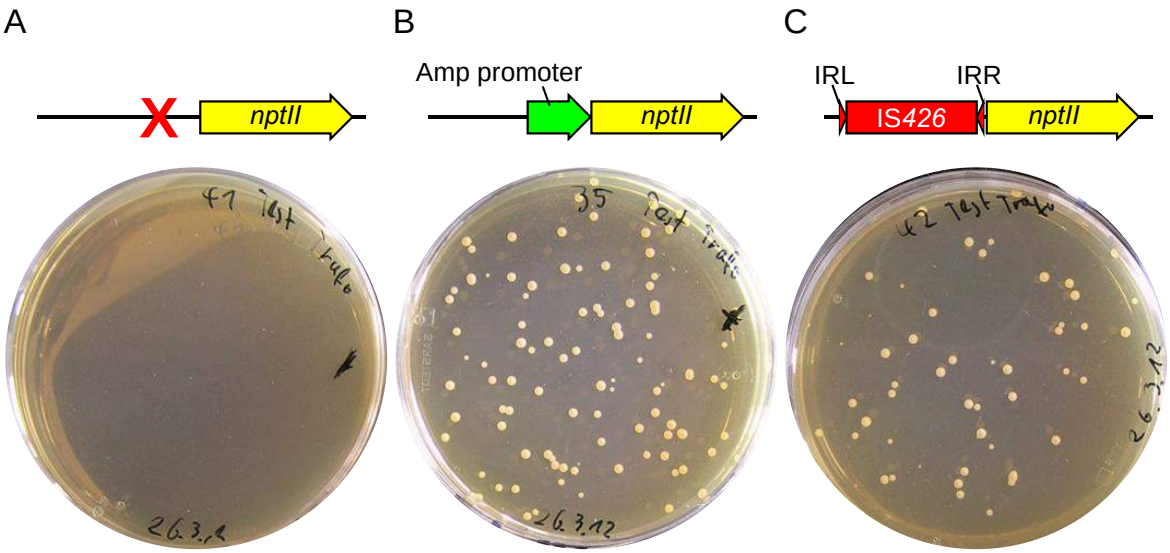


Figure 4.9: IS*426* is able to control neighbouring gene activity. **A**: Without any promoter sequence, bacteria did not survive selection on kanamycin. **B**: The amp promoter transcribed *nptII* and colonies grew on the selection medium. **C**: IS*426* was able to mediate expression of NptII and colonies grew on kanamycin containing medium.

To better characterize the insertion site of IS*426* in pIPTmKanR, 12 PCR products amplified by primers 591 and 397 (see Fig. 4.6) from different colonies were sequenced.

To see if the size of the colony has an influence on the locus of insertion, the PCR products of three small, three medium and three large colonies were chosen and the precise insertion locus was determined. Additionally, the colony PCR products of the three colonies where IS*426* seemed to have integrated further upstream were sequenced.

The sequencing results showed that depending on the size of the colony, IS*426* had a different insertion site. In the large colonies (no. 2, 3 and 5) IS*426* inserted 3 bp upstream of *nptII*. In the medium and small sized colonies IS*426* inserted in five cases 29 bp upstream of the start codon and in one case 30 bp upstream. In three cases IS*426* insertion occurred 231 bp upstream of `ATG`. However, this relative large distance to the start codon was not reflected in the size of the respective colonies, since they were not significantly smaller. It is noteworthy that, with the exception of colony 27 (insertion 30 bp upstream of *nptII*), all insertions were in either one of three positions; 3 bp, 29 bp or 231 bp upstream of the *nptII* start codon (Fig. 4.10).

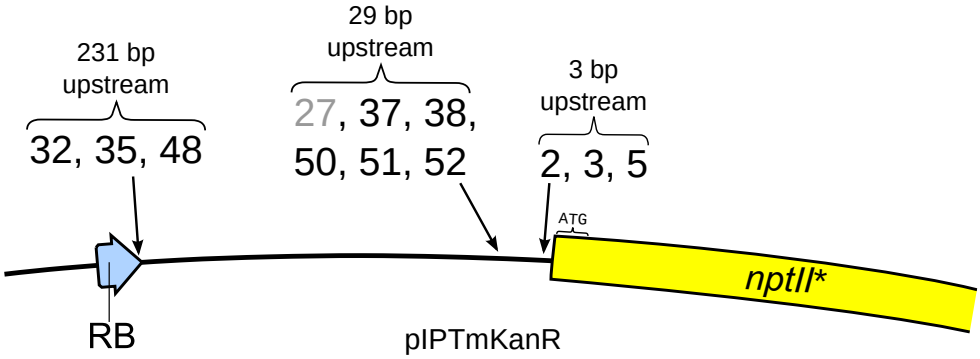


Figure 4.10: Position of IS*426* insertions in different colonies The distance of the IS*426* insertion relative to the *nptII* start codon in sequenced colonies is annotated. In colony 27 (light grey), the insertion was 30 bp upstream of *nptII*.

Deletion of IS*426* from the genome

IS*426* was found to be a major hotspot of *A*chrDNA transfer to plants by Ülker et al. (2008a). Other studies also reported its transposition into binary vectors (Vanderleyden et al., 1986; Fortin et al., 1993; Rawat et al., 2009). Additionally, as described in 4.1.1, it integrated frequently into pIPTmKanR and was able to trigger the transcription of the adjacent *nptII*. Because IS*426* is highly active and repeatedly interfered with plant transformation, an *A. tumefaciens* strain devoid of both IS*426* copies should be developed.

To accomplish this, two IS*426* knockout plasmids were generated, one to replace

IS*426* copyI by the spectinomycin antibiotic resistance gene *aadAI* and a second one to replace IS*426* copyII by the kanamycin resistance gene *nptII*.

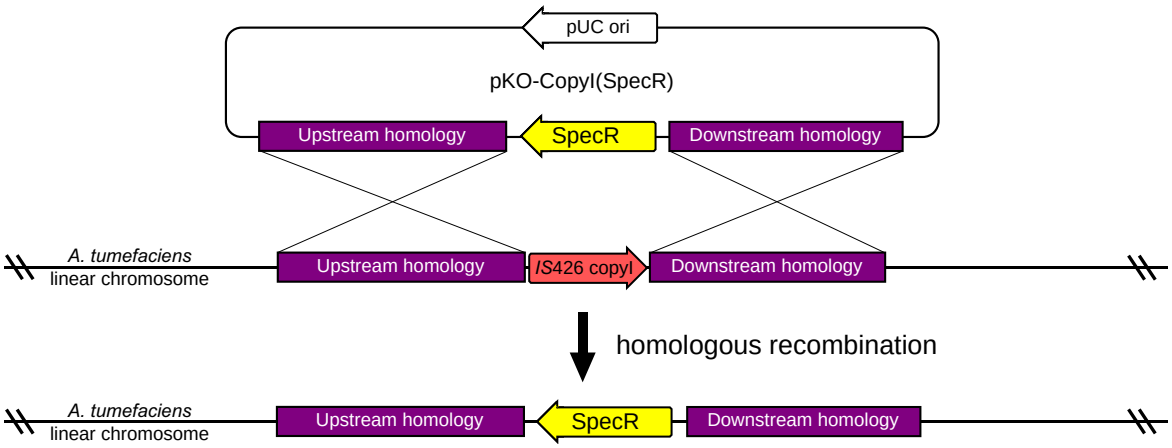


Figure 4.11: Deletion of IS*426* copyI Schematic drawing of IS*426* copyI replacement by SpecR via double homologous recombination.

The Ti-plasmid cured *A. tumefaciens* strain A136 was chosen to delete its IS*426* copies. Since A136 does not contain any Ti- or helper-plasmid, it offers the possibility to introduce a new binary vector system later. 3 kbp upstream and 3 kbp downstream homology regions were PCR amplified for each IS*426* copy. In a second step, the respective antibiotic resistance genes were amplified. *aadAI* was amplified from pBasicS1 and *nptII* was amplified from pCR™-Blunt II-TOPO®. The pUC ori for replication in *E. coli* was excised from pBasicS1. All four DNA fragments (two homology regions + selectable marker + pUC ori) were ligated and resulted in pKO-CopyI(SpecR) and pKO-CopyII(KanR). Since plasmids containing the pUC ori alone cannot replicate in *A. tumefaciens*, under selective pressure of the respective antibiotic only bacteria which integrated the plasmid into their chromosomal DNA by HR will survive. Because of the homology regions, it is likely that one region recombines with the chromosome, thereby integrating the whole plasmid into the chromosome. However, in some rare cases both HR regions will recombine with the chromosome, thereby only exchanging IS*426* with the resistance gene (Fig. 4.11).

A136 cells were transformed by electroporation with pKO-CopyI(SpecR) to delete IS*426* copyI by exchanging it with *SpecR*. The electroporated bacteria were incubated without selection and then plated on LB medium containing spectinomycin. Surviving colonies were tested for double HR by colony PCR.

A positive colony was used to generate a new set of competent *A. tumefaciens* cells and these were used to delete the second copy of IS*426* from the chromosome.

Analogous to deletion of the first copy, cells were transformed by pKO-CopyII(KanR) and selected on LB medium containing spectinomycin and kanamycin.

A third colony PCR with primers 174 and 341 amplifying both IS*426* copies was performed to confirm the absence of IS*426* (Fig. 4.12 A and B). To further show the absence of both IS*426* copies in the genome of *A. tumefaciens* strain A136, a DNA blot analysis was conducted. Genomic DNA was isolated and cut with SalI and the blotting was performed. The used probe is indicated in figure 4.12 A. The probe hybridizes to both IS*426* full copies and the partial copy on the At plasmid.

Both PCR and DNA blot confirmed that subsequently both complete IS*426* copies were replaced by antibiotic resistance genes and that the strain was free of intact IS*426* (Fig.4.12 C).

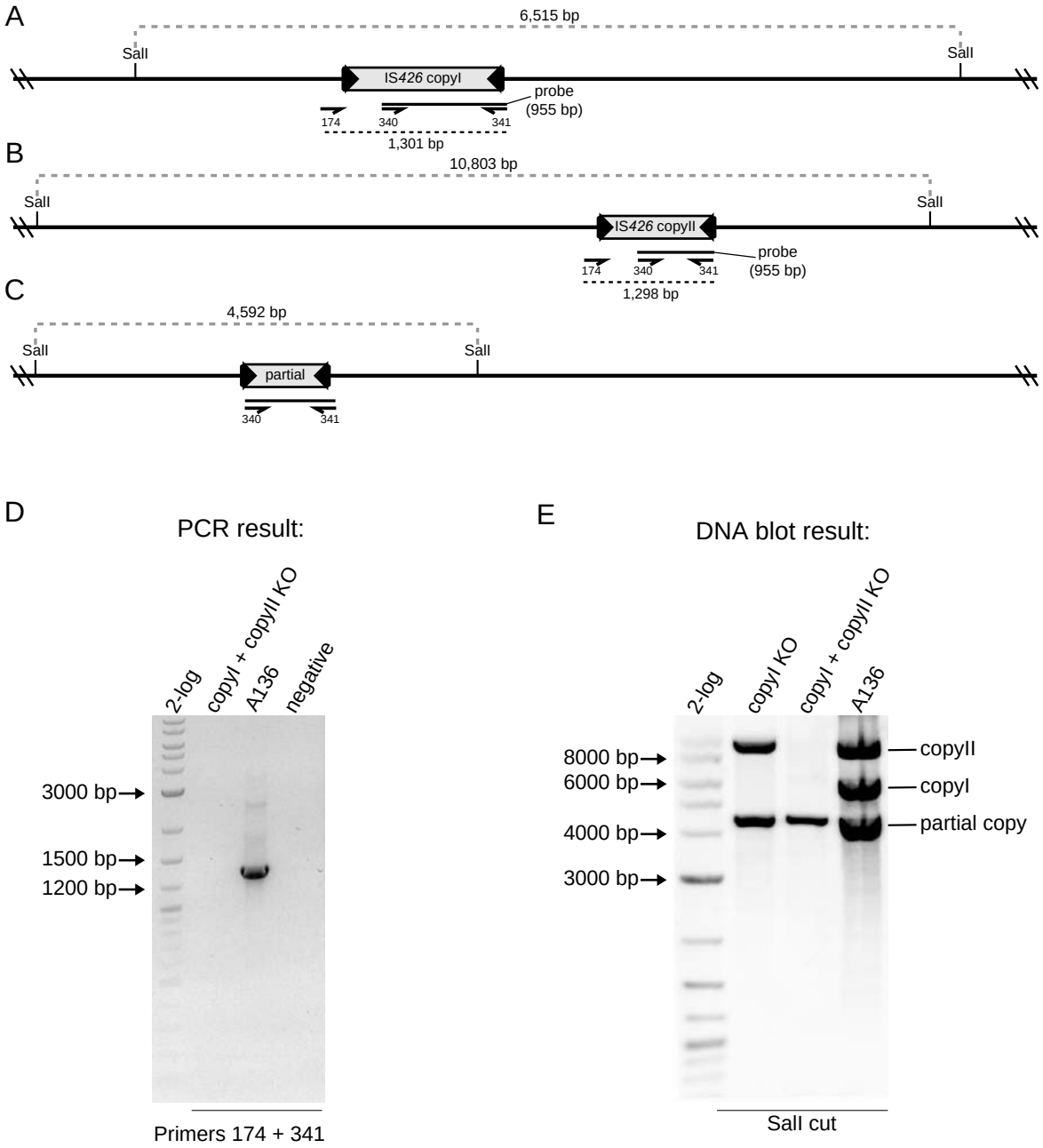


Figure 4.12: Confirming the deletion of IS*426*. Arrows indicate primer binding sites. **A - C:** DNA blot strategy to detect the three copies of IS*426*. Cutting of genomic DNA with SalI leads to three distinguishable IS*426* fragments, detectable by the 340 - 341 probe. PCR with primers 174 and 341 leads to a band of 1,301 (1,298) bp size for copy-I and copy-II. **D**: Result of the PCR with primers 174 and 341. No band amplified in the strain where IS*426* was deleted, confirming the deletion of IS*426*. A136 was used as a positive control and water as a negative control. **E**: DNA blot analysis of the IS*426* deletion. All three copies were visible in the original A136 strain. Subsequently both copies were deleted.

4.2 T-DNA independent *A*chrDNA transfer

When Ülker et al. (2008a) analysed the T-DNA insertion lines for the presence of bacterial DNA, *A*chrDNA was always found flanking the T-DNA either on the right or left side. However, this observation was biased since analysis of flanking sequence tags (FSTs), as it was done in their study, would not show an *A*chrDNA which integrated spatially separated (unlinked) from the T-DNA in the plant genome. Therefore, it is possible that T-DNA and *A*chrDNA were transferred independently of each other.

If it is the case, the conclusion made by Ülker et al. that one out of 250 transformed plants contains *A*chrDNA might have been an underestimation. All the cases where T-DNA and *A*chrDNA did not integrate into the same locus were not included in their estimation. This would also mean that no simultaneous T-DNA transfer would be necessary for the transfer of *A*chrDNA to plant cells.

To test the possibility of a T-DNA independent *A*chrDNA transfer, two different attempts were made. First, the ability of IS*426* to transfer to plant cells independently of T-DNA was analysed (4.2.2). In a second series of experiments, other hotspots of *A*chrDNA transfer were tested for their ability to be transported to plant cells without a simultaneous T-DNA transfer (4.2.3).

4.2.1 pBasic vectors

In order to test a T-DNA independent transfer from *Agrobacterium* to plant cells, a vector was needed which has an origin of replication allowing plasmid replication in *A. tumefaciens*. Furthermore, because new sequences should be tested for their ability to mediate DNA transfer to plants, the vector itself should not have this ability (should not contain T-DNA borders). To monitor the transfer to plant cells, it should either contain a *gfp* (*green fluorescent protein*) or a *pat* (*phosphinothricin acetyl transferase*) gene. *gfp* was used for transient transformation assays with *N. benthamiana* and *pat* for floral dip transformation assays with *A. thaliana*. Sequences with the potential ability to mediate DNA transfer were ligated into these vectors. If a sequence had the ability to initiate DNA transfer to plants, GFP positive or BASTA® resistant cells/plants should be obtained.

These testing vectors were generated in a three step process. The spectinomycin resistance gene from pBUcc120 was ligated with the pUC ori and MCS amplified from pCR®2.1-TOPO. The resulting vector was termed pBasic. Next, the broad host range origin of replication pVS1-ori/sta (Itoh et al., 1984) was amplified from pBUcc120.

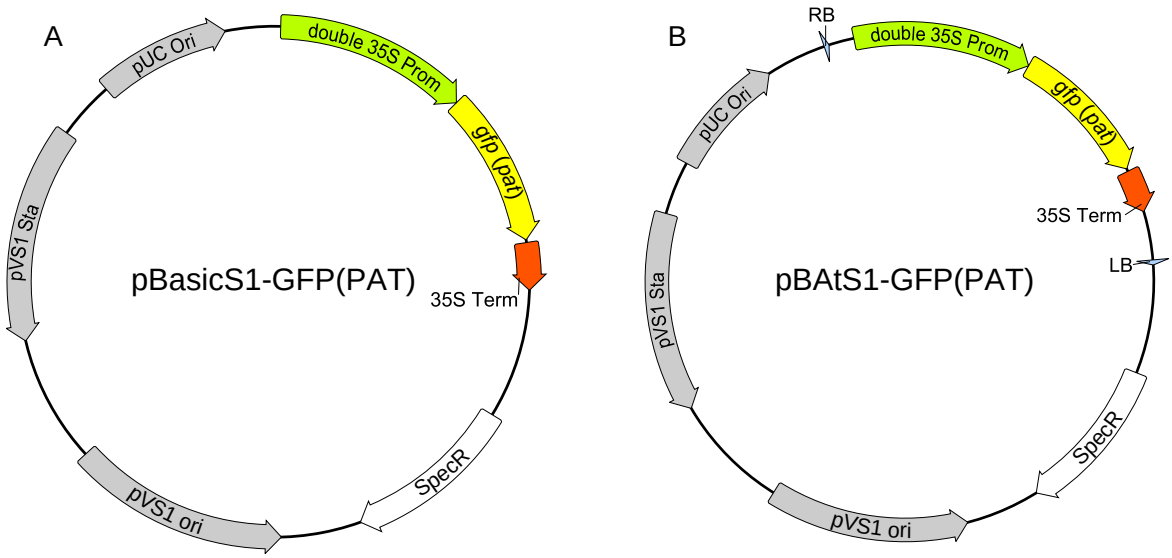


Figure 4.13: Map of pBasicS1-GFP(PAT) and pBAtS1-GFP(PAT). A: pBasicS1-GFP(PAT) does not contain any border sequences and does not transfer DNA to plant cells. DNA sequences were ligated in this vector, in order to test their DNA transfer ability. This vector was used as a negative control in all following experiments. **B**: pBAtS1-GFP(PAT) contains border sequences and transfers DNA to plant cells. This vector was used as a positive control in all following experiments.

The resulting vector was able to replicate in *A. tumefaciens* and was termed pBasicS1. This vector was used to integrate either a GFP expression cassette or the *pat* gene. The resulting vectors were called pBasicS1-GFP or pBasicS1-PAT (Fig. 4.13 A).

The vectors were used to transform *A. tumefaciens* GV3101(pMP90) cells. Transformed bacteria were used for plant transformation by either *N. benthamiana* leaf infiltration (pBasicS1-GFP) or *A. thaliana* floral dip (pBasicS1-PAT). Since the vectors do not have border sequences, no DNA transfer should take place. As expected, no GFP expressing plant cells or BASTA® surviving *A. thaliana* plants could be obtained (data not shown). These strains were used as negative controls in all following experiments. In order to have a positive control, the Ti-plasmid's border sequences were integrated into pBasicS1-GFP and pBasicS1-PAT, resulting in pBAtS1-GFP and pBAtS1-PAT respectively (Fig. 4.13 B). Plant transformation with GV3101(pMP90) cells containing these plasmids resulted in GFP expressing *N. benthamiana* leaf cells and and BASTA® resistant *A. thaliana* plants. They were used as positive controls in all following experiments.

4.2.2 IS*426* cannot be transferred to plant cells independently

IS*426* seems to play an important role in *A*chrDNA transfer and is among the most frequently transferred hotspots (Ülker et al., 2008a). In three different experiments the ability of IS*426* to transfer to the plant cells without the simultaneous transfer of a T-DNA was assessed.

IS*426* inverted repeats are not sufficient to mediate DNA transfer to plant cells

Initially, it was tested if the left and right inverted repeats of IS*426* are sufficient to mediate transport to the plant cell. Since T-DNA and IS*426* are delimited by flanking border repeats, the simplest way IS*426* could be transferred to plant cells is that its inverted repeats are recognized by a DNA transfer machinery (e.g. the VirD1/VirD2 complex or its transposase) which transfers IS*426* to plant cells.

The plasmid pBasicS1-GFP which contains the pUC and pVS1 origin of replication for plasmid replication in *E. coli* and *A. tumefaciens*, a *gfp* under the control of the cauliflower mosaic virus 35S promoter and the spectinomycin resistance gene *aadaI* was used to test this hypothesis. Since the plasmid does not contain T-DNA border sequences, no DNA can be transferred from this plasmid to plant cells.

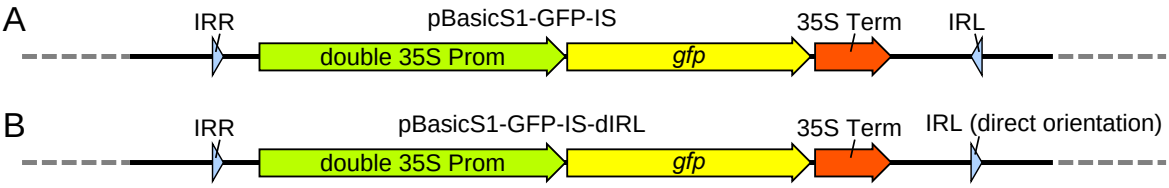


Figure 4.14: Maps of pBasicS1-GFP-IS and pBasicS1-GFP-dIRL. A: The inverted repeats (≈ 230 bp each) of IS*426* were arranged in the original (inverted) orientation. **B**: The inverted repeats were arranged in direct orientation. Backbone contains the pUC and pVS1 ori for replication in *A. tumefaciens* and *E. coli* and a SpecR gene.

Inverted repeats IRL and IRR of IS*426*, together with flanking sequences (≈ 230 bp for each repeat) were PCR amplified from genomic DNA of *A. tumefaciens* strain C58. The PCR fragments were ligated upstream and downstream of *gfp* in the same orientation they have, flanking the IS*426* transposase in the chromosome. The plasmid was named pBasicS1-GFP-IS (Fig. 4.14 A).

Since the borders on the T-DNA are in direct orientation while the terminal repeats of IS*426* are in inverted orientation, it was tested if IRR and IRL in direct orientation mediate transfer of enclosed DNA. Analogous to pBasicS1-GFP-IS, a second plasmid

was generated where the orientation of IRL was changed from inverted to direct. The plasmid was named pBasicS1-GFP-IS-dIRL (Fig. 4.14 B).

These plasmids were used to transform GV3101(pMP90) cells. Transformed bacteria were infiltrated into *N. benthamiana* leaves and after three days leaves were screened for GFP expression. Since the IS*426* repeats were flanking the *gfp* gene, GFP expressing plant cells would indicate the transfer of the *gfp* gene, mediated by the IS*426* repeats. GV3101(pMP90) cells containing pBasicS1-GFP or pBAtS1-GFP were used as negative and positive controls respectively. No GFP expression could be detected in plant cells for pBasicS1-GFP-IS, pBasicS1-GFP-IS-dIRL and pBasicS1-GFP. The positive control pBAtS1-GFP (contains RB and LB) led to GFP expressing plant cells.

In order to analyse if DNA gets transferred to plant cells in a stable *A. thaliana* transformation, the *gfp* gene was exchanged to a *pat* gene, which allows screening for stably transformed *A. thaliana* plants by BASTA® selection. Transformation with both constructs (pBasicS1PAT-IS and pBasicS1PAT-dIRL) did not lead to herbicide resistant plants. Thus, the IS*426* inverted repeats alone do not seem to mediate the transfer of IS*426* to plant cells (data not shown).

Insertion of *gfp* into IS*426*

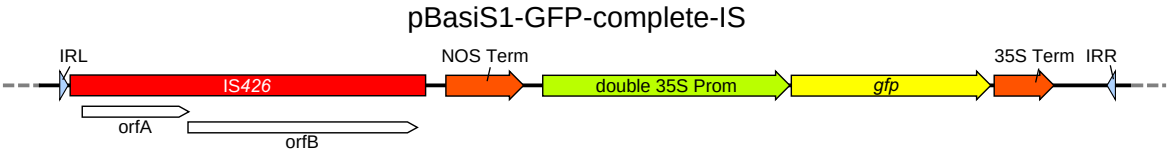


Figure 4.15: Map of pBasicS1-GFP-complete-IS. A *gfp* gene and a NOS terminator were introduced into IS*426*. Open reading frames A and B (orfA/B) of IS*426* are indicated. Backbone contains the pUC and pVS1 ori for replication in *A. tumefaciens* and *E. coli* and a SpecR gene.

In contrast to the prior experiment where only the borders were included, a second experiment was conducted in which the complete IS*426* was labelled with a GFP expression cassette. IS*426* was amplified by two PCR reactions. The first PCR covered the IRL and both open reading frames, while the second contained only the IRR. These fragments were cloned into the pBasicS1 multiple cloning site. In a second step a GFP expression cassette was introduced into the multiple cloning site, resulting in IS*426* with a *gfp* gene between orfB and IRR. To prevent the transposase promoter from driving the expression of GFP, a NOS terminator was inserted between orfB and *gfp*. The resulting plasmid was named pBasicS1-GFP-complete-IS (Fig. 4.15).

The plasmid was used to transform *A. tumefaciens* GV3101(pMP90) cells. Transformed bacteria were infiltrated into *N. benthamiana* leaves followed by a screening for GFP expressing cells using pBatS1-GFP and pBasicS1-GFP strains as positive and negative controls. Only infiltration of the positive control resulted in GFP expressing *N. benthamiana* leaf cells (data not shown). Therefore, even if the complete IS element is present, no DNA transfer to plant cells was detectable.

Insertion of IS*426* into vector backbones

To see if an IS*426* insertion into the backbone of a vector can lead to DNA transfer to plants, a third experiment was conducted. It was previously reported that IS*426* is frequently inserting itself into binary T-DNA vectors (Vanderleyden et al. 1986; Fortin et al. 1993; Rawat et al. 2009; this work). A transposition of IS*426* into the vector backbone might lead to homologous recombination with a chromosomal IS*426* copy. This would result in a complete integration of the T-DNA vector into the chromosome, thereby being flanked by two IS*426* copies. This in turn might lead to a transfer of IS*426* and plasmid to the plant cell.

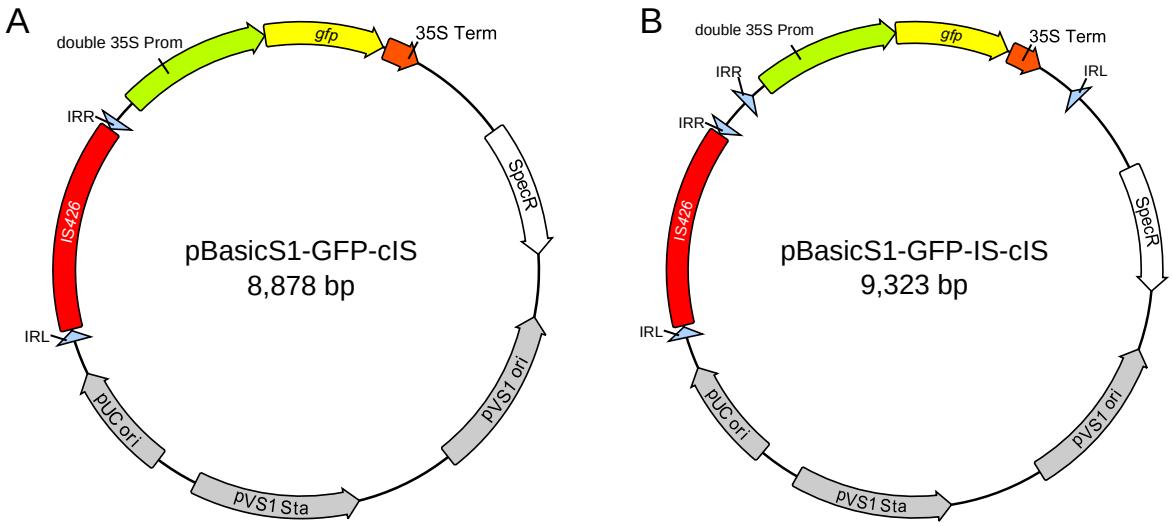


Figure 4.16: Map of pBasicS1-GFP-cIS and pBasicS1-GFPIS-cIS. A: IS*426* was inserted into the backbone of pBasicS1-GFP. **B:** IS*426* was inserted into the backbone of pBasicS1-GFP-IS.

The complete IS*426* copy-I was amplified from genomic DNA of *A. tumefaciens* strain C58. The PCR fragment was ligated into the backbone region of pBasicS1-GFP (contains no T-DNA borders) resulting in pBasicS1-GFP-cIS (Fig. 4.16 A). The plasmid was used to transform GV3101(pMP90) cells and *N. benthamiana* leaf cells

were infiltrated. In this experiment is was only assessed, if the presence of IS*426* in the backbone of a *gfp* containing vector can lead to transfer of *gfp* to plant cells. It was not assessed, if the integration of the vector into the chromosome is an intermediate step in this process. Nevertheless, no GFP expressing cells could be observed (data not shown), indicating that the presence of IS*426* in the vector backbone does not lead to DNA transfer to plant cells.

Additionally, the complete IS*426* sequence was PCR amplified and ligated into the backbone of pBasicS1-GFP-IS, resulting in pBasicS1-GFP-IS-cIS (Fig. 4.16 B). This plasmid contains the borders of IS*426* flanking a *gfp* gene, next to the complete IS*426*. Similar to the situation a in compound transposon (two insertion sequences flanking a random DNA fragment) in this vector two IS*426* sequences were next to each other. One was thereby complete and one was only represented by its inverted repeats, flanking *gfp*.

GV3101(pMP90) cells were transformed with the plasmids and the resulting strains were infiltrated into *N. benthamiana* leaves with the prior used controls. Only the positive control led to GFP expressing leaf cells (data not shown).

The above described experiments led to believe that IS*426* is not able to transfer to plant cells on its own.

4.2.3 Insertion of reporter genes into the bacterial genome

With the exception of IS*426*, so far the transfer of *A*chrDNA to plant cells was only observed *in silico* in FSTs obtained from online databases and in some examples where transgenic plants were analysed by DNA blot or plasmid rescue analysis (Ülker et al., 2008a). This *A*chrDNA transfer always happened unintentionally during generation of the insertion mutants and was only discovered later by Ülker et al.. However, to investigate the underlying mechanism and the key factors involved in it, it was necessary to reproduce the *A*chrDNA transfer in the laboratory. Furthermore, a rapid assay was desirable without the need for extensive plant cultivation and screening.

For this reason, an assay was developed which allowed a fast and reproducible monitoring of *A*chrDNA transfer to plant cells. Two reporter genes (*gfp* for a transient transformation assay in *N. benthamiana* and *pat* for a stable transformation assay in *A. thaliana*) were introduced into different loci of the genome of *A. tumefaciens*. Since the reporter genes were stably integrated into the bacterial chromosomes, their expression in plant cells after transformation would indicate a transfer of bacterial chromosomal DNA to plants.

The reporter genes were cloned into suicide vectors containing a selectable marker and a sequence (HR region) which was homologous to different regions of the bacterial chromosomes. Vectors containing the pUC origin of replication cannot replicate in *A. tumefaciens*. Hence, under selective pressure by spectinomycin, only bacteria which integrated the vector containing the reporter genes and the selectable marker into their chromosome survive (Fig. 4.17). The large HR regions thereby mediate the integration into the desired locus. The homology regions were chosen in a way that bacterial genes do not get disrupted upon integration of the plasmid.

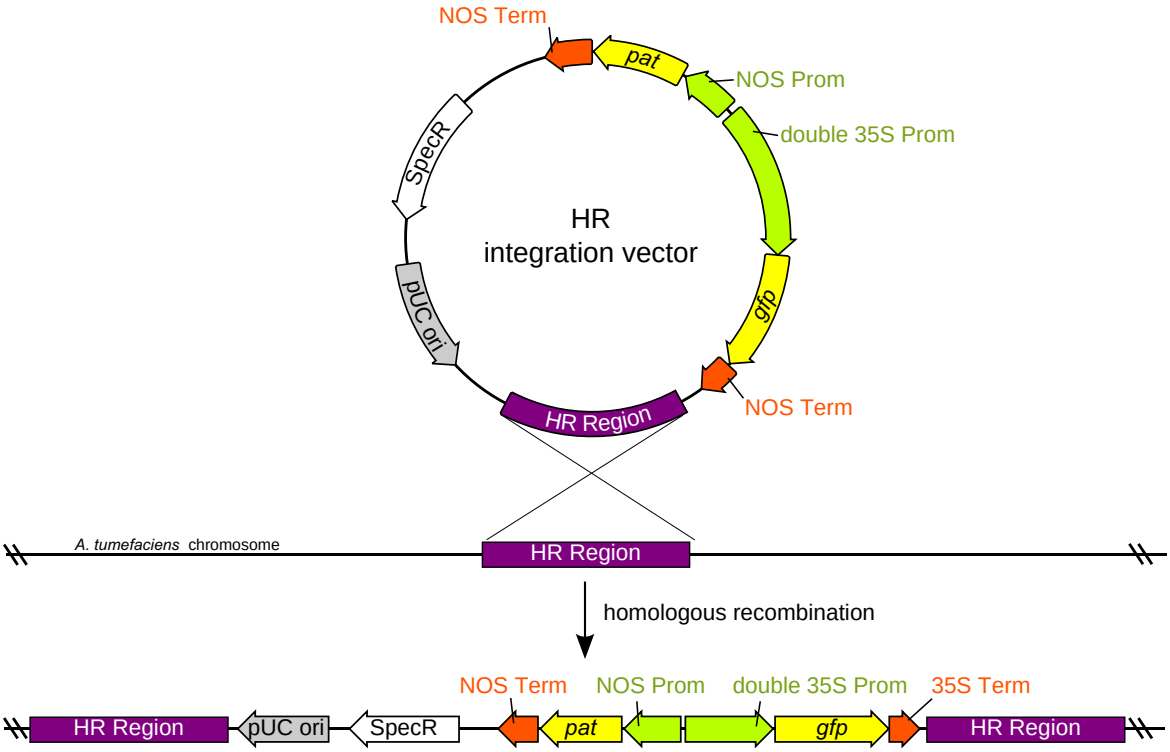


Figure 4.17: Suicide-plasmid mediated integration into the chromosome by homologous recombination. Homology region (HR Region) in HR integration vector leads to its complete integration into the chromosome and thereby integrates reporter genes in the respective locus. The plasmid has no origin of replication for *A. tumefaciens*. Thus, only bacteria where the vector integrated into the chromosome survive spectinomycin selection.

The decisions which chromosomal loci are suitable for tagging with the reporter genes were based on the observations made by Ülker et al. (2008a). As described in 1.3.2, the authors were able to identify regions of the linear *A. tumefaciens* chromosome which were more often transferred to plant cells than other regions (hotspots of *A*chrDNA transfer). Reporter genes were integrated into two regions within the main accumulation of hotspots, which occurs roughly between 1.3 Mb and 1.5 Mb on the linear chromosome (see Fig. 1.7). These two insertions should represent the main hotspot region. Since this region is large (200 kbp) and because it is not clear where the limits of each hotspot are, it cannot be predicted which of the different hotspots is actually tagged by these insertions.

In addition to the integrations in the hotspot region, two integrations were conducted in a region of the chromosome which does not contain any hotspots. These integrations served as negative controls for the transfer. Reporter genes were also integrated into one region of the circular chromosome which was found in GABI-Kat T-DNA insertion line 052H10 by Ülker et al. (2008a) (all insertion shown in Fig. 4.18).

The selected regions are described below (Nomenclature: HS: hotspot, NE: negative control, LC: linear chromosome, CC: circular chromosome).

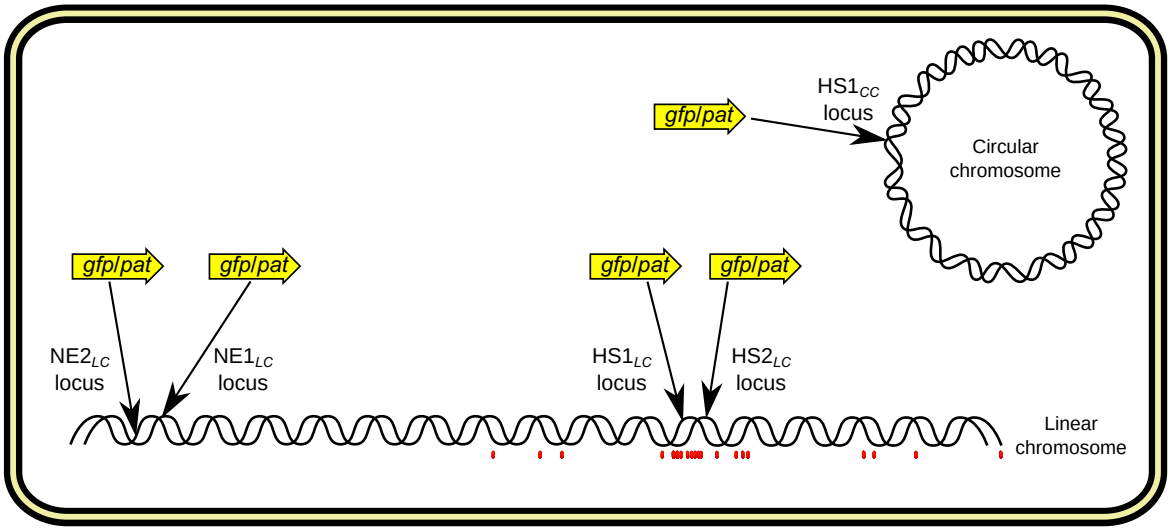


Figure 4.18: Reporter genes were inserted in the genome of *A. tumefaciens*. Reporter genes were inserted in the bacterial genome by homologous recombination of HR insertion vectors in five locations. Red diamonds indicate hotspots of *A*chrDNA transfer, detected by Ülker et al. (2008a).

$HS1_{LC}$ (1^{st} integration into the hotspot region on linear chromosome) This locus should represent the most frequently transferred region of the linear chro-

mosome. The *glxA* gene (Atu4225) was amplified and cloned into the integration plasmid to serve as homology region (base 1,350,145 - base 1,351,319). It is located 243 bp upstream of the RB-like1 sequence, identified by Ülker et al. (2008a) in GABI-Kat line 086C02. If the RB-like1 sequence is involved in the transfer of *A*chrDNA, the reporter genes, integrated into the $HS1_{LC}$ locus, should be transferred plant cells.

$HS2_{LC}$ (2^{nd} integration into the hotspot region on linear chromosome) A second locus within the main accumulation of hotspots was chosen for reporter gene integration. It is located 33 kbp downstream of the $HS1_{LC}$ locus in a region, which was frequently transferred to plant cells. As homology region, an ABC transporter (Atu4259) was amplified from the linear chromosome (base 1,386,626 - base 1,384,692).

$HS1_{CC}$ (integration into a region on the circular chromosome which was found in plant cells) Ülker et al. (2008a) detected this locus during their rescue experiments in the *A. thaliana* GABI-Kat mutant line 052H10. It originates from the circular chromosome. A gene belonging to the GntR family was amplified (base 2,471,078 - base 2,473,047) and ligated into the integration plasmid to serve as homology region. Since the hotspots of frequent *A*chrDNA transfer of the circular chromosome were not identified by Ülker et al. (2008a), it is not known how frequently this locus is actually transferred.

$NE1_{LC}$ (negative control 1 on linear chromosome) As negative control two insertions were made in chromosomal regions that were never found in plant cells. Since this is the case for the first quarter of the linear chromosome, the first insertion side was chosen randomly in this region. The *mfs* gene was used as homology region for the integration plasmid (base 201,238 - base 203,430).

$NE2_{LC}$ (negative control 2 on linear chromosome) For the second negative control the *picA* locus was chosen to serve as homology region. Is is also located in the first quarter of the linear chromosome and therefore was never detected in plant cells. Furthermore, it was reported that insertions in this locus do not affect growth or plant transformation efficiency of the bacterium (Lee et al., 2001; Oltmanns et al., 2010).

These regions were PCR amplified and ligated into the HR integration vectors containing the reporter genes. The vectors were used to transform three different *A.*

tumefaciens strains (GV3101(pMP90), A136 and ATΔ*virD2*) by electroporation as described in 3.1.4. In the following text, the chromosomal loci in which the reporter genes were integration will be referred to as either $HS1_{LC}$, $HS2_{LC}$, $HS1_{CC}$, $NE1_{LC}$ or $NE2_{LC}$. The bacterial strain having the reporter genes inserted into the respective locus will be referred to as a combination of the locus name, " ::*pat*/*gfp*" and the name of the strain which was used for integration in parenthesis. For example, strain A136 carrying the reporter gene insertion in the $HS1_{CC}$ locus is termed $HS1_{CC}$::*pat*/*gfp*(A136).

4.2.4 DNA blot analysis of vector integrations

To confirm vector integrations into the correct loci, DNA blot analyses were performed. The transformed cells were plated on LB medium containing the necessary antibiotics for the selection of the respective strain and spectinomycin for selection of the inserted vector. Grown colonies were cultured and genomic DNA was isolated. A 490 bp internal probe, hybridizing to *gpf*, was used to confirm all recombinations in all strains. To further confirm the correct insertion in the $HS1_{LC}$ locus in GV3101(pMP90), additionally an external probe, hybridizing downstream of the homology region on the linear chromosome, was used (Fig. 4.19 A).

DNA blot analysis of vector insertions in the $HS1_{LC}$ locus

Figure 4.19 shows the results of the vector integrations in the $HS1_{LC}$ locus. For the analysis with the external probe, genomic DNA was cut with BamHI. In case the vector did not insert, a 11,694 bp fragment should be visible. If the vector inserted correctly, a band of 4,648 bp should be detectable. Figure 4.19 C shows the result of the DNA blot using the external probe of two tested colonies. The blot confirmed the integration in the correct locus, since a clear band shift was visible when comparing it to untransformed cells.

For the analysis with the internal probe, genomic DNA of GV3101(pMP90), A136 and ATΔ*virD2* cells containing the insertions was cut with NdeI. Correctly inserted vectors would lead to a 2,604 bp fragment on the blot while untransformed bacteria should not result in any band, since no *gfp* was present. Figure 4.19 D shows that for each strain a band of the correct size was visible. However, a second band of approximately 6 kbp was visible for the $HS1_{LC}$::*pat*/*gfp*(ATΔ*virD2*) strain. It is possible that the vector inserted a second time in the same locus or apart from it. However, since a band of the correct size was visible, this strain was used for further experiments. No

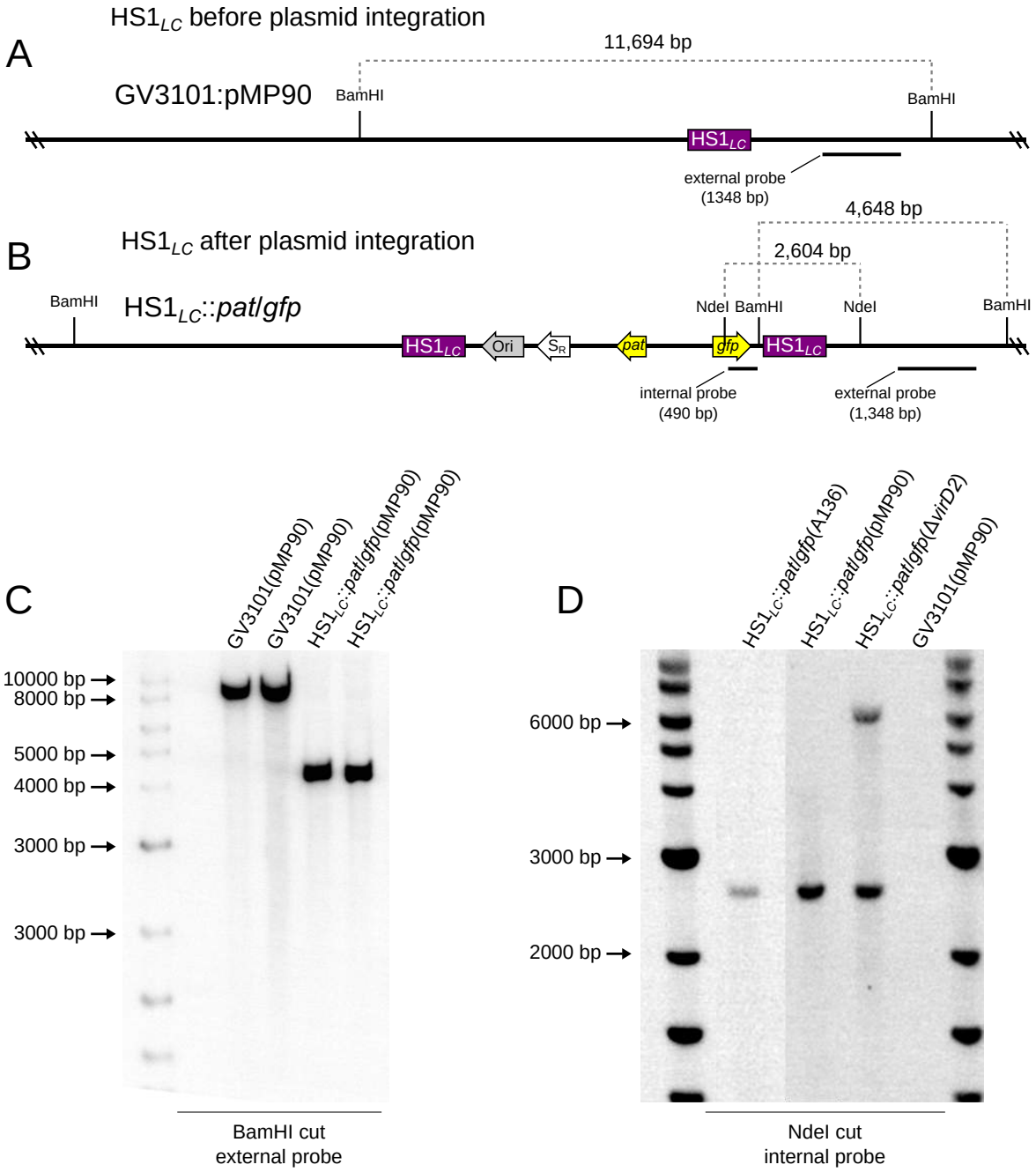


Figure 4.19: DNA blot analysis of vector insertions in the HS1$_{LC}$ locus. A, **B**: Map of the locus with and without vector insertion. **C**: DNA blot result using the external probe. The blot confirmed the correct integration in GV3101(pMP90). **D**: DNA Blot result using the internal probe in different strains[1]. The blot confirmed the correct integration in all strains. In HS2$_{LC}$::*pat*/*gfp*(ATΔ*virD2*) a second, larger band was visible, indicating a second integration.

[1]For simplicity, blot pictures were combined. Original blots can be found in supplement (Fig. S1).

band was detectable for the untransformed GV3101(pMP90) strain.

DNA blot analysis of vector insertion in the $HS2_{LC}$ and $HS1_{CC}$ locus

To confirm the vector integrations in the $HS2_{LC}$ locus, genomic DNA of all strains was isolated and cut with EcoRV. Correct insertions would result in a band of 3,673 bp. The blot confirmed the correct integration in all *A. tumefaciens* strains (Fig. 4.20 A and C).

Genomic DNA of strains having the vector integrated in the $HS1_{CC}$ locus was cut with EcoRV. Correctly inserted plasmids resulted in a band with a size of 3,551 bp (Fig. 4.20 B, D). The band was present in all strains. However, in the case of $HS1_{CC}$::*pat*/*gfp*(ATΔ*virD2*) a second larger band was visible. This band could have either resulted from a second insertion of the vector or because of an incomplete digestion by EcoRV, since the amount of genomic DNA used for this strain was very high. Because a band of correct size was visible, it was decided to continue with this strain. No band could be observed for untransformed *A. tumefaciens* A136 cells (Fig. 4.20 D).

DNA blot analysis of vector insertion in the $NE1_{LC}$ and $NE2_{LC}$ locus

An integration of the recombination vector in the $NE1_{LC}$ locus (negative control) was only conducted in GV3101(pMP90) cells. NdeI cut genomic DNA should result in a band with a size of 5,177 bp (Fig. 4.21 A). The DNA blot showed a band with the correct size but additionally a second large band, possibly originating from a second integration (Fig. 4.21 C).

For the second negative control, the vector was integrated into the $NE2_{LC}$ locus. Genomic DNA cut with EcoRV should result in a band with a size of 3,641 bp. Indeed, the band was visible for all strains. EcoRV cut untransformed ATΔ*virD2* genomic DNA did not result in any band (Fig. 4.21 B and D).

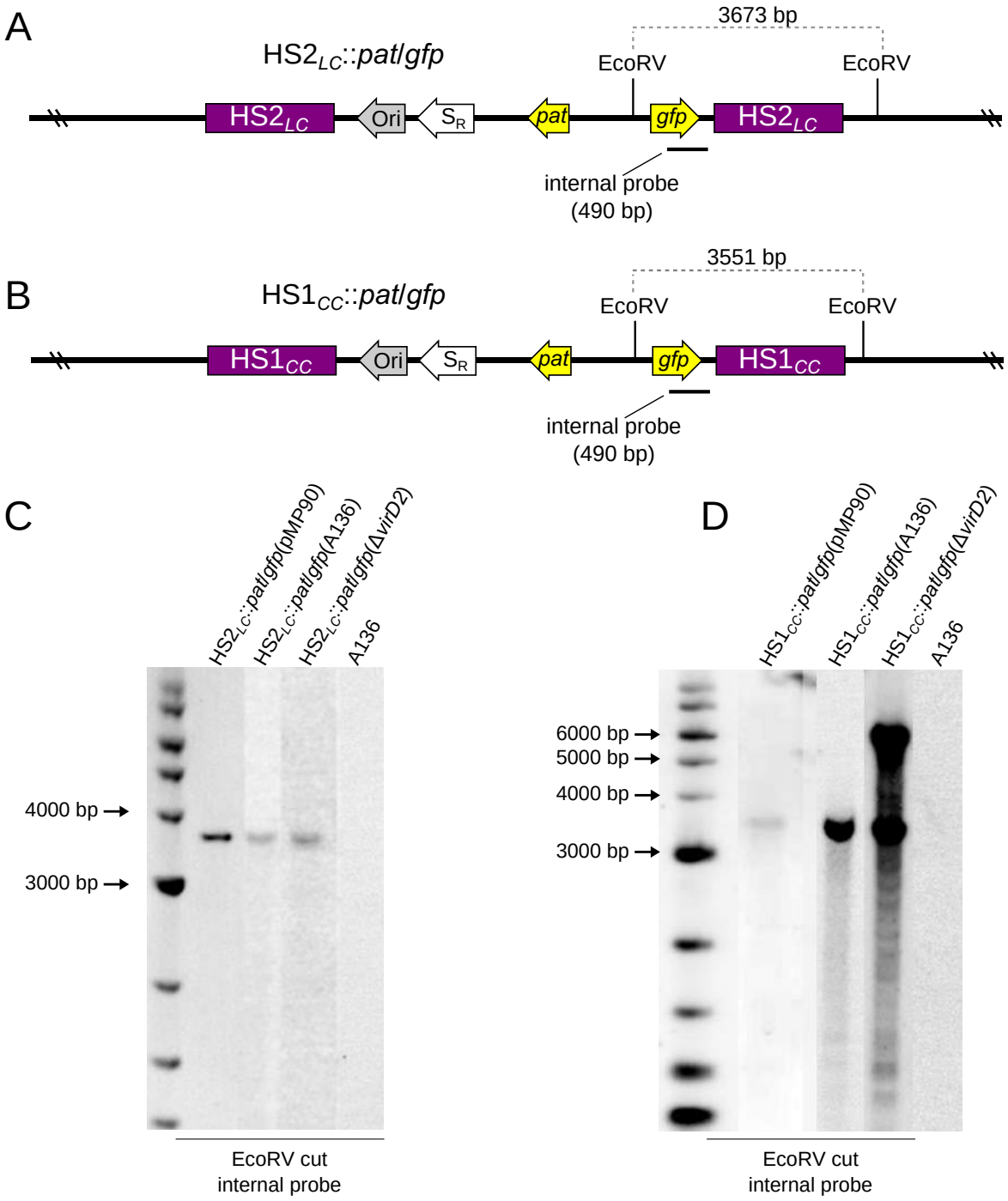


Figure 4.20: DNA blot analysis of insertions in $HS2_{LC}$ and $HS1_{CC}$ loci[2]. **A and B**: Strategy to confirm the correct insertions in $HS2_{LC}$ and $HS1_{CC}$. **C**: DNA blot analysis confirmed the correct integrations in $HS2_{LC}$ in all strains. **D**: The correct integrations in $HS1_{CC}$ in all strains were confirmed. For $HS1_{CC}$::*pat*/*gfp*(ATΔ*virD2*) a second, larger band was visible. Untransformed A136 cells did not lead to any band.

[2]For simplicity, blot pictures were combined. Original blots can be found in supplement (Fig. S1).

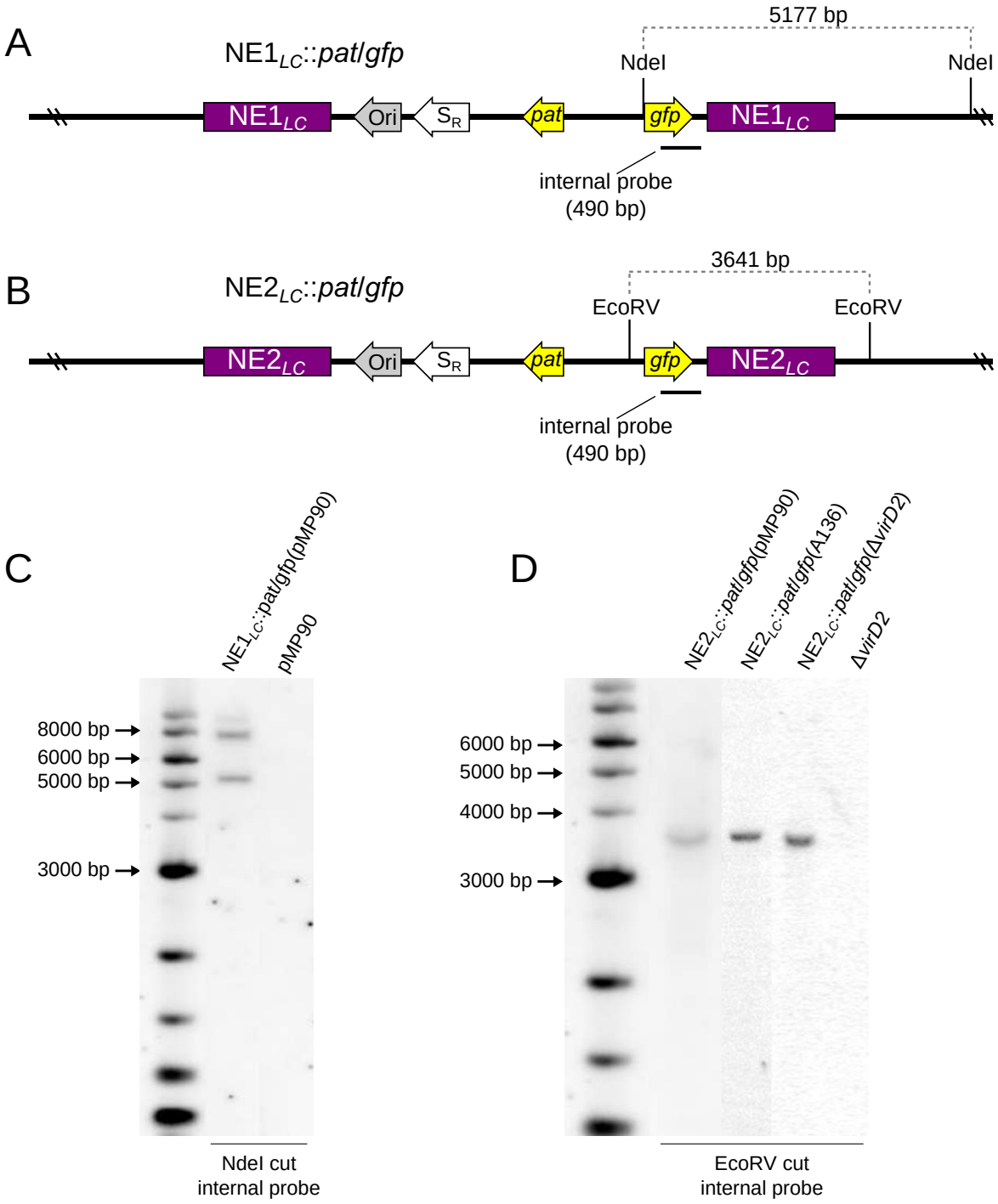


Figure 4.21: DNA blot analysis of insertions in NE1$_{LC}$ and NE2$_{LC}$ loci. A and B: Strategy to confirm the correct insertions in NE1$_{LC}$ and NE2$_{LC}$. **C**: DNA blot analysis confirmed the correct integration in HS1$_{CC}$::*pat*/*gfp*(ATΔ*virD2*). Additional larger bands were visible. **D**: The correct integration into the NE2$_{LC}$ locus in all strains was confirmed[3].

[3]For simplicity, blot pictures were combined. Original blots can be found in supplement (Fig. S1).

4.2.5 Detection of *A*chrDNA transfer by insertion of reporter genes into the bacterial chromosome

Strains containing the marker genes integrated into the chromosome were used to either infiltrate leaves of *N. benthamiana* or for transformation of *A. thaliana* plants by floral-dip. Infiltrated leaves were screened for GFP expressing cells. Since the bacteria did not contain any T-DNA or borders, plant cells could only express GFP if a transfer of chromosomal DNA with the integrated reporter genes took place. Analogous, the harvested seeds of dipped *A. thaliana* plants were screened by applying the herbicide BASTA®. Only plants to which the *pat* gene was transferred from the chromosome can survive selection.

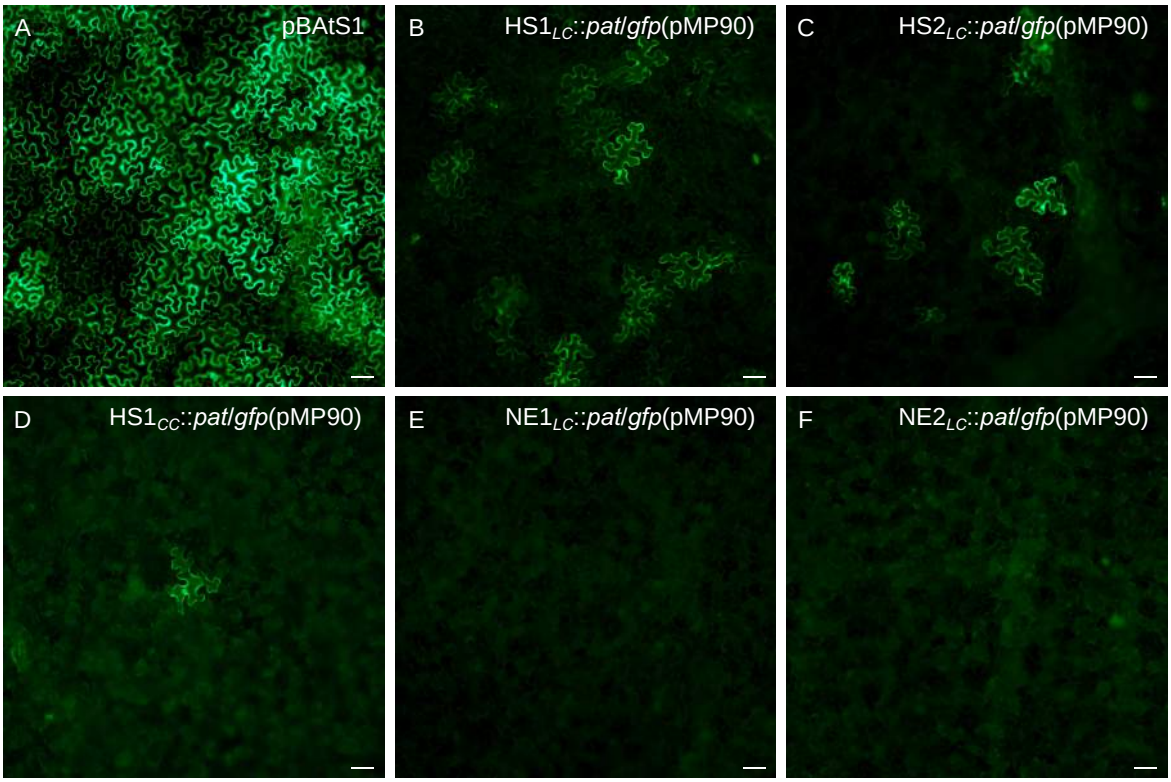


Figure 4.22: Results of *N. benthamiana* leaf infiltrations with GV3101(pMP90) cells having *gfp* inserted in different chromosomal loci: **A**: Positive control GV3101(pMP90) pBAtS1-GFP, *gfp* was located within the T-DNA with intact borders (normal case). **B**: *gfp* insertion in HS1$_{LC}$ locus. **C**: insertion in HS2$_{LC}$ locus. **D**: insertion in HS1$_{CC}$ locus. **E**, **F**: insertion in NE1$_{LC}$ and NE2$_{LC}$ locus. Note: in B, C, D, E and F *A. tumefaciens* cells do not have any T-DNA vector with borders. Scale bar: 50 μm.

To see if any of the *gfp* tagged chromosomal loci are transferred to plant cells, the five tagged GV3101(pMP90) strains were used for *N. benthamiana* leaf infiltra-

tion. Depending on the tagged locus, different transfer efficiencies could be observed. Infiltration of HS1$_{LC}$::*pat*/*gfp*(pMP90), representing the most prominent hotspots, resulted in the highest number of GFP expressing cells ($\approx$ 1 in 10 cells). In repetitions this number varied, probably depending on factors like the age and size of the leaf. Nevertheless, this strain always led to the highest number of GFP expressing cells (Fig. 4.22 B).

HS2$_{LC}$::*gfp*(pMP90) infiltrated leaves showed GFP expressing cells as well, indicating that bacterial chromosomal DNA was transferred. However, less leaf cells were expressing GFP ($\approx$ 1 in 50 cells; Fig. 4.22 C).

Infiltration of HS1$_{CC}$::*pat*/*gfp*(pMP90) resulted in roughly one out of 1,000 GFP expressing cells (Fig. 4.22 D). Both negative controls NE1$_{LC}$::*pat*/*gfp*(pMP90) and NE2$_{LC}$::*pat*/*gfp*(pMP90) did not lead to any GFP expression (Fig. 4.22 E and F).

These results indicate that transfer of chromosomal DNA took place from certain chromosomal regions in strain GV3101(pMP90). HS1$_{LC}$::*pat*/*gfp*(pMP90) and HS2$_{LC}$::*pat*/*gfp*(pMP90) thereby showed the highest transfer rate to the plant cell while *gfp* from the HS1$_{CC}$ locus was transferred to a much lesser extend. Both negative controls never led to a GFP expressing plant cell. Thus, the transfer of *A*chrDNA was successfully reproduced and the results by Ülker et al. (2008a) independently confirmed.

Since the homologous integration vectors also carry the *pat* gene for plant selection, the five GV3101(pMP90) based strains were also used to transform *A. thaliana* plants by floral dip. No stably transformed plant could be obtained. However, the floral dip assays were not done extensively (see discussion 5.4.5).

4.2.6 Vir-proteins are involved in the transfer of *A*chrDNA

To answer the question if the Vir proteins and in particular the T-DNA "pilot" protein VirD2 are involved in the transfer of *A*chrDNA, the five chromosomal *gfp* insertions were conducted in two additional *A. tumefaciens* strains. To test if the Vir proteins are involved in the *A*chrDNA transfer process, all five HR insertions were conducted in *A. tumefaciens* strain A136, which lacks the Ti-plasmid harbouring the *vir*-genes. In contrast to the insertions in strain GV3101(pMP90) (Fig. 4.23 A, D, G), leaves infiltrated with insertions in strain A136 did not show any GFP expressing cells, irrespective of the integration locus. This indicates a role of the *vir*-genes encoded on the Ti-plasmid in the transfer of *A*chrDNA (Fig. 4.23 B, E, H).

To test whether the T-DNA pilot protein VirD2 is involved in the transfer, an-

other *A. tumefaciens* strain was tagged with GFP in the loci described above. The ATΔ*virD2* strain is based on *A. tumefaciens* strain C58 but its *virD2* gene was completely deleted while the other *vir* genes are still present. Tobacco leaf infiltration with the different tagged loci in ATΔ*virD2* did not lead to any GFP expressing cells, suggesting that VirD2 is involved in the transfer of the tagged chromosomal regions (Fig. 4.23 C, F, I).

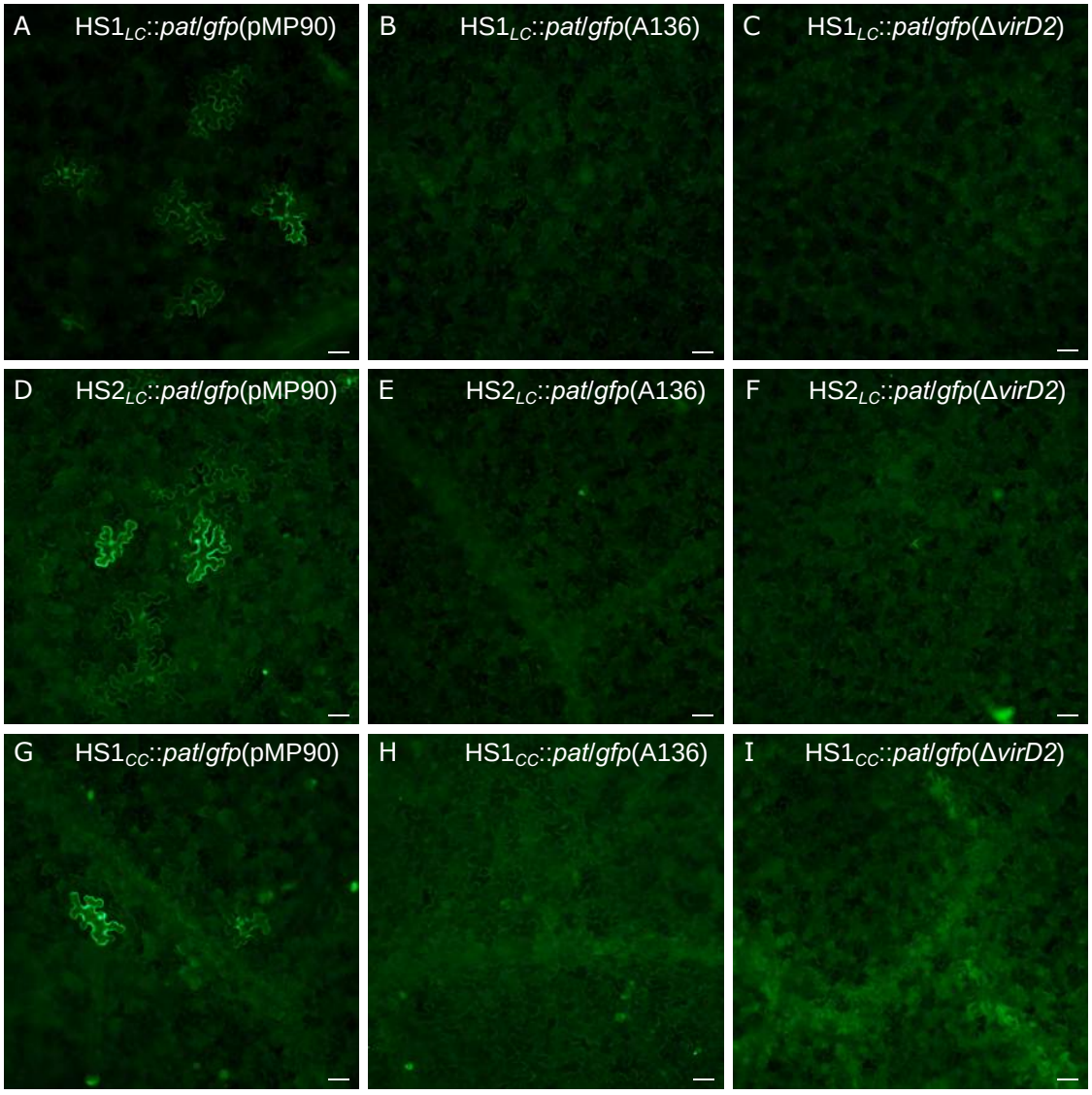


Figure 4.23: Results of the comparison of three different *A. tumefaciens* strains carrying *gfp* insertions. **A**, **D**, **G**: Different loci tagged in GV3101(pMP90) resulted in the transfer of *A*chrDNA. **B**, **E**, **H**: Loci tagged in Ti plasmid cured *A. tumefaciens* strain A136 did not lead to GFP expressing plant cells. **C**, **F**, **I**: Loci tagged in *A. tumefaciens* strain ATΔ*virD2* without *virD2* were not transferred to plant cells. Scale bar: 50 μm.

4.3 Identification of chromosomal sequences as starting points for DNA transfer

Since for some *gfp* tagged regions DNA transfer was observed ($HS1_{LC}$, $HS2_{LC}$ and $HS1_{CC}$) while for others it was not ($NE1_{LC}$ and $NE2_{LC}$), it is likely that certain chromosomal sequences in proximity to the hotspot regions are responsible for the transfer. Therefore, the next step was to find the exact sequences on the chromosomes which mediated the transfer of *A*chrDNA. For this reason, an assay was developed which allows the fast testing of different sequences for their ability to act as a starting point for the transfer of DNA to plant cells. Candidate sequences were PCR amplified from genomic DNA of *A. tumefaciens* and ligated into pBasicS1-GFP, a plasmid containing an origin of replication for *A. tumefaciens* and for *E. coli*, a spectinomycin resistance gene, a GFP expression cassette but no T-DNA borders. Due to lacking border sequences, a GV3101(pMP90) strain containing this plasmid is not able to mediate the transfer of plasmid DNA to plant cells and hence never results in GFP expressing cells when infiltrating *N. benthamiana* leaves. However, if a sequence which is able to mediate DNA transfer to plants is integrated into pBasicS1-GFP, the complete plasmid DNA, including the *gfp* expression cassette, will be transferred to plant cells and GFP expressing cells should be detectable. The empty pBasicS1-GFP plasmid was used as a negative control for all tested chromosomal DNA fragments.

4.3.1 Sequences covering the locus of integration mediate only a limited DNA transfer

In a first attempt to locate the exact sequence which was mediating the transfer of chromosomal DNA, PCR fragments covering all the loci which served as target for homologous recombination in section 4.2.3 were tested. Four 5 kbp long PCR fragments were amplified from *A. tumefaciens* strain C58 genomic DNA, including the loci of $HS1_{LC}$, $HS2_{LC}$, $HS1_{CC}$ and $NE2_{LC}$ with approximately 1 kbp overhang on each site (see Fig.4.24 A). The PCR fragment covering the $NE2_{LC}$ locus was used as a negative control, since integration of *gfp* in this locus did not result in GFP expressing plant cells in section 4.2.5. The PCR fragments were ligated into pBasicS1-GFP. All four PCR fragments had a size of approximately 5 kbp. The resulting vectors were termed p5kb-$HS1_{LC}$-GFP, p5kb-$HS2_{LC}$-GFP, p5kb-$HS1_{CC}$-GFP and p5kb-$NE2_{LC}$- GFP and used to transform GV3101(pMP90) cells.

These four different *A. tumefaciens* strains were infiltrated into *N. benthamiana* leaves followed by a screening for GFP expression. Similar to the results obtained in section 4.2.5, the 5 kbp fragments covering $HS1_{LC}$ and $HS2_{LC}$ led to GFP expressing cells. However, the amount of fluorescent cells was much lower compared to the strains where GFP was integrated into the bacterial chromosome by homologous recombination (Fig. 4.24 B and C). The strains covering $HS1_{CC}$ and $NE2_{LC}$ did not result in any GFP expressing cells (Fig. 4.24 D and E).

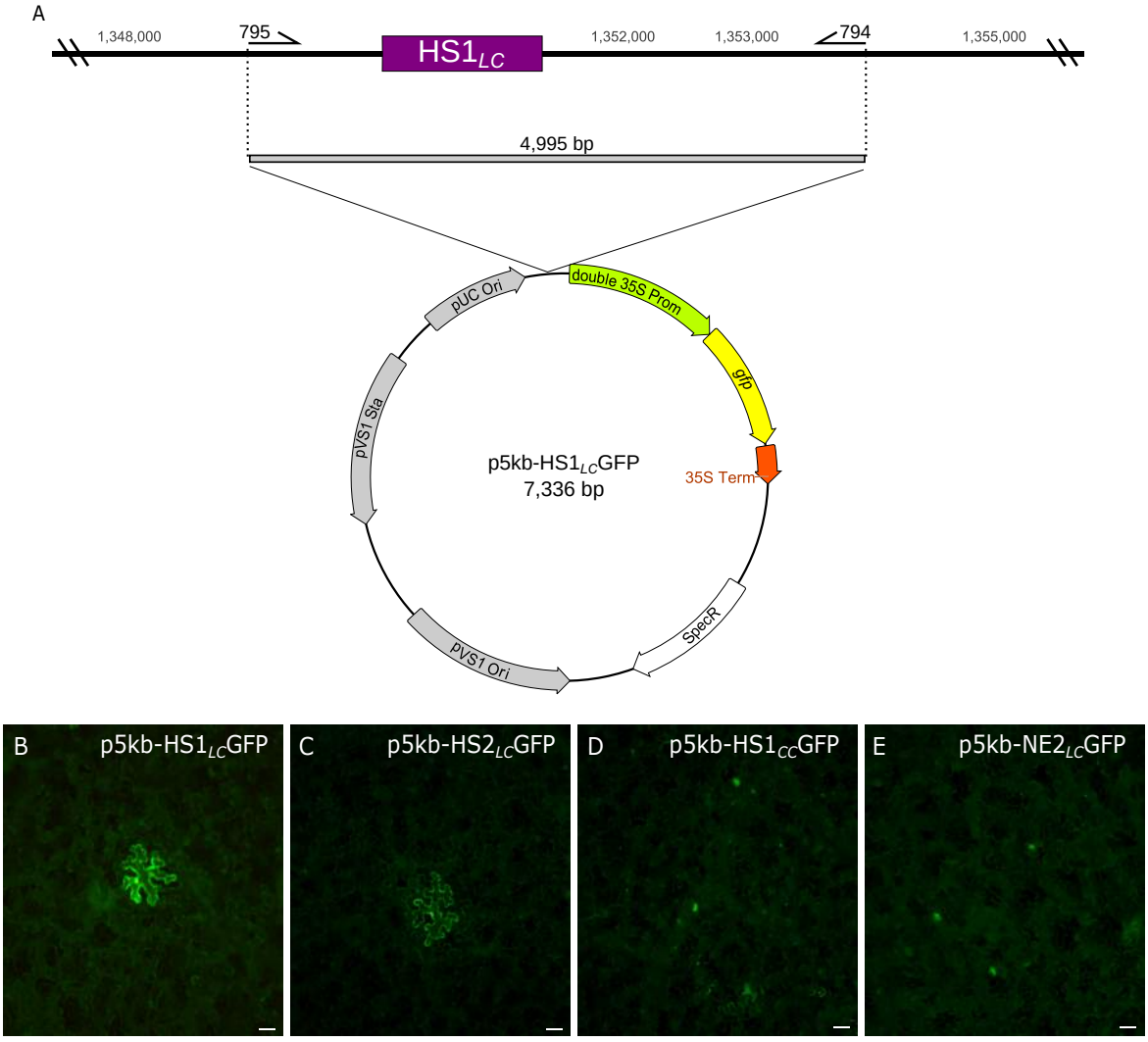


Figure 4.24: PCR Fragments covering the insertion regions were tested for their ability to transfer DNA. A: The $HS1_{LC}$ region was amplified and ligated to pBasicS1-GFP, leading to plasmid p5kb-$HS1_{LC}$-GFP. The same was done for the other insertion loci. **B - E**: Results of the infiltration assay using the different plasmids. p5kb-$HS1_{LC}$-GFP and p5kb-$HS2_{LC}$-GFP resulted in GFP expressing plant cells, while p5kb-$HS1_{CC}$-GFP and p5kb-$NE2_{LC}$- GFP did not. Scale bar: 50 μm.

If p5kb-HS1$_{LC}$-GFP and p5kb-HS2$_{LC}$-GFP would actually contain the sequence responsible for the transfer of the integrated reporter genes in section 4.2.3, one would expect the amount of GFP positive leaf cells to be as high as in the strains having the *gfp* inserted in the respective locus. It could be even higher because the sequence is now located on a plasmid which has a copy number of 7 - 9 copies per cell (Lee and Gelvin, 2008), instead of only one copy in the chromosome. The fact that the amount of fluorescent cells is lower suggests that another sequence, not included in the 5 kbp PCR fragments, is involved in the transfer.

4.3.2 Candidate sequences on the linear chromosome

To find more candidate sequences, an *in silico* BLAST analysis was conducted. The chromosomes of *A. tumefaciens* C58 were screened for sequences which show homology to the Ti-plasmid's right and left borders (border-like) or to origin of transfer consensus sequences of the IncP or IncQ type (oriT-like), which are normally involved in plasmid conjugation. The oriT-like sequences were included in the screening because of their sequence homology to the T-DNA borders (Pansegrau and Lanka, 1991; Waters et al., 1991; Pansegrau et al., 1993; Lessl and Lanka, 1994). Furthermore, it has been shown that VirD2 can cleave oriT sequences (Pansegrau et al., 1993) and that the VirD1/VirD2 complex can mediate DNA transfer to plant cells starting from oriT sequences (Dube et al., 2004).

BLAST analysis resulted in no perfect RB, LB or oriT sequences on the chromosomes. Depending on the mismatches allowed, hundreds to thousands of sequences with a similarity between 60 % to 80 % were detected. Since these were too many to test, the locations of the sequences was included in the choice. Candidate sequences responsible for the transfer of *A*chrDNA should be in close proximity to the hotspots of *A*chrDNA transfer, found in the T-DNA insertion lines by Ülker et al. (2008a).

Based on these assumptions nine different chromosomal sequences, showing homology to either an origin of transfer or the T-DNA borders, were tested for their ability to act as a starting point for *A*chrDNA transfer to plant cells. These sequences were PCR amplified and ligated into pBasicS1-GFP, analogous to the testing of the 5 kbp sequences in section 4.3.1. The empty plasmid pBasicS1-GFP was used as a negative control. *A. tumefaciens* containing the plasmid with the respective sequence to test was used to infiltrate leaves of *N. benthamiana*. Detection of GFP in leaf cells would show the ability of the sequence to mediate the transfer of DNA from the plasmid to the plant cell. Figure 4.25 shows the location of the candidate sequences of the linear

chromosome relative to the HR vector insertion loci from section 4.2.3, the hotspots of frequent *A*chrDNA transfer as well as both IS*426* copies.

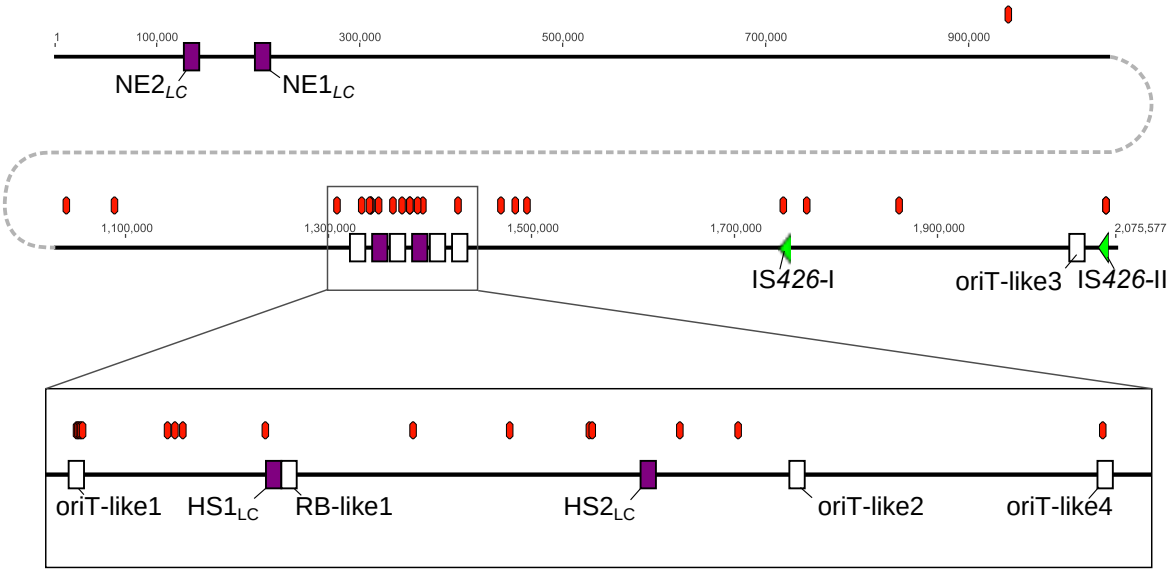


Figure 4.25: Linear chromosome with reporter gene insertion sites and putative candidate sequences mediating T-DNA transfer. Purple squares indicate the positions of the different regions used for reporter gene integration by HR. White squares show the position of putative sequences mediating the transfer of *A*chrDNA. Red diamonds indicate hotspots of *A*chrDNA transfer, as described by Ülker et al. (2008a). Green triangles indicate the position of the two IS*426* copies.

RB-like1 from GABI-Kat line 086C02

In their study Ülker et al. (2008a) found a region on the linear chromosome which showed homology to the Ti-plasmid's RB and was located in close proximity to the most prominent hotspots of *A*chrDNA transfer. This region even showed homology to the overdrive region found next to many RBs (Fig. 4.26 A). This "RB-like" sequence was found in GABI-Kat T-DNA insertion line 086C02 (Ülker et al., 2008a). During a normal T-DNA transfer, the RB is cleaved between nucleotide 3 and 4. Intriguingly, the cleavage side in the RB-like sequence found in the *A. thaliana* mutant line also appeared to be cleaved at exactly this position (Fig. 4.26 B), making it a good candidate for a sequence with the potential to mediate the transfer of chromosomal DNA to plant cells.

The RB-like1 sequence was amplified and ligated two times in pBasicS1-GFP, upstream and downstream of the *gfp* gene in direct orientation, to mimic both RB and

LB of the Ti-plasmid (Fig. 4.26 C). However, after *A. tumefaciens* transformation and *N. benthamiana* infiltration no GFP expressing cells could be detected (Fig.4.26 D).

In order to analyse if DNA gets transferred during a stable *A. thaliana* transformation, the *gfp* gene was exchanged to a *pat* gene which allows screening for stably transformed *A. thaliana* plants by BASTA® selection. Transformation with pBasicS1-PAT-RB-like1 did not lead to herbicide resistant plants. This indicates that the RB-like1 sequence by itself did not mediate DNA transfer to plant cells.

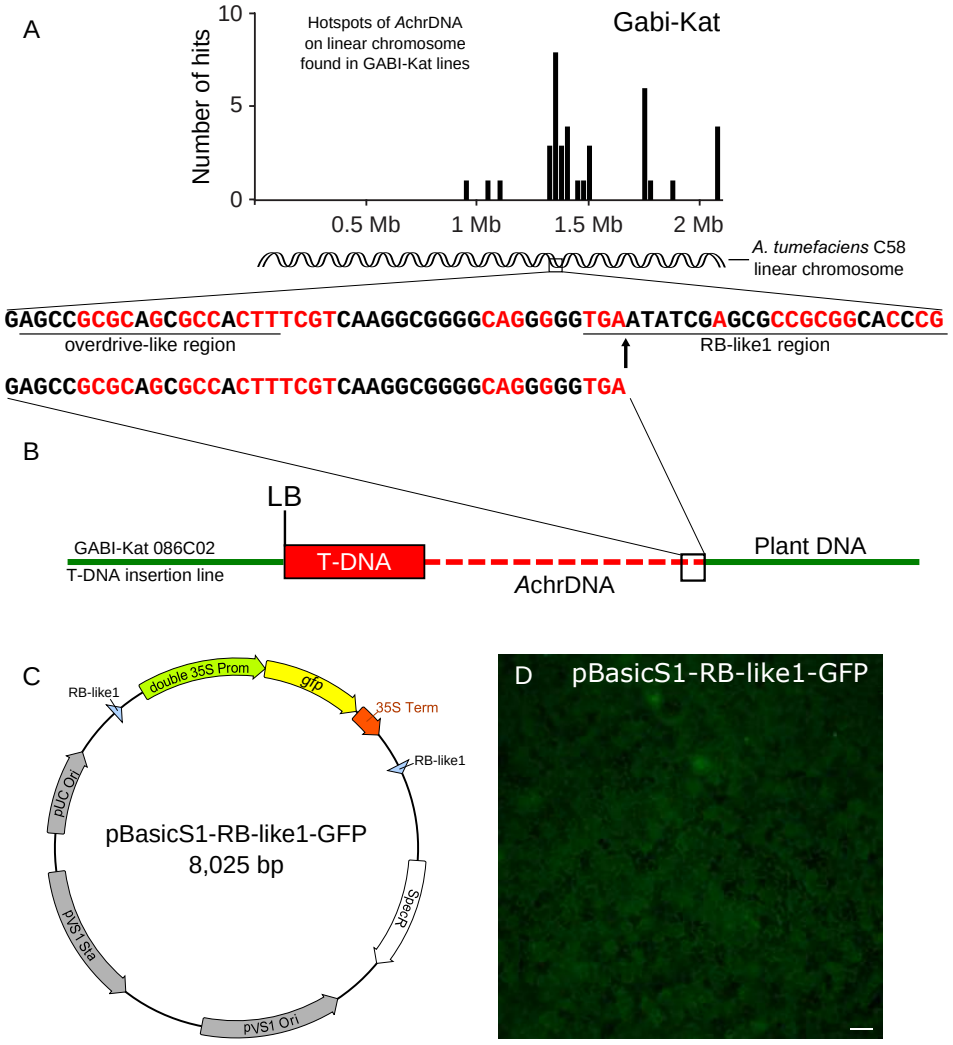
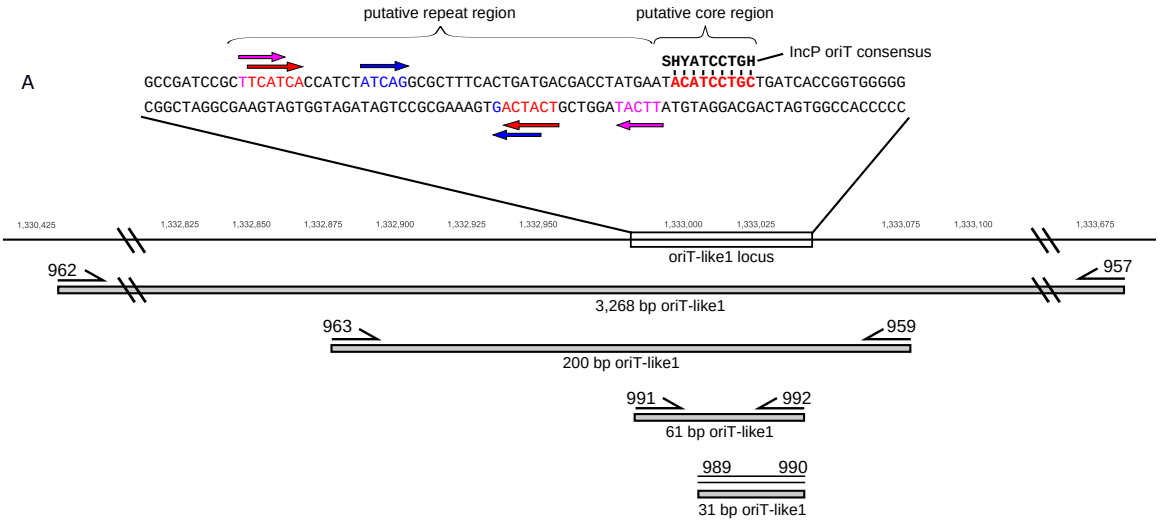


Figure 4.26: Testing of RB-like1 to mediate DNA transfer. **A**: Location of the RB-like1 sequence on the linear chromosome (Ülker et al., 2008a). Overdrive-like sequence and RB-like sequence are underlined. Red nucleotides show homology to either the RB of pTi-C58 or the RB consensus sequence. The arrow indicates the position where a normal RB would be nicked. **B**: Sequence found in GABI-Kat line 086C02. Drawings adapted from Ülker et al. (2008a). **C**: RB-like1 testing plasmid. **D** Infiltration of pBasicS1-RB-like1-GFP in *N. benthamiana* leaves. Scale bar: 50 μm.

The oriT-like1 sequence

By screening the linear chromosome for a sequence homologous the consensus motif of an IncP oriT, the oriT-like1 sequence was discovered. It was located 17 kbp upstream of $HS1_{LC}$ within the coding sequence of an ABC transporter (Atu4209). Interestingly, this coding sequence was found in five independent GABI-Kat *A. thaliana* T-DNA insertion lines by Ülker et al. (2008a) and thus constitutes a hotspot of *A*chrDNA transfer.



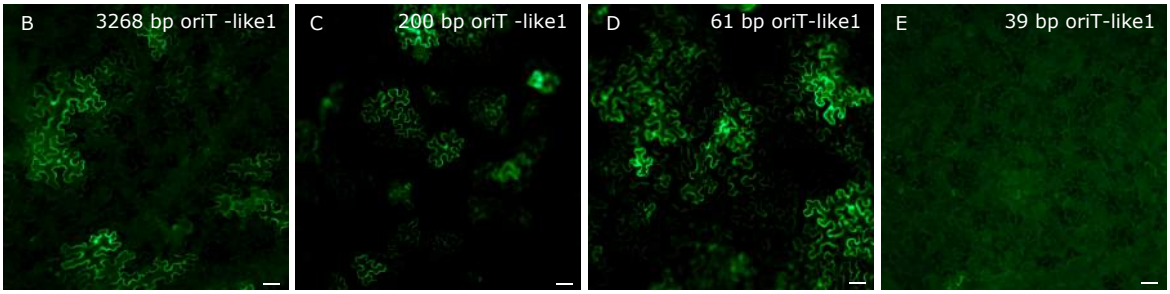


Figure 4.27: Testing of the oriT-like1 sequences to mediate DNA transfer. Arrows indicate primer binding sites. **A**: The oriT-like1 region and amplified PCR products for insertion in pBasicS1-GFP. Potential inverted repeat sequences are indicated by coloured arrows. Core region is indicated by bold letters. **B** - **E**: Results of the testing of the different fragments to mediate DNA transfer to plant cells. Except for the 31 bp fragment, all fragments led to a similar amount of GFP expressing plant cells. Scale bar: 50 μm.

Close to the oriT-like1 core region, short inverted repeat sequences could be detected, which is a typical feature of oriTs. Because of its orientation in the chromosome it would mediate the transfer of regions which are located downstream and thus might be responsible for the transfer of $HS1_{LC}$. A 3,268 bp PCR fragment containing

oriT-like1 with extensive coverage of the upstream and downstream sequences was amplified (Fig. 4.27 A). The fragment was ligated into plasmid pBasicS1-GFP and used to transform GV3101(pMP90) cells. The *N. benthamiana* infiltration assay was used to asses the transfer potential of this sequence.

Figure 4.27 B shows the result of the assay. The 3,268 bp oriT-like1 fragment led to GFP expressing cells in a comparable amount to the HS1$_{LC}$::*pat*/*gfp*(pMP90) strain. Thus, it is possible that the sequence responsible for the transfer is located within the 3,268 base pairs.

To further narrow down the starting sequence, three additional PCR fragments were generated and ligated into pBasicS1-GFP. A 200 bp PCR fragment covered the core region, the repeat region and short sequences upstream and downstream. Additionally, a 60 bp PCR fragment covering exactly the core and the repeat region and a 39 bp fragment only covering the core region was generated (Fig. 4.27 A). The resulting plasmids were used to transform GV3101(pMP90) cells followed by the *N. benthamiana* leaf infiltration assay.

As can be seen in figure 4.27 C and D, the shorter sequences led to a comparable amount of GFP expressing leaf cells as the 3,268 bp fragment did. Thus, the 61 bp fragment seems to harbour the sequence necessary for the transfer. The fragment which only contains the core region but not the repeats did not lead to any GFP expressing leaf cells (Fig. .4.27 E), suggesting that the repeat region is a necessary prerequisite for the transfer.

Since the oriT-like1 region was able to mediate DNA transfer to plant cells, more effort was put into testing oriT-like sequences.

The oriT-like2 sequence

Another sequence with the putative ability to mediate the transfer of *A*chrDNA was located on the linear chromosome within a gene called *accC* (Atu4273, Fig. 4.28 A). It showed homology to the oriT (IncP) consensus sequence. The sequence was located approximately 12 kbp downstream of the $HS2_{LC}$ insertion locus, flanking the accumulation of hotspots (Fig. 4.25). Because of its orientation it would mediate the transfer of DNA upstream of it and thus might be involved in transfer of the $HS2_{LC}$ locus.

A 318 bp PCR fragment was amplified and ligated to pBasicS1-GFP followed by *N. benthamiana* infiltration. No GFP expressing cells could be detected (Fig. 4.28 C).

oriT-like3 sequence in a conjugation locus

It was reported by Leloup et al. (2002) that besides the virulence T4SS and the two systems involved in conjugation of Ti- and the At-plasmid (Trb and AvhB), the linear chromosome harbours a 4^{th} cryptic locus with putative DNA mobilizing abilities. It contains a gene encoding for a homologue to the coupling protein VirD4 and other genes which show similarity to known bacterial transfer and conjugation systems. In their study, the authors did not assess if this cryptic conjugation locus is active or not. They also reported the existence of an oriT-like region, located between orf1 and orf2 of the cryptic transfer locus but did not analyse its function either (Leloup et al., 2002). Ülker et al. (2008a) did not report the occurrence of any hotspots of *A*chrDNA transfer close to this locus. Nevertheless, using primers 1048 and 1049, a 551 bp fragment containing the oriT-like3 sequence was amplified and ligated to pBasicS1-GFP, followed by the *N. benthamiana* infiltration assay. No GFP expression could be detected by microscopy of the leaf (Fig. 4.28 B and D).

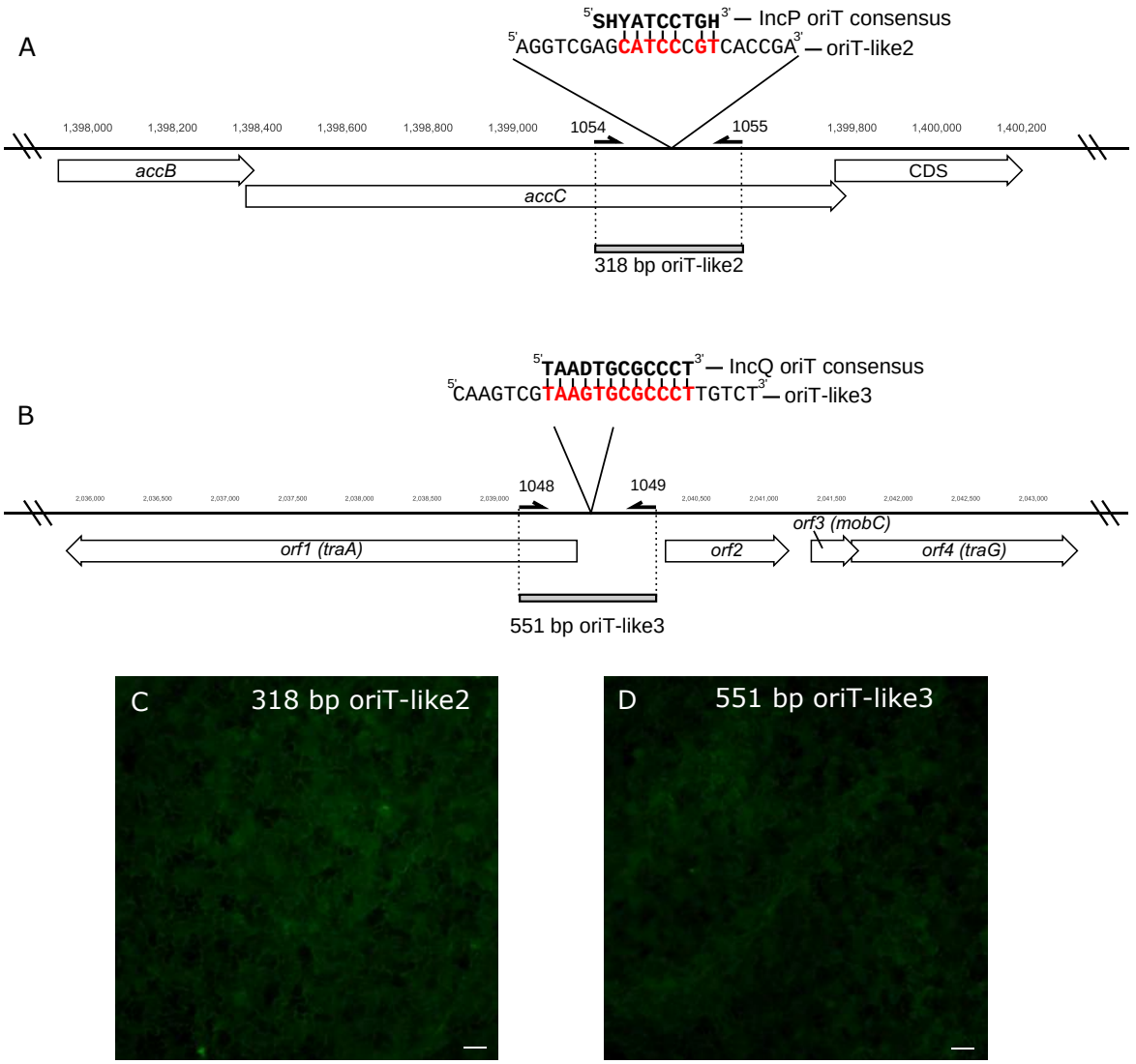


Figure 4.28: Testing of oriT-like2 and 3 to mediate DNA transfer. Arrows indicate primer binding sites. **A**: The oriT-like2 region was PCR amplified and resulted in a 318 bp fragment. **B**: The oriT-like3 region from the cryptic conjugation locus was amplified and resulted in a 551 bp fragment. **C** and **D**: Both fragments did not lead to any GFP expressing plant cells. Scale bar: 50 μm.

The oriT-like4 sequence

Two oriT-like sequences were located on the linear chromosome in a conserved hypothetical protein CDS (Atu4304; Fig. 4.29 A). Both showed homology to the IncP oriT consensus sequence. Intriguingly, parts of this conserved hypothetical protein CDS were detected by Ülker et al. (2008a) in the GABI-Kat T-DNA insertion mutant line 133G07. Thus, one of the two oriT-like sequence (referred to as oriT-like4) was potentially involved in the *A*chrDNA transfer to this line. To test this, three PCR fragments were amplified comprising the potential origin of *A*chrDNA transfer. In one 904 bp fragment the complete locus was covered. To further narrow down the putative origin of transfer, the region was split into two shorter fragments, each containing one oriT-like sequence (Fig. 4.29 A). The three fragments were ligated into pBasicS1-GFP and the infiltration assay was performed. Non of the three constructs led to GFP expression in *N. benthamiana* leaf cells (Fig. 4.29 B, C and D).

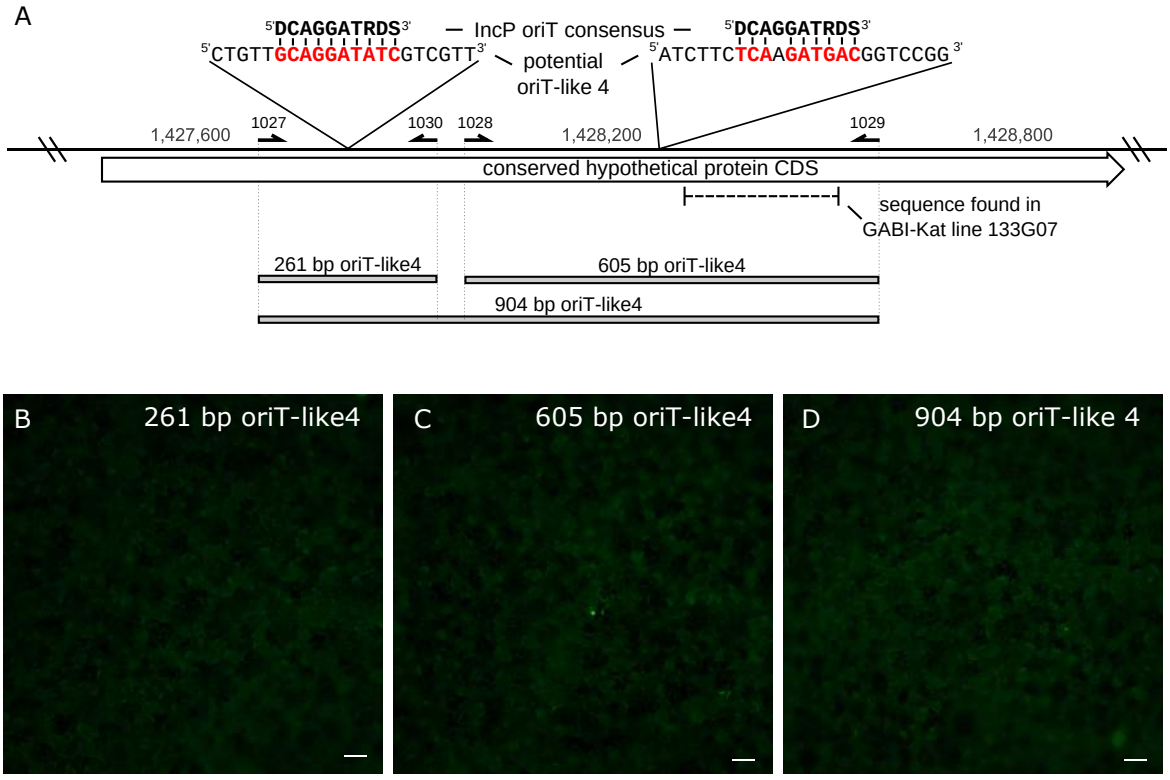


Figure 4.29: Testing of oriT-like4 to mediate DNA transfer Arrows indicate primer binding sites. **A**: Three fragments covering the oriT-like4 sequence were amplified. **B**, **C** and **D**: Non of the three fragments was able to mediate DNA transfer to plant cells. Scale bar: 50 μm.

4.3.3 Candidate sequences on the circular chromosome

Ülker et al. (2008a) analysed the hotspots of frequently transferred *A*chrDNAs of the linear chromosome but did not assess the ones on the circular chromosome. Nevertheless, they could show that also *A*chrDNA originated from the circular chromosome was transferred to plant cells. Additionally, in this work it was shown that a *gfp* reporter gene inserted into the $HS1_{CC}$ locus on the circular chromosome resulted in GFP expressing plant cells after leaf infiltration, indicating that *A*chrDNA transfer from the circular chromosome takes place (see 4.2.5).

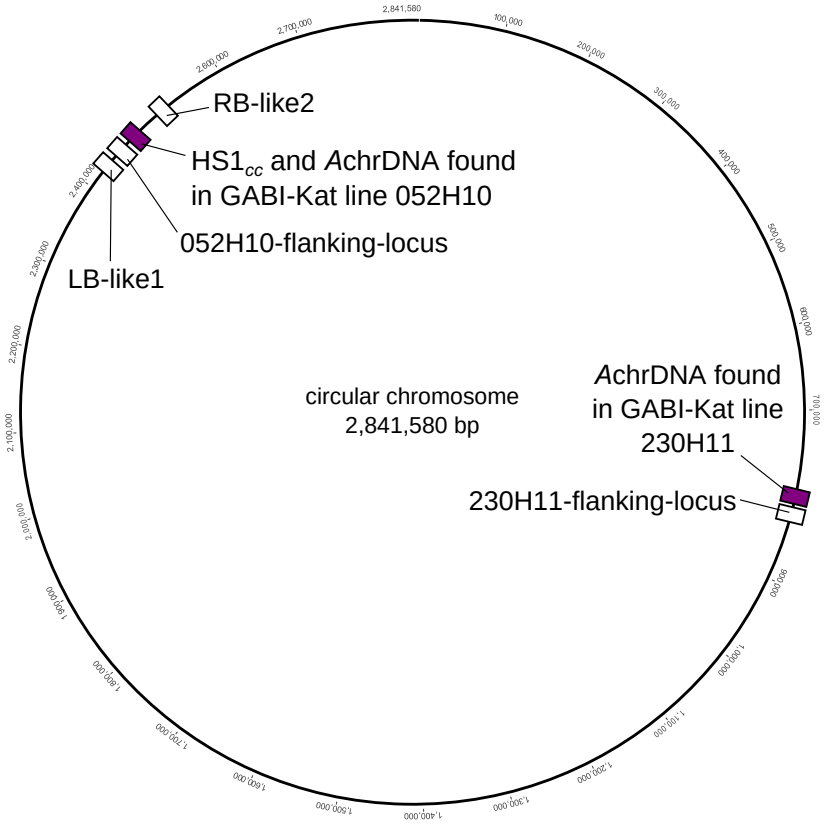


Figure 4.30: Circular chromosome with *pat*/*gfp* insertion sites and putative candidate sequences mediating T-DNA transfer. Purple squares indicate the positions of the region used for reporter gene integration by HR and sequences which were found in GABI-Kat T-DNA insertion lines by Ülker et al. (2008a). White squares show the positions of putative sequences mediating the transfer of *A*chrDNA that were tested in this work.

To understand which region on the circular chromosome is mediating *A*chrDNA transfer, the chromosome was screened for sequences homologous to the T-DNA borders and oriT consensus sequences. Four candidate sequences were tested for their ability to mediate the transfer from *A. tumefaciens* to plant cells (Fig. 4.30).

Testing of the 230H11-flanking-locus

In order to find the sequence responsible for the transfer of an *A*chrDNA fragment to GABI-Kat line 230H11, the area flanking the *A*chrDNA fragment was tested. This region did not contain a specific sequence which showed extensive homology to the Ti-borders or oriTs but harboured many which showed only weak homologies ($\approx 65\,\%$ identity). A large 2,344 bp fragment was amplified which covered many of these sequences (Fig. 4.31 A). However, when this fragment was integrated into pBasicS1-GFP and the infiltration assay was performed, no GFP expressing cells were visible (Fig. 4.31 D).

LB-like1 and and the 052H10-flanking-locus

In an attempt to locate the sequence which was responsible for the *A*chrDNA transfer from the circular chromosome to GABI-Kat line 052H10, two different sequences were tested.

The first candidate sequence was found approximately 30 kb upstream from the *A*chrDNA fragment transferred to *A. thaliana* T-DNA insertion line 052H10 and had 68 % identity with the T-DNA left border (LB-like1; Fig. 4.31 B). The second sequence was chosen because of its location directly flanking the chromosomal DNA which was transferred to GABI-Kat line 052H10. Thus, it potentially harboured a sequence responsible for the transfer to this GABI-Kat line (052H10-flanking-locus; Fig. 4.31 C). This region did not contain a specific sequence which showed extensive homology to the Ti-borders or oriTs but harboured many which showed weak homologies.

Using these sequences for the *A*chrDNA transfer assay, no GFP expression was induced in the *N. benthamiana* leaf cells upon infiltration (Fig.4.31 E, F)

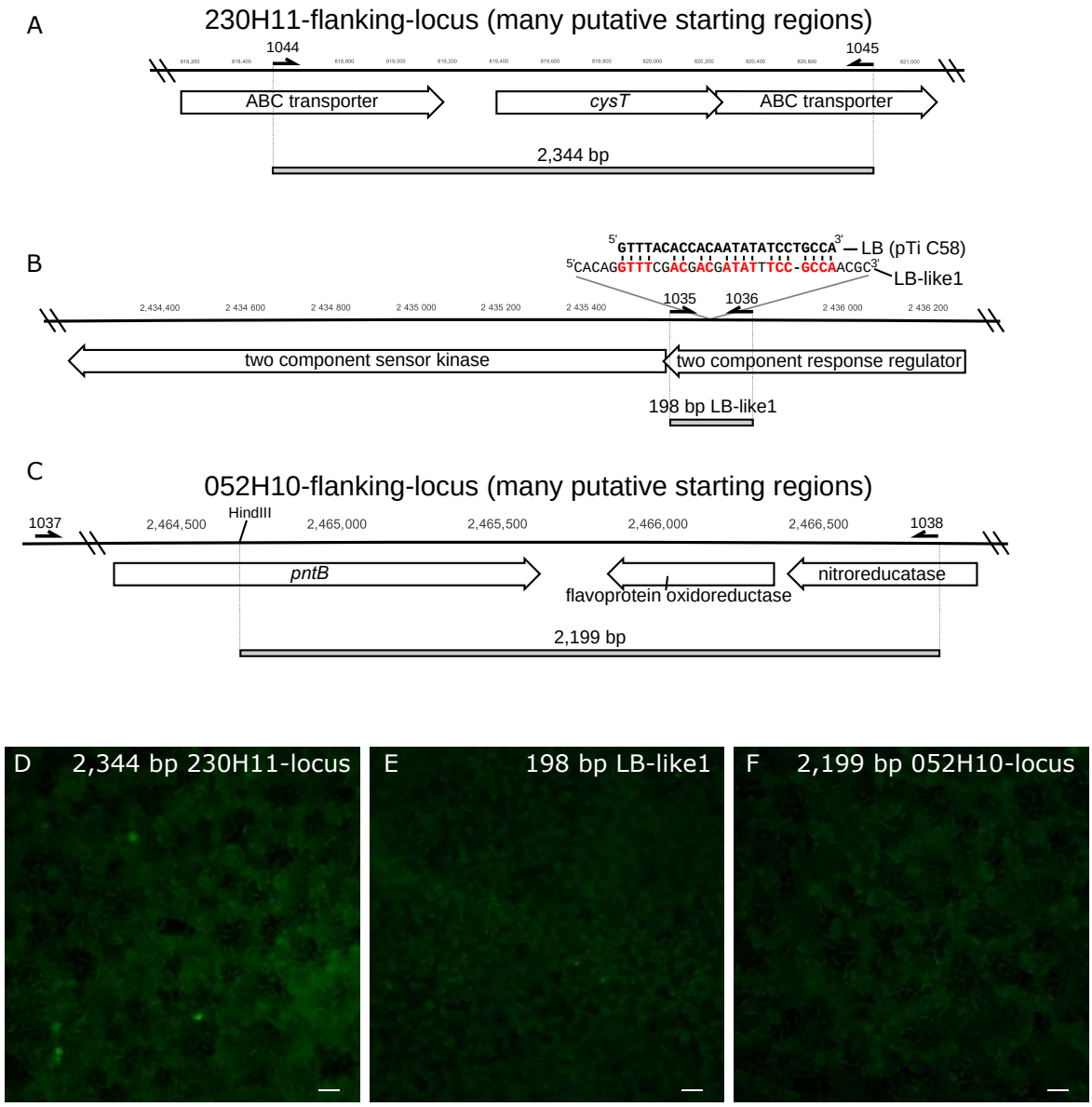


Figure 4.31: Testing of 230H11-locus, LB-like1 and 052H10-locus. Arrows indicate primer binding sites. **A**: Testing of 230H11-flanking-locus by amplifying a 2,344 bp fragment. **B**: A 198 bp LB-like fragment was tested. **C**: A large sequence flanking the 052H10 locus was tested for harbouring a sequence able to mediate DNA transfer. **D**, **E** and **F**: Non of the three fragments mediated DNA transfer to plant cells. Scale bar: 50 μm.

RB-like2

The last tested sequence (RB-like2) showed homology to the Ti-plasmid's RB. This sequence was chosen because it was located directly flanking an *A*chrDNA region, found in GABI-Kat line 052H10. The sequence was located in a gene called *purC* (Atu2548) approximately 46 kb downstream of the $HS1_{CC}$ locus. A 221 bp fragment was generated, covering the RB-like2 region with extensive overhangs. A second fragment of 31 bp, only including the putative core region, was generated (Fig. 4.32 A). The fragments were ligated into pBasicS1-GFP. *N. benthamiana* infiltration led to GFP expressing plant cells. Both sequences resulted in a comparable amount of GFP positive leaf cells (≈ one out of 1,000 cells showing fluorescence), indicating that the 31 bp fragment harbours the sequence responsible for the transfer (Fig. 4.32 B and C). However, compared to the oriT-like1 sequence the number of GFP expressing cells was much lower.

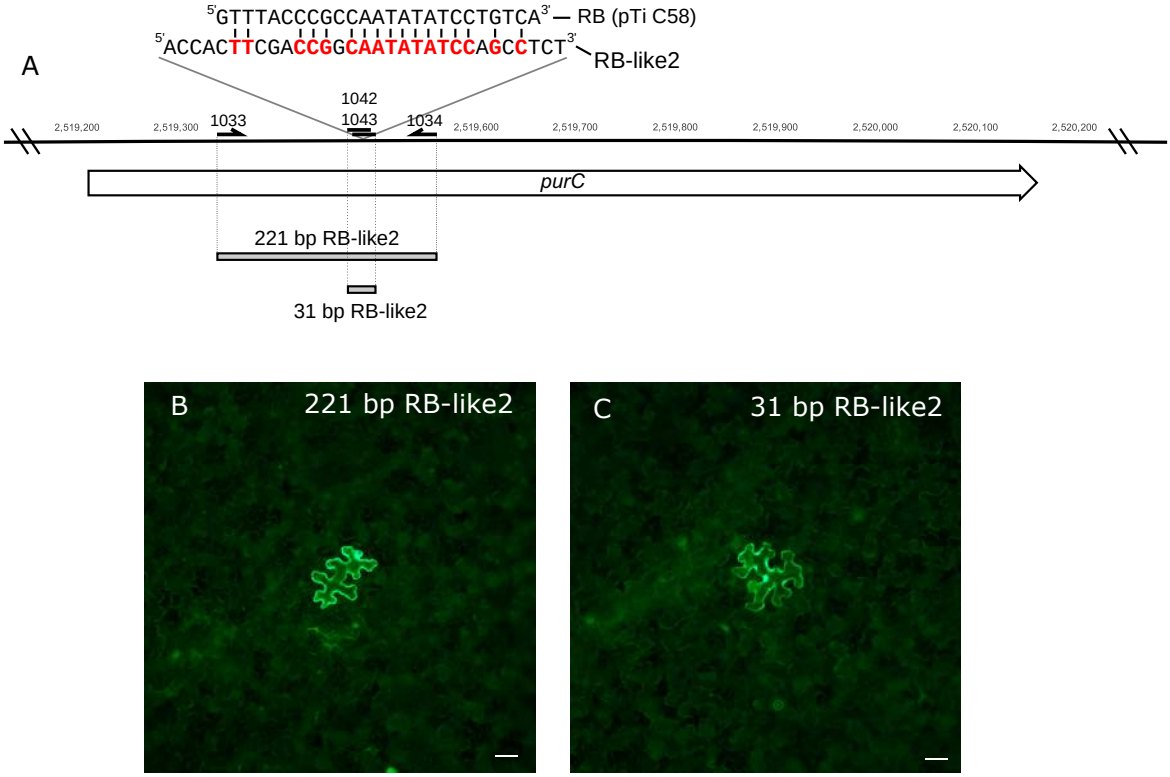


Figure 4.32: RB-like2 is able to mediate DNA transfer to plant cells Arrows indicate primer binding sites. **A**: Two fragments covering the RB-like2 sequence were tested for having the ability to mediate DNA transfer to plant cells. **B** and **C**: Both fragments were able to mediate DNA transfer. Scale bar: 50 μm.

4.3.4 VirD2 is involved in the transfer starting from oriT-like1 and RB-like2

In the previous experiments, two sequences (oriT-like1 and RB-like2) were identified which were able to mediate the transfer of DNA from a plasmid to plant cells. To further characterize the transfer and the proteins involved in it, *A. tumefaciens* strain ATΔ*virD2* was transformed with the pBasicS1-GFP-oriT-like1(200 bp) and pBasicS1-GFP-RB-like2(31 bp). In *A. tumefaciens* strain GV3101(pMP90) both plasmids led to GFP expressing *N. benthamiana* cells (Fig. 4.27 C; Fig. 4.32 B) However, when the assay was performed with strain ATΔ*virD2*, no fluorescence could be detected (Fig. 4.33). These results point towards an involvement of VirD2 in the transfer of plasmid DNA mediated by these two sequences .

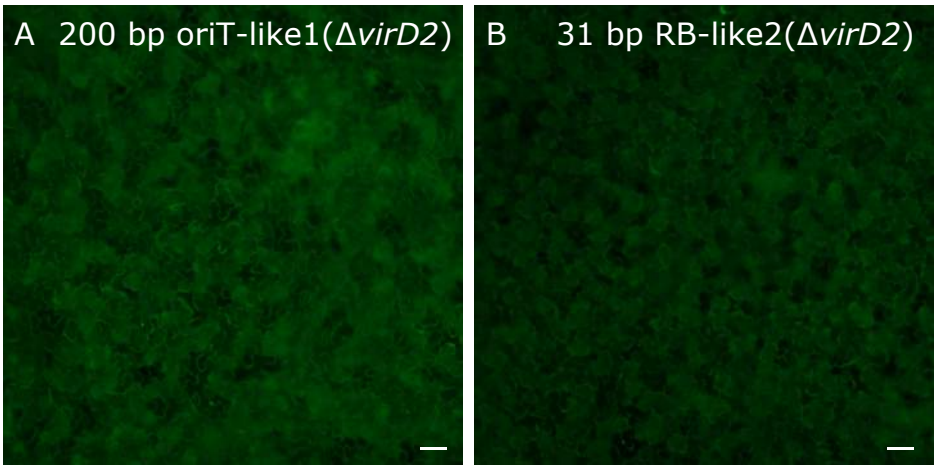


Figure 4.33: Testing of oriT-like1 and RB-like2 to mediate DNA transfer to plant cells in the absence of VirD2. A and **B**: No GFP expressing plant cells can be detected after transformation with pBasicS1-GFP-oriT-like1(ATΔ*virD2*) and pBasicS1-GFP-RB-like2(ATΔ*virD2*). Scale bar: 50 μm.

4.4 Activity of bacterial promoters in plant cells

Since *A. tumefaciens* transfers *A*chrDNA to plant cells, the question arises if this transfer is beneficial for the bacterium. As discussed in more detail in section 5.7, many *A*chrDNA fragments found in plants harboured coding sequences for ABC (ATP binding cassette) transporters. It was hypothesized that these ABC transporters might be involved in the plant transformation process, for example by exporting opines from

the plant cells. A necessary prerequisite for this would be that the bacterial encoded genes are expressed in plant cells.

Parts of one ABC transporter (Atu4209) were found by Ülker et al. (2008a) in five different T-DNA insertion lines. The fact that this gene was often observed in plant cells might indicate that it also gets expressed there.

The promoter sequences of this ABC transporter and the one from an upstream located second ABC transporter (Atu4208) were tested for their activity in plant cells (Fig. 4.34 A). The promoter of a third gene (*atsA*; Atu4255) was also tested (Fig. 4.34 B). *atsA* was found in two independent T-DNA insertion lines and was reported to be potentially involved in virulence and attachment of the bacterium to host cells (Matthysse et al., 2000).

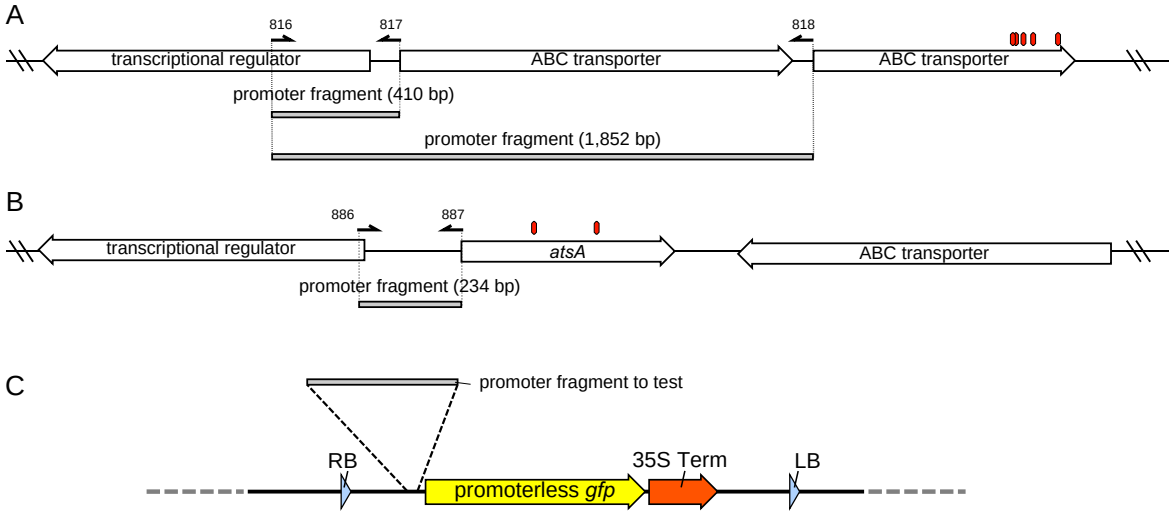


Figure 4.34: Bacterial promoter sequences were tested for their activity in plant cells. Arrows indicate primer binding sites. **A**: Promoter sequences of two adjacent ABC transporters were amplified. **B**: Promoter sequence of *atsA* was amplified. **C**: Promoter sequences were inserted in front of a promoterless *gfp* gene. Red diamonds indicate sequences which were transferred to GABI-Kat T-DNA insertion lines analysed by Ülker et al. (2008a).

The promoter sequences were amplified and ligated in front of a promoterless *gfp* gene, which was located in the T-DNA region of pBAtS1 (Fig. 4.34 C). Plasmids were used to transform *A. tumefaciens* GV3101(pMP90) cells which were subsequently infiltrated into leaves of *N. benthamiana*. Since both T-DNA borders were present on the vector, the *gfp* gene with the test-promoter sequence was transferred to plant cells. If the promoter is active, GFP expressing cells should be visible. However, infiltration of the three constructs did not lead to GFP expressing cells (data not shown). This indicates that the promoter is either not active in plant cells or that its activity is too

weak to result in enough GFP to be detectable.

4.5 Deletion of oriT-like1 nearly eliminates *A*chrDNA transfer

When comparing the strain carrying the reporter gene insertions in the $HS1_{LC}$ locus ($HS1_{LC}$::*pat*/*hfp*(pMP90)) to the strain having the oriT-like1 sequence on a plasmid (pBasicS1-GFP-oriT-like1), similar amounts of fluorescent cells could be observed (Fig. 4.22 and 4.27). It was further analysed if the oriT-like1 sequence was actually the sequence responsible for the transfer of the $HS1_{LC}$ locus. For this reason, the oriT-like1 sequence was deleted from the genome of the $HS1_{LC}$::*pat*/*hfp*(pMP90) strain. Similar to the deletion of IS*426*, the oriT-like1 sequence was replaced by the kanamycin resistance gene *nptII* by homologous recombination. To achieve this, the plasmid pKO-oriT-like1 was generated. This suicide vector cannot be replicated in *A. tumefaciens* and carries the *nptII* gene conferring kanamycin resistance between two homology regions.

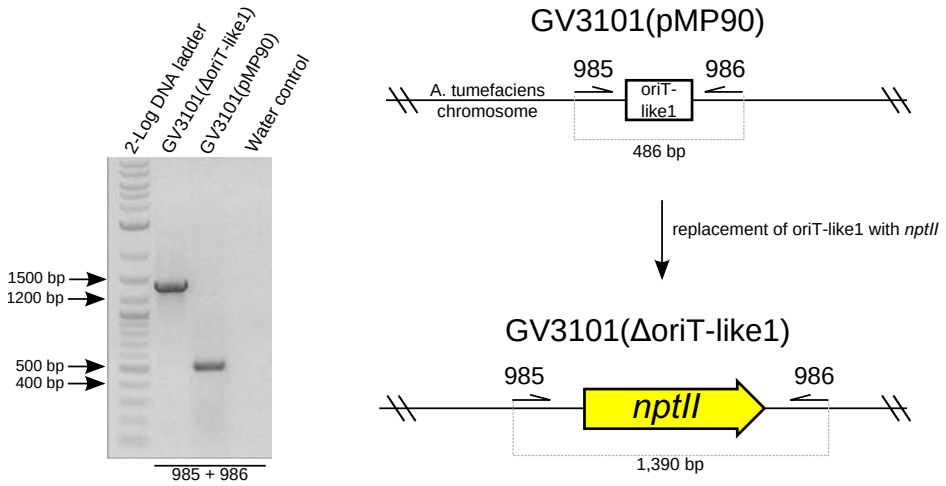


Figure 4.35: Confirmation of oriT-like1 deletion. Arrows indicate primer binding sites. The replacement of oriT-like1 by *nptII* was confirmed by PCR with primers 985 and 986 which anneal upstream and downstream of oriT-like1. If the replacement took place, a 1,300 bp band instead of a 486 bp band is amplified.

Regions upstream and downstream from the oriT-like1 sequence were amplified from the linear chromosome. Both PCR fragments were ligated upstream and downstream of the *nptII* selectable marker. The vector was used to transform competent $HS1_{LC}$::*pat*/*hfp*(pMP90) cells. Cells were selected on spectinomycin and kanamycin.

Resistant cells were screened by colony PCR for replacement of the oriT sequence by *nptII*. Because of the size difference of oriT-like1 and *nptII*, the different genotypes could be distinguished (Fig. 4.35).

Successfully transformed bacteria were used to infiltrate *N. benthamiana* leaves. Fluorescence was compared to strain HS1$_{LC}$::*pat*/*hfp*(pMP90) without the deletion. Figure 4.36 C and D shows that the strain where the oriT-like1 sequence was replaced by *nptII* showed less fluorescent cells than the one where the oriT-like1 sequence was still present (≈ 1 out of 1,000 compared to 1 out of 10 fluorescent cells; Fig. 4.36 C and D). This indicates an involvement of the oriT-like1 sequence in the transfer of the HS1$_{LC}$ locus. However, after deletion of the chromosomal oriT-like1 sequence, the transfer did not stop completely as still some fluorescent cells were detectable.

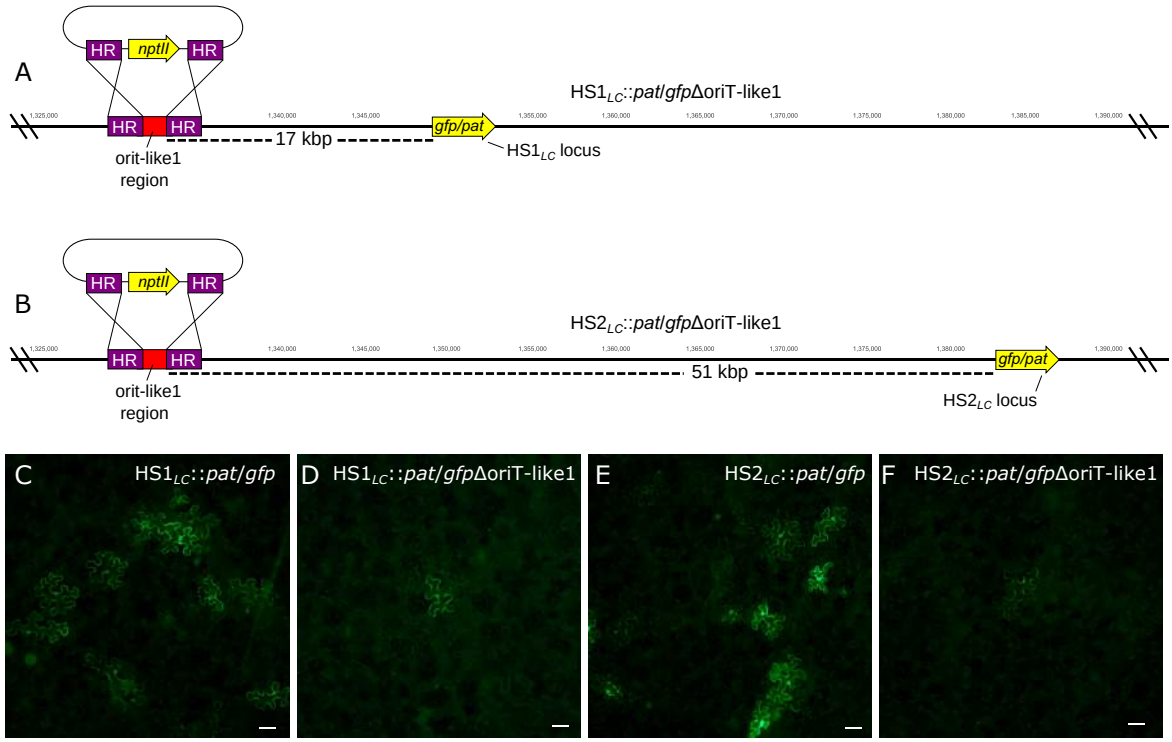


Figure 4.36: Deletion of the oriT-like1 locus in reporter gene tagged *A. tumefaciens* strains. A: oriT-like1 was replaced by *nptII* in HS1$_{LC}$::*pat*/*gfp*(pMP90). **B**: The oriT-like1 sequence was replaced by *nptII* in HS2$_{LC}$::*pat*/*gfp*(pMP90). **C, D**: Comparison of the transfer of the *gfp* tagged HS1$_{LC}$ locus with and without the oriT-like1 sequence. Deletion of oriT-like1 resulted in less GFP expressing plant cells. **E, F**: Comparison of the transfer of the *gfp* tagged HS2$_{LC}$ locus with and without the oriT-like1 sequence. Deletion of oriT-like1 resulted in less GFP expressing plant cells.

In an additional experiment the question if the oriT-like1 sequence was also responsible for transfer of the HS2$_{LC}$ locus was addressed. The oriT-like1 sequence is

located about 51 kbp upstream of the HS2$_{LC}$ locus. The oriT-like1 sequence was also replaced by *nptII* in HS2$_{LC}$::*gfp*/*pat*(pMP90) analogous to its replacement in HS1$_{LC}$::*gfp*/*pat*(pMP90)(Fig. 4.36 B). After successful transformation and screening for a colony in which the oriT-like1 was replaced by *nptII*, the bacteria were used for tobacco leaf infiltration. Strains containing the *gfp* insertion in the HS2$_{LC}$ locus with or without the oriT-like1 sequence where compared to each other.

Similar to the results obtained by the deletion in the HS1$_{LC}$::*gfp*/*pat*(pMP90) strain, deletion of oriT-like1 in HS2$_{LC}$::*gfp*/*pat*(pMP90) led to a decreased number of GFP expressing cells (Fig. 4.36 E, F). However, few GFP expressing cells were still observable. This shows that the oriT-like1 sequence might also have been involved in the transfer of the HS2$_{LC}$ locus. The results confirm that the oriT-like1 sequence on the linear chromosome is involved in transfer of *A*chrDNA to the plant cell.

This information was used to generate a new *A. tumefaciens* strain which has a restricted *A*chrDNA transfer potential. For this reason, the oriT-like1 sequence was deleted in *A. tumefaciens* GV3101(pMP90). The resulting strain was termed GV3101(pMP90)ΔoriT-like1 and can be used for future plant transformation with a decreased chance of co-transferring undesired bacterial chromosomal DNA.

5 Discussion

5.1 IPTmKanR assay

In order to test the hypothesis that T-DNA occasionally integrates into *Agrobacterium*'s own chromosome, vector pIPTmKanR was generated and the IPTmKanR assay was conducted (see 4.1.1). The assay is a variation of the classical promoter/gene-trapping which is used for reverse genetic analysis of gene functions in eukaryotes (Springer, 2000; Stanford et al., 2001).

5.1.1 Kanamycin resistant bacteria did not gain resistance because of a T-DNA insertion in the own chromosome

In this study, the hypothetical integration of a T-strand into the bacterial chromosome would lead to expression of a kanamycin resistance gene. The assay resulted in the formation of kanamycin resistant colonies. However, resistance was predominantly mediated by transposition of the insertion sequence IS*426* into pIPTmKanR (discussed in 5.2) and by plasmid rearrangements, which led to *nptII* transcription by the promoter of the spectinomycin resistance gene in the vector backbone. These rare rearrangement cases are selected for by applying a high selective pressure (alternatives are discussed below).

Inverse PCR analysis of the colonies in which kanamycin resistance was not mediated by IS*426* insertion or plasmid rearrangements, did not show a T-strand integration into the chromosome. In these colonies, resistance potentially developed without the involvement of pIPTmKanR. It was reported that activation of bacterial ABC transporters, which are able to actively pump out toxic substances, can lead to antibiotic resistant bacteria (Levy, 1992; Dawson and Locher, 2006). The genome of *A. tumefaciens* encodes an unusual high number of ABC transporters and only few ones have been characterized so far (Wood et al., 2001). Activation of such transporters, for example mediated by IS*426*, could potentially lead to kanamycin resistant bacteria.

IS*426* is a good candidate for gene activation, since it is an active IS element and was shown to integrate into different positions in the bacterial genome (Vanderleyden et al., 1986; Luo and Farrand, 1999). Additionally, in this study it was shown that it has the ability to act as a promoter for flanking genes (discussed in 5.2.1).

5.1.2 No T-DNA insertion in the bacterial chromosome could be detected

Using the IPTmKanR assay, the integration of T-strands into the bacterial chromosomes could not be shown. This might be due to shortcomings of the assay. For example a particular strong chromosomal promoter would be needed to generate enough NptII proteins to mediate kanamycin resistance. The T-DNA integration next to a strong bacterial promoter might be a rare event and analysis of more resistant bacteria would be necessary to detect one. However, in a project with the goal to identify constitutively active *A. tumefaciens* promoter sequences, Pratibha Kamble of the PME-Group cloned random *A. tumefaciens* chromosomal DNA fragments in front of a promoterless *nptII* gene. She found several sequences which were able to drive the expression of NptII, thus allowing the bacterium to grow in the presence of kanamycin (P. Kamble and B. Ülker, unpublished). Thus, in theory the bacterial chromosomes harbour many promoters which could drive the expression of the T-DNA encoded NptII.

The assay could be improved by shortening the distance between the RB and the *nptII* start codon. After release of the T-strand, there is a distance of 228 nt between the *nptII* start codon and the 5' end of the T-strand. If a T-DNA inserts next to a weak chromosomal promoter, this distance might lead to preliminary dissociation of the RNA polymerase and a weak NptII expression, not allowing the formation of a colony.

However, the failure to detect T-DNA insertion events in the *A. tumefaciens* bacterial chromosome might simply show that such insertions do not happen. Indeed, no *A. tumefaciens* encoded proteins are known to mediate such an integration (Gelvin, 2008). Furthermore, if the integration into the chromosome is a natural event and not an artificial laboratory artefact, in some *A. tumefaciens* strains remnants of inserted T-DNAs should have been detectable, like it is the case for plasmids which at some point in evolution integrated into bacterial chromosomes (Hagblom et al., 1986; Dempsey and Dubnau, 1989).

5.1.3 Conclusion and outlook: IPTmKanR assay

The hypothesis that T-DNA first integrates into the chromosome and is in a second step re-cleaved and together with *A*chrDNA transported to the plant cell could not be confirmed by the IPTmKanR assay and further experiments were conducted on the assumption that such integrations do not happen. However, because of the many false positive colonies, which gained their resistance either by a rearrangement of the plasmid, by an insertion of IS*426* or independently of pIPTmKanR, it cannot be excluded that true T-DNA insertion events were missed.

In order to eliminate the cases where IS*426* integrated in front of the reporter gene, the assay should be performed in the bacterial strain devoid of IS*426*, which was generated in this study. Using promoterless reporter genes like *gfp*, *luciferase* or *lacZ*, instead of an antibiotic resistance gene, could theoretically eliminate false positive colonies since no strong selective pressure is applied. Colonies showing expression of the reporter gene would constitute potential T-DNA integration cases. Both Alexandra Stirnberg and Hamed Al Ghaithi of the PME-Group tried to use promoterless *gfp* and *lacZ* genes to find the stimulus which is triggering transposition of IS*426* (Stirnberg, 2011; Al Ghaithi, 2012). These attempts were not successful because *Agrobacterium* carries an uncharacterised *lacZ* gene on its chromosome and, even if a strong bacterial promoter was used, no GFP positive *A. tumefaciens* cells could be detected. These obstacles have to be overcome first before using these genes for finding T-DNA insertions into the bacterial chromosomes. Furthermore, without antibiotic selection, extensive screening of many colonies would be required.

5.2 IS*426* insertions

5.2.1 IS*426* can control neighbouring gene expression

In the majority of kanamycin resistant colonies the transcription of *nptII* was mediated by an IS*426* insertion upstream of its start codon. Thereby, the transposable element always integrated in the same orientation, with its IRR facing the *nptII* start codon. This suggests that the inherent promoter activity of IS*426*, responsible for the transcription of *nptII*, is unidirectional towards the IRR. In this study the ability of IS*426* to mediate transcription of flanking genes was demonstrated for the first time. Control of neighbouring gene expression is a common feature of insertion sequences and has been demonstrated for many of them (e.g. IS*1*, IS*2*, IS*5* (Mahillon and Chan-

dler, 1998) or IS*406* (Scordilis et al., 1987). Some (e.g. IS*1*, IS*2* and IS*5*) harbour an outwardly directed -35 promoter region in their terminal repeats. If these elements insert in the right distance to a chromosomal -10 promoter region, a strong promoter can be formed. Alternatively, it is also possible that an endogenous transposase promoter influences genes beyond the IS element's terminal repeats (Charlier et al., 1982; Ciampi et al., 1982; López de Felipe et al., 1996). The latter seems to be the case for IS*426*. The location of the element, responsible for controlling neighbouring genes, was further investigated in the Master's thesis of Alexandra Stirnberg. Her results indicate a location within a 241 bp sequence, covering the IRL and the first part of orfA in the 5' region of IS*426* (Stirnberg, 2011).

5.2.2 IS*426* frequently integrates into Ti-plasmids and binary vectors

The transposable element IS*426* was reported to frequently integrate into the T-DNA region of Ti-plasmids, leading to non-pathogenic *A. tumefaciens* variants (Vanderleyden et al., 1986; Fortin et al., 1993; Llop et al., 2009). It is not known if the Ti-plasmid is a more attractive place for insertions of IS*426* or if insertions into genes involved in pathogenicity are just more likely to be discovered by scientists, since they often cause phenotypes. Insertions into chromosomal regions (Luo and Farrand, 1999) and transgenes on binary vectors (Rawat et al., 2009) have also been observed. Interestingly, IS*426* also integrated in the helper plasmid of the common binary *Agrobacterium* strain GV3101(pMP90), thereby probably inactivating the *virK* gene (Stirnberg, 2011, Masterthesis). However, it is not clear if this transposition led to a phenotype in this strain as the function of VirK in *A. tumefaciens* virulence is not well characterized (Hattori et al., 2001).

In this work, both chromosomal copies of IS*426* inserted into pIPTmKanR in the kanamycin resistant colonies (distinguishable by an additional base triplet in copyI). However, it is not clear if both contain an active transposase (autonomous transposition) or if the transposase of one copy is inactive (non-autonomous transposition) and is reliant on the *trans* activity of the other one (Hartl et al., 1992).

In order to prevent the unintended integration of IS*426* into plasmids, it would be important to understand how its transposition is triggered. The fact that the *A. tumefaciens* C58 wild type contains only two copies of the insertion sequence, even though it replicates via a copy and paste mechanism, points towards a relative

stability of the element. The antibiotic stress during the IPTmKanR assay might trigger an increased transposition of IS*426*. However, information about stress induced activation of insertion sequences is scarce. Two publications linked elevated metal-ion concentrations and UV-light to an increased transposition of two IS elements in *E. coli* (Eichenbaum and Livneh, 1998; Brocklehurst and Morby, 2000). The *A. tumefaciens* C58 genome contains a tetracycline resistance gene (*tetA*) on its linear chromosome. It is normally kept inactive by its repressor (*tetR*). However, under tetracycline selection IS*426* was to found to insert into *tetR*, thereby disrupting it and thus allowing expression of *tetA* and survival of the bacterium (Luo and Farrand, 1999). This represents another case where IS*426* transposition is linked to bacterial resistance in the presence of antibiotics.

5.2.3 pIPTmKanR constitutes an insertion sequence trap

Using pIPTmKanR, T-DNA insertions into the chromosomes could not be detected. However, the plasmid was able to efficiently trap IS*426* via the insertion sequence's ability to act as a promoter for neighbouring genes.

Generally, so called transposon traps are utilized to investigate if a bacterial genome contains active transposable elements (Solenberg and Burgett, 1989; Cirillo et al., 1991; Guilhot et al., 1992). pIPTmkanR can also serve this purpose. It can specifically trap transposable elements which are capable of controlling neighbouring gene expression. It contains the broad host range pVS1 origin of replication and is stably kept in a wide range of bacteria (Itoh et al., 1984). Thus, it might prove useful for characterization studies of transposable elements in other prokaryotes.

5.2.4 IS*426* could not transfer to plant cells independently

Ülker et al. (2008a) analysed three independent GABI-Kat *A. thaliana* T-DNA mutant lines carrying an IS*426* insertion in their chromosome (343H01, 146B12 and 135B06). This gave rise to the question, if IS*426* is transferred to plant cells independently of a co-transferred T-DNA. For this reason, in three independent approaches IS*426* was tested for harbouring a sequence which can mediate the transfer of a *gfp* gene to plant cells (4.2.2). However, non of the three approaches led to a detectable DNA transfer. Therefore, it can be assumed that the T-DNA transfer is a necessary prerequisite for the transfer of IS*426* to plant cells.

The GABI-Kat lines which carried the IS*426* insertion showed complex insertion

patterns. In each case, two copies of the T-DNA were flanking a truncated IS*426* sequence. One of the T-DNAs was always next to the right inverted repeat of the IS element (Ülker et al., 2008a). The observation that two T-DNAs were flanking a truncated IS*426* might have resulted from an integration of the insertion sequence into the T-DNA region in one bacterium. Since two T-DNAs were found in the plant genome, a second T-strand integration occurred. The second T-strand probably originated from a plasmid where IS*426* did not integrate into the T-DNA. If two T-DNAs are integrating into the genome of one plant cell, they frequently integrate into the same locus and often cause complex insertion patterns, including truncated T-DNA sequences (De Neve et al., 1997). Thus, the complex insertion patterns observed by Ülker et al. (2008a) are probably caused by two T-DNA insertions into the same locus, one with and one without IS*426*. Such an event was already shown for the octopine *A. tumefaciens* strain LBA4404 by Kim and An (2012). The authors observed an insertion of the transposon Tn*5393* into the T-DNA region. Since not all bacteria carried this insertion prior to transformation, plants were transformed with two different T-DNAs, either carrying the transposon or not. This led frequently to similar complex insertion patterns containing deletions as observed by Ülker et al. (2008a) (Kim and An, 2012).

5.2.5 Conclusion and outlook: IS*426* transposition

In this work *A. tumefaciens* insertion sequence IS*426* was trapped by using vector pIPTmKanR. Furthermore, the ability of IS*426* to activate flanking genes was demonstrated. It was also shown that IS*426* is not able to transfer to plants on its own. The transposition into a T-DNA followed by plant transformation is probably the reason for detection of IS*426* in plant cells.

Since IS*426* frequently interfered with plant transformation, information about the stimuli which are triggering its transposition would be valuable. The pIPTmKanR vector could potentially be used to determine these triggers. The promoterless *nptII* could be exchanged by a promoterless *lacZ* gene. After transformation of *A. tumefaciens*, bacteria should be exposed to different stimuli, potentially causing transposition (e.g. salt, heat, light etc.) and plated on X-Gal containing LB agar. If a stimulus resulted in an elevated IS*426* transposition, an increased number of blue colonies would appear. However, since *A. tumefaciens* carries an uncharacterised *lacZ* gene in its genome (Al Ghaithi, 2012), this gene has to be deleted first.

5.3 Deletion of IS*426* led to a strain with a decreased risk of *A*chrDNA transfer

IS*426* was shown to frequently integrate into chromosomal DNA, plasmids, *vir*-genes or transgenes on binary vectors, to activate neighbouring genes, to cause gene disruptions and to be unintentionally transferred to plant cells (Vanderleyden et al., 1986; Fortin et al., 1993; Ülker et al., 2008a; Rawat et al., 2009; Stirnberg, 2011; this work). Because of these unpredictable and unwanted consequences, it would be desirable to use an *A. tumefaciens* strain for plant transformation, in which IS*426* is deleted and thus cannot cause any of these issues.

For this reason, both copies of IS*426* were deleted from the linear chromosome of *A. tumefaciens* strain A136. A136 is a derivative of *A. tumefaciens* strain C58 which was cured of its Ti plasmid. Thus, it offers the flexibility to generate a novel binary vector system by introducing a new helper plasmid to the strain (work of Max Schelski, PME-Group).

To remove IS*426* from the genome, two IS*426* deletion vectors were generated. IS*426* copyI and II were subsequently replaced by a spectinomycin and kanamycin resistance gene (see 4.1.1). The partial copy of the At-plasmid was not deleted, since the sequence is already truncated and contains many mutations. Furthermore, it never inserted in pIPTmKanR, making it unlikely to be an active transposable element. The replacements resulted in an *A. tumefaciens* strain devoid of IS*426* (A136ΔIS*426*).

5.3.1 Conclusion and outlook: strain with decreased risk of *A*chrDNA transfer

In this work, the first step towards a more bio-safe and reliable *A. tumefaciens* plant transformation strain was done by deletion of both IS*426* copies. The next step would be to excise the inserted antibiotic resistance genes, in order to allow more flexibility in the choice of binary vectors and helper plasmids which are compatible with this strain. Afterwards, the oriT-like1/RB-like2 sequences should be deleted as it was done for oriT-like1 in this work (see 4.5). This would additionally decrease the risk of transferring *A*chrDNA to plants. To complete the vector, a helper plasmid should be introduced to allow convenient plant transformation.

In order to further decrease the risk of *A*chrDNA transfer, a gene which encodes for a plant-toxic protein could be integrated into different regions of the bacterial

genome. Every time *A*chrDNA is co-transferred with the T-DNA, no plant could be regenerated and only "clean" T-DNA insertions would be obtained. For counter-selection, the *barnase* gene from *Bacillus amyloliquefaciens* containing an intron could be used, as it was done to counter-select for transferred vector backbone upon left border skipping (Hanson et al., 1999). However, the transfer of genes encoding for toxic proteins might cause other biosafety problems and should be used with caution.

5.4 Insertion of reporter genes into the bacterial chromosomes

So far, the transfer of *A*chrDNA always occurred unintentionally during generation of T-DNA insertion mutants and was only discovered later by analysis of FSTs (Ülker et al., 2008a). In order to investigate the underlying mechanism of this transfer, it had to be reproduced, which was successfully achieved in this work. Introduction of reporter genes into different regions of the bacterial chromosomes by homologous recombination, followed by plant transformation, led to transfer of the reporter genes from bacterial chromosomes to plant cells, thus confirming the results by Ülker et al. (2008a).

Depending on the chromosomal regions, where the reporter genes where inserted, different amounts of GFP expressing plant cells could be observed.

5.4.1 The $NE1_{LC}$ and $NE2_{LC}$ locus did not transfer to plant cells

The $NE1_{LC}$ and $NE2_{LC}$ loci were chosen for insertion of the *pat*/*gfp*-vector, because they are located in a chromosomal region, which was never found in plant cells by Ülker et al. (2008a). Therefore, these integrations served as negative controls for the transfer of *A*chrDNA to plant cells.

Integrations into the desired loci were tested by DNA blot analysis. A single integration of the vector into the $NE1_{LC}$ region of GV3101(pMP90) was not achieved. A second larger band, indicating a second integration of the vector, was detectable on the membrane (see 4.21). It was not assessed if the vector inserted two times into the same locus or if the second insertion happened somewhere else in the genome. However, a band of the correct size was also detectable, indicating a vector insertion in the desired locus. Thus, the strain $NE1_{LC}$::*patgfp*(pMP90) was used for *N. benthamiana* infiltration. Because of the difficulties obtaining a single insertion into the

NE1$_{LC}$ locus in *A. tumefaciens* GV3101(pMP90) strain, insertions in this locus were not attempted in strains A136 and ATΔ*virD2*. DNA blot analysis of vector insertions into the NE2$_{LC}$ locus did only result in one detectable band of the correct size for all three strains. It was shown in two publications that disruption of this locus does not interfere with virulence of the bacterium (Lee et al., 2001; Oltmanns et al., 2010), making this locus a good choice for the integration of reporter genes.

Infiltration of NE1$_{LC}$::*pat*/*gfp*(pMP90) and NE2$_{LC}$::*pat*/*gfp*(pMP90) never led to any GFP expressing plant cells (Fig. 4.22 E and F). The absence of GFP expression in plant cells after infiltration showed that no transfer of the NE1$_{LC}$ and NE2$_{LC}$ loci took place. The result is in agreement with the observation by Ülker et al. (2008a) that this region does not contain hotspots of *A*chrDNA transfer. The most probable reason for the absence of *A*chrDNA transfer in this region is that the first 500 kbp of the linear chromosome do not harbour any sequence which can be used for initiating a transfer of DNA.

Furthermore, the absence of GFP signal in *N. benthamiana* leaves shows that the vector backbone used for integration of the reporter genes does not contain any sequence (e.g. border like) which is able to mediate DNA transfer to plants. Similarly, it shows that the detection of a GFP is not due to its expression in bacteria after leaf infiltration.

Thus, DNA transfer is exclusively dependent on the locus of integration in the chromosome.

5.4.2 The HS1$_{LC}$ locus is frequently transferred to plant cells

The correct insertion of the vector into the HS1$_{LC}$ locus was confirmed by DNA blot analysis for all strains. In strain ATΔ*virD2* a second larger band was visible, indicating a second insertion of the vector (Fig. 4.19 D). Since an insertion into the correct locus was also achieved, strain HS1$_{LC}$::*pat*/*gfp*(ATΔ*virD2*) was still used for *N. benthamiana* infiltration. However, the results obtained by this strain have to be interpreted with care (see below).

HS1$_{LC}$ is one of two tagged loci in the hotspots of frequently transferred *A*chrDNAs on the linear chromosome. The locus was chosen because of its location 243 bp upstream of the RB-like1 sequence, identified by Ülker et al. (2008a) (see Fig. 4.25). The RB-like1 sequence was potentially involved in the transfer of an 18 kbp *A*chrDNA sequence to GABI-Kat line 086C02. Because of the close proximity and its orientation in the chromosome, RB-like1 would also mediate the transfer of a reporter gene

integrated into the $HS1_{LC}$ locus.

The by $HS1_{LC}$::*pat*/*gfp*(pMP90) transformed *N. benthamiana* leaf epidermis cells indeed showed expression of GFP, indicating that a transfer of the $HS1_{LC}$ locus from the chromosome to the plant DNA took place. Compared to $HS2_{LC}$ and $HS1_{CC}$, *gfp* insertion in the $HS1_{LC}$ locus resulted in the highest number of GFP expressing cells (≈ one out of 10 cells; Fig. 4.22 B). The positive control with a "normal" T-DNA resulted in a uniform GFP expression of virtually all cells (Fig. 4.22 A). The fact that this is not the case for $HS1_{LC}$::*pat*/*gfp*(pMP90) shows that the transfer from this locus is less efficient than it is with a "normal" T-DNA. Rommens et al. (2005) showed that the T-DNA's border sequence can be degenerated to a certain degree and still mediate DNA transfer to plant cells. The degree of degeneration thereby influenced the efficiency of T-DNA transfer. It is likely that such a degenerated border sequence is located on the linear chromosome and mediates the transfer of the $HS1_{LC}$ locus, but less efficiently than a T-DNA RB does.

By insertion of *gfp* in the $HS1_{LC}$ locus the transfer of *A*chrDNA to plant cells was successfully visualized. However, the assay only showed that chromosomal DNA was transferred but did not show from which sequence the transfer started and where it was terminated.

The RB-like1 sequence does not mediate transfer of $HS1_{LC}$

The next step was to determine which sequence is mediating this transfer. For this reason, candidate sequences were PCR amplified and ligated to a vector which is stably kept in *A. tumefaciens*. Furthermore, the vector carries a *gfp*-gene and does not harbour any sequences which can mediate DNA transfer to plants. The resulting vectors were used to transform GV3101(pMP90) cells and *N. benthamiana* leaf cells were infiltrated. Ligation of a candidate sequence which is able to act as a "border", is expected to lead to transfer of the complete vector to plant cells indicated by GFP expression, as it is the case for vectors which contain only one border repeat (Horsch and Klee, 1986).

Because of the close proximity of the RB-like1 sequence to the integrated vector in the bacterial chromosome, its homology to RB as well as the potential involvement in the transfer of *A*chrDNA to GABI-Kat line 086C02, it was tested for its ability to mediate DNA transfer to plants (see 4.3.2). Surprisingly, no GFP expressing cells could be observed and hence RB-like1 does not seem to be able to act as a starting sequence for DNA transfer on its own.

A 5 kbp fragment mediated DNA transfer to a weak extend

In an additional attempt to locate the sequence responsible for the transfer of $HS1_{LC}$, a chromosomal sequence with a size of 5 kbp was tested for its transfer ability (see 4.3.1). It included the $HS1_{LC}$ integration locus and flanking sequences upstream and downstream. This sequence led to GFP expressing plant cells. However, the amount of fluorescent cells was much less compared to the $HS1_{LC}$::*pat*/*gfp*(pMP90) strain with the stably integrated *gfp*. If the sequence, which is mediating the *gfp* transfer in $HS1_{LC}$::*pat*/*gfp*(pMP90) would be located within the 5 kbp, an equal or higher amount of GFP expressing cells would be expected.

The lower amount of GFP expressing cells could be explained by two possibilities. It might be the case that the 5 kbp fragment does not contain the sequence responsible for the transfer of *gfp* in $HS1_{LC}$::*pat*/*gfp*(pMP90), but harbours an additional, weaker starting sequence, mediating the low of *gfp* transfer. Alternatively, the 5 kbp fragment might occasionally lead to a complete integration of the vector into the genome by homologous recombination. The so integrated vector is then transferred to the plant cell the same way, the vector integrated into the $HS1_{LC}$ locus is transferred. Since the transfer of $HS1_{LC}$ is likely mediated by oriT-like1, this possibility could be excluded by transforming an oriT-like1 deletion strain with the 5 kbp containing vector.

Since the high amount of GFP expressing cells cannot be explained by the 5 kbp sequence, further PCR fragments were tested. However, it cannot be excluded that the 5 kbp fragment contains an additional DNA transfer starting sequence.

The oriT-like1 sequence mediates transfer of $HS1_{LC}$ to plant cells

The conjugative mobilization of plasmids is initiated at origin of transfer (oriT) sequences. These sequences share high degree of sequence identity with the Ti plasmid's border sequences (Pansegrau and Lanka, 1991; Waters et al., 1991). Furthermore, it was shown that VirD2 can nick oriT sequences *in vitro* (Pansegrau et al., 1993) and *in vivo* (Dube et al., 2004). Because of the close relation of oriTs and Ti borders, not only border-like sequences but also chromosomal oriT-like sequences were analysed for their ability to mediate the transfer of *A*chrDNA to plants.

The oriT-like1 sequence is located within a gene coding for an ABC-transporter. This transporter was found in five independent GABI-Kat T-DNA insertion lines and thus constitutes a hotspot of transferred *A*chrDNA. Similar to oriT sequences, the oriT-like1 sequence contains a core region and upstream short putative inverted

repeat stretches could be found (see Fig. 4.27). A 61 bp fragment, harbouring the core and repeat region, was able to mediate DNA transfer to plant cells, while a shorter fragment including only the core region failed. Thus, the repeat region of this sequence seems to be necessary for the transfer.

The amount of GFP expressing plant cells, initiated by pBasicS1-GFP-oriT-like1, was comparable to the amount which was observed by the integration of *gfp* in chromosomal $HS1_{LC}$ locus ($HS1_{LC}$::*pat*/*gfp*(pMP90)). By comparing oriT-like1 to known oriTs, it can be deduced that the transfer of sequences located downstream of oriT-like1 is mediated. Thus, oriT-like1 seems to be the sequence responsible for the transfer of the $HS1_{LC}$ locus. Indeed, deletion of oriT-like1 from the linear chromosome resulted in a dramatically reduced transfer rate of the $HS1_{LC}$ locus (discussed in 5.6). Furthermore, most of the hotspots of frequently transferred *A*chrDNA are located downstream of oriT-like1. Therefore, the transfer of most of these hotspots is likely mediated by the oriT-like1 sequence. Since there is no clear point of termination, the strand mobilization which started at oriT-like1 might get terminated randomly, resulting in the transfer of *A*chrDNA fragments of different sizes. However, not all of the *A*chrDNA fragments observed by Ülker et al. (2008a) have the oriT-like1 sequence at the end. This might be explained by deletions which occurred by the simultaneous integration of T-DNA and *A*chrDNA, since integrations of more than one T-DNA often leads to truncations and complex insertion patters.

oriT-like1 instead of RB-like1 could have mediated the transfer of *A*chrDNA to GABI-Kat line 086C02

Possibly, the transfer of the 18 kbp *A*chrDNA fragment from the bacterial chromosome to GABI-Kat line 086C02 was not initiated at the RB-like1 sequence as hypothesised by Ülker et al. (2008a), but terminated there. This would mean that a sequence located approximately 18 kbp upstream of RB-like1 was responsible for this transfer. Intriguingly, the oriT-like1 sequence is located exactly 18.5 kbp upstream of RB-like1. In this scenario, the transfer of the 18 kbp sequence was initiated at the oriT-like1 sequence and terminated at the RB-like1 sequence, meaning that the RB-like1 sequence is actually an LB-like. Prediction, if a sequence has termination (left border) or initiation (right border) character is difficult, since it is not really understood which parts are defining it (Podevin et al., 2006). It might be that the RB-like1 sequence allows termination of T-strand mobilization, but not its initiation.

5.4.3 The $HS2_{LC}$ locus is frequently transferred to plant cells

Reporter genes were integrated into a second locus within hotspot region on the linear chromosome ($HS2_{LC}$). The $HS2_{LC}$ locus is located 33 kbp downstream of $HS1_{LC}$ (see Fig. 4.25). It was chosen because of its central location within the accumulation of frequently transferred *A*chrDNAs. Successful integration was confirmed by DNA blot analysis in all strains (Fig 4.20 C). Similar to the $HS1_{LC}$ locus, *gfp* inserted into the $HS2_{LC}$ locus in strain GV3101(pMP90) was transferred and resulted in GFP expressing plant cells. However, the number of expressing cells was lower ($\approx$ one out of 50 cells; Fig. 4.22 C).

This might indicate that the transfer of the $HS2_{LC}$ locus is initiated at a sequence, which is less effective in DNA mobilization than the one responsible for the transfer of $HS1_{LC}$. Due to its location and orientation on the chromosome, the oriT-like3 sequence could have potentially been involved in the transfer of $HS2_{LC}$ and was tested for its ability to transfer DNA to plant cells. However, no DNA transfer was detectable when transforming plants with pBasicS1-GFP-oriT-like3, indicating that oriT-like3 is not involved in the transfer of $HS2_{LC}$. In contrast, a 5 kbp fragment including the locus of integration resulted in GFP expressing cells. Like the 5 kbp fragment including the $HS1_{LC}$ locus, the amount of fluorescent cells was lower than when using the $HS2_{LC}$::*pat*/*gfp*(pMP90) strain, indicating that the 5 kbp fragment does not harbour the sequence responsible for the transfer.

Alternatively, the transfer of $HS1_{LC}$ and $HS2_{LC}$ might not represent two separate events. Both loci could be located on one large mobilized *A*chrDNA strand with a size of at least 33 kbp. The fact that less of the $HS2_{LC}$ locus is transferred might indicate that the transfer starts upstream of $HS1_{LC}$, covers $HS1_{LC}$ and $HS2_{LC}$ and terminates downstream of $HS2_{LC}$. Sometimes termination might occur earlier, resulting in a lower transfer of $HS2_{LC}$ compared to $HS1_{LC}$. Indeed, deletion of oriT-like1 resulted in a drastically decreased amount of transfer of the $HS1_{LC}$ locus but also of the $HS2_{LC}$ locus (discussed in 5.6), indicating that both loci are transferred as one large *A*chrDNA fragment.

5.4.4 The $HS1_{CC}$ locus is transferred to plant cells

In this study it was also assessed how *A*chrDNAs of the circular chromosome could be transferred to plant cells. For this reason the *gfp* reporter gene was inserted into the $HS1_{CC}$ locus. The $HS1_{CC}$ locus was chosen for marker gene integration because

of its location within an *A*chrDNA sequence found in GABI-Kat T-DNA insertion line 052H10 by Ülker et al. (2008a). Since the authors did not asses the hotspots of frequent *A*chrDNA transfer of the circular chromosome, it is unknown if the HS1$_{CC}$ locus is a hotspot of *A*chrDNA transfer or if its transfer to line 052H10 was a unique event.

DNA blot analysis confirmed integration of the vector into the correct locus in GV3101(pMP90), A136 and ATΔ*virD2*. In ATΔ*virD2* a second band was detectable, indicating an additional insertion. Since a vector insertion in the correct locus was confirmed, the strain was still used for subsequent experiments. *N. benthamiana* leaf infiltration with HS1$_{CC}$::*pat*/*gfp*(pMP90) resulted in GFP expressing leaf cells, indicating that a transfer of *A*chrDNA took place. The number of GFP positive cells was the lowest compared to HS1$_{LC}$ and HS2$_{LC}$::*pat*/*gfp*(pMP90). Likely, the sequence mediating the transfer of this locus is less efficient or located further afar.

To identify the starting point of the transfer, four PCR fragments were amplified and tested for harbouring a sequence, able to mediate DNA transfer to plant cells. The LB-like1 fragment, the 052H10-flanking-locus fragment and the 5 kbp fragment covering the HS1$_{CC}$ integration locus did not mediate DNA transfer to plant cells.

The RB-like2 sequence mediates *A*chrDNA transfer from the circular chromosome

In contrast, the RB-like2 sequence located almost 50 kbp downstream of the HS1$_{CC}$ integration locus resulted in GFP expressing plant cells (see 4.3.3). Compared to oriT-like1, the amount of GFP expressing cells was much lower, indicating that the immanent sequence is not as efficient in mediating DNA transfer. Because of its orientation, the RB-like2 sequence would mediate the transfer of DNA which is located upstream of it. Thus, it would be possible that this sequence is responsible for the transfer of HS1$_{CC}$ and the *A*chrDNA transferred to GABI-Kat line 052H10.

However, it cannot be excluded that another sequence is responsible for the transfer of HS1$_{CC}$. In order to test this, RB-like 2 should be deleted in HS1$_{CC}$::*pat*/*gfp*, like it was done in HS1$_{LC}$/HS2$_{LC}$::*pat*/*gfp*ΔoriT-like1. If, after deletion of RB-like2, no GFP expressing cells are detectable, the involvement of RB-like2 as the only element in the transfer of HS1$_{CC}$ would be confirmed.

5.4.5 Conclusion outlook: Identification of sequences mediating DNA transfer

By aligning RB, LB and oriT sequences to the chromosomes of *A. tumefaciens*, followed by cloning an *N. benthamiana* infiltration, two sequences were identified which are able to mediate DNA transfer to plant cells. The methodology to first screen for potential candidate sequences *in silico*, to clone them in a vector and then to test their transfer ability proved to be somewhat ineffective. However, if more of these non-border sequences, mediating DNA transfer to plant cells, are identified, likely a more elaborate consensus sequence could be deduced. This in turn should facilitate the prediction if a chromosomal sequence can mediate DNA transfer.

Alternatively, complete genomic DNA of *A. tumefaciens* could be cut into short fragments, followed by ligation into pBasicS1-GFP. Thereby the average length of the fragments could be influenced by the choice of restriction enzyme. Thus, a library of random *Agrobacterium* DNA fragments in pBasicS1-GFP would be generated. This library could be used for *A. tumefaciens* transformation and *N. benthamiana* leaf infiltration. If infiltration results in GFP expressing plant cells, the responsible sequence could be determined by sequencing of the respective vector. However, this approach would probably be even more cumbersome and repeated detection of the same sequence would be likely.

Quantification of the *A*chrDNA transfer process

Since it was shown that the oriT-like1 and RB-like2 sequences are able to mediate DNA transfer to plant cells, it is important to quantify the efficiency of this process. Tobacco leaf infiltration and screening for GFP expressing cells makes it difficult to accomplish this. The transfer process is very sensitive and many factors can influence the amount of GFP expressing cells (e.g. age of the plant, size of the leaf, area of infiltration etc.). Furthermore, counting GFP expressing cells is error prone and might lead to false results.

Transforming *A. thaliana* plants by floral dip should allow quantification of the transfer process. By comparing the number of BASTA® resistant plants obtained by transformation with pBasicS1-PAT-oriT-like1 or pBasicS1-PAT-RB-like2 to the number obtained by a normal T-DNA transformation, the efficiency of the transfer process could be assessed. For this experiment, the *gfp* gene in pBasicS1-GFP-oriT-like1 and pBasicS1-GFP-RB-like2 should be replaced by a *pat* gene.

In this work, it was attempted to transform *A. thaliana* plants with an *A. tumefaciens* strain which has the *pat* and *gfp* reporter genes inserted into the $HS1_{LC}$ locus on the linear chromosome ($HS1_{LC}$::*gfp*/*pat*(pMP90)). When this strain was used for *N. benthamiana* leaf infiltration, GFP expressing cells could be observed ($\approx$ 1 out of 10 cells showed fluorescence; Fig. 4.22). When the same strain was used for transformation of *A. thaliana*, no BASTA® resistant plants could be obtained. However, these experiments were not done extensively and environmental factors such as fungi and larvae growing in the soil might have caused early death of actually transformed plants.

Compared to the transformation of *N. benthamiana* leaf cells, transformation of *A. thaliana* is less efficient. Therefore it is possible that more plants have to be screened to obtain one, which was transformed by $HS1_{LC}$::*gfp*/*pat*(pMP90).

Alternatively, instead of leaf infiltration it would be possible to infect *N. benthamiana* explants, followed by regeneration of transformed calli (Rommens et al., 2005). Counting the regenerated calli and comparing them to a positive control would allow a better quantification of the *A*chrDNA transfer, initiated from these sequences.

5.5 Involvement of Vir proteins and VirD2

By integration of *gfp* into different loci in *A. tumefaciens* strain GV3101(pMP90), the transfer of *A*chrDNA to plant cells could be visualized. In order to determine the proteins orchestrating the transfer process, reporter genes were inserted into the respective loci in *A. tumefaciens* strains A136 and the mutated strain ATΔ*virD2*.

Since the Vir-proteins are important for all stages of the T-DNA transfer to plant cells, it is very likely that they are also involved in the transfer of the *gfp* labelled chromosomal sequences. However, since the mobilization and transfer of T-DNA is closely related to the mobilization and transfer of plasmid DNA during conjugation (Stachel et al., 1986; Pansegrau and Lanka, 1991; Lessl and Lanka, 1994), it is also possible that components of a conjugation system are participating in *A*chrDNA transfer. It has been shown that components of the virulence system can substitute for components of the conjugation system and vice versa. For example the broad host range plasmid RSF1010 can be conjugated between bacteria using the *virD4*/*virB* encoded T4SS (Beijersbergen et al., 1992). The same plasmid can be transferred to plant cells in the absence of VirD2, if the plasmid's own mobilization proteins are present (Buchanan-Wollaston et al., 1987; Bravo-Angel et al., 1999).

5.5.1 The Vir proteins are involved in the *A*chrDNA transfer process

In order to test the involvement of the Vir proteins, reporter genes were integrated into the $HS1_{LC}$, $HS2_{LC}$, $HS1_{CC}$ and $NE2_{LC}$ locus of *A. tumefaciens* strain A136. A136 is an avirulent strain based on *A. tumefaciens* strain C58. It was cured of its Ti-plasmid and thus lacks all the genes involved in virulence and pTi conjugation (Watson et al., 1975). *N. benthamiana* infiltration with the four strains did not lead to any GFP expressing plant cells (Fig. 4.23), indicating an involvement of proteins encoded on the Ti-plasmid in the transfer of *A*chrDNA. However, from this experiment it cannot be determined if the complete transfer is orchestrated by the Vir-proteins, or by the components of the Trb system which is involved in Ti-plasmid conjugation.

It has been shown that the broad host range plasmid RSF1010 can be transferred from *A. tumefaciens* to plant cells, starting from its oriT sequence (Buchanan-Wollaston et al., 1987). Instead of VirD2, mobilization of the transferred strand is mediated (less efficiently) by the mobilization proteins MobA, MobB and MobC, encoded on RSF1010, They are normally involved in the conjugative transfer of RFS1010 between bacteria. After mobilization by the Mob-proteins, the strand is transferred to the plant cell and integrated into the nucleus with the help of the Vir proteins (Bravo-Angel et al., 1999). Because strain A136 also lacks the proteins involved in pTi conjugation, it would be possible that these proteins are involved in the transfer of *A*chrDNA.

5.5.2 VirD2 is necessary for the *A*chrDNA transfer process

By testing the transfer of *A*chrDNA in an *A. tumefaciens* mutant strain which lacks the T-DNA "pilot" protein VirD2, the proteins involved in the *A*chrDNA transfer process should be further narrowed down. For this reason, the reporter genes were integrated into the $HS1_{LC}$, $HS2_{LC}$, $HS1_{CC}$ and $NE2_{LC}$ locus of *A. tumefaciens* mutant strain ATΔ*virD2*. The strain is based on C58 and carries a helper plasmid on which *virD2* was deleted, while the remaining virulence and conjugation genes are still present (Bravo-Angel et al., 1998). During his master thesis in the PME-Group, Lamprinos Frantzeskakis re-introduced *virD2* into ATΔ*virD2* and was able to restore virulence in this strain. Hence, except for the missing *virD2*, the strain is functional and can be used to study the involvement of VirD2 in the *A*chrDNA transfer process.

When using the ATΔ*virD2* strain for plant transformation, no expressing plant

cells could be detected. This indicates that without a functional VirD2 no transfer of *A*chrDNA takes place. Thus, it seems likely that mobilization of *A*chrDNA is mediated by the VirD1/VirD2 complex. In this scenario, the complex occasionally binds to chromosomal sequences which resemble Ti-border/oriT sequences. It was shown that the complex is able to tolerate certain sequence variations in the borders and still mediates strand mobilization. Depending on the amount of degeneration, the transfer was impaired (Rommens et al., 2005). Since the chromosomal sequences are not resembling perfect border sequences, the transfer of *A*chrDNA is less effective than normal T-DNA transfer.

DNA blot analysis of the correct reporter gene integration in the $HS1_{LC}$ and $HS1_{CC}$ locus in strain ATΔ*virD2* resulted not only in a band of the expected size but also in an additional band. This indicates that the integration-vector integrated a second time into the bacterial genome. No *A*chrDNA transfer could be observed using these two strains for plant transformation. Because of the second band, this result has to be regarded with some caution.

However, DNA blot analysis of the vector integration into the $HS2_{LC}$ locus in ATΔ*virD2* resulted in only one band of the correct size. Using this strain for transformation, also no *A*chrDNA transfer was detectable, indicating that VirD2 is generally necessary for the transfer of all the reporter gene tagged loci. This was supported by the observation that no DNA was transferred to plant cells from the vectors pBasicS1-GFP-oriT-like1 and pBasicS1-GFP-RB-like2 in ATΔ*virD2* cells (Fig. 4.33).

5.5.3 Conclusion and outlook: involvement of VirD2 in *A*chrDNA transfer

The results indicate that the transfer of *A*chrDNA to plant cells is likely mediated by VirD2. VirD2 seems to be able to bind to chromosomal sequences and from there to initiate the DNA transfer to plant cells. It remains to be shown, if *A*chrDNA is transferred to the plant cell via the VirD4/VirB type IV secretion system or if it uses one of the channels involved in conjugation of pTi or pAt. To further characterize the proteins involved in this transfer, *A. tumefaciens* knock out mutants for different *vir*-genes (e.g. *virB*, *virD4*) should be used for transformation by pBasicS1-GFP-oriTlike1 and pBasicS1-GFP-RB-like2. This way, the key enzymes involved in *A*chrDNA transfer to plants could be further characterised.

The transfer of chromosomal DNA between bacteria, starting from oriT-like se-

quence, was recently shown in the human pathogen *Helicobacter pylori*. Similar to the observations made in this study, the authors described the binding of a VirD2-like protein to a chromosomal oriT-like sequence. Upon binding, the protein introduced a nick in the oriT-like sequence and stayed covalently bound to the 5'-cleaved end (Grove et al., 2013). These observations substantiate the findingsq made in this work, that VirD2 can bind to chromosomal oriT-like sequences and from there mediate DNA transfer.

5.6 The oriT-like1 sequence might initiate transfer of the entire hotspot region

The in 4.5 described experiments were designed to analyse, if the oriT-like1 sequence is responsible for the transfer of $HS1_{LC}$ and $HS2_{LC}$. The oriT-like1 sequence was deleted in the strains $HS1_{LC}$::*pat*/*gfp*(pMP90) and $HS2_{LC}$::*pat*/*gfp*(pMP90). Strains with and without the oriT-like1 sequence were compared to each other, regarding their ability to transfer the integrated reporter genes to plant cells.

Both strains led to a drastically reduced amount of GFP expressing plant cells. Thus, it is likely that the transfer of the $HS1_{LC}$, as well as of the $HS2_{LC}$ locus, is initiated at the oriT-like1 sequence. oriT-like1 is located 17 kbp upstream of $HS1_{LC}$ and 50 kbp upstream of $HS2_{LC}$. The fact that the oriT-like1 sequence might be responsible for the transfer of a locus which is located 50 kbp downstream, was a surprising result. Because most of the hotspots are located between the $HS1_{LC}$ and $HS2_{LC}$ locus, it is possible that oriT-like1 mediates the transfer of most of these hotspots. In this scenario, the VirD1/VirD2 complex would bind to oriT-like1 and initiate the transfer of large downstream located fragments. During transfer and integration, these fragments would occasionally get truncated, resulting in the integrations observed by Ülker et al. (2008a).

The oriT-like1 sequence is located within the coding sequence of a membrane spanning ABC transporter (Atu4209). By replacing oriT-like1 with the *nptII* gene, the ABC transporter was probably disrupted. Thus, the reduced amount of GFP expressing cells could potentially also be due to the loss of the protein function and not because of the oriT-like1 sequence. However, this is unlikely because this particular ABC transporter was never identified to be involved in *A. tumefaciens* mediated plant transformation (Rudder et al., 2014) and generally its function is not known. Furthermore, because of the high number of genes encoding for ABC transporter in

A. tumefaciens (Wood et al., 2001), it is likely that other ones can substitute for this particular transporter. To completely exclude this possibility, the complete ABC transporter gene should be re-introduced in these strains on a plasmid. If the same amount of GFP expressing cells would be visible after *N. benthamiana* leaf infiltration, the involvement of the transporter in the transfer of *A*chrDNA could be entirely ruled out.

Deletion of oriT-like1 did not completely stop transfer of the $HS1_{LC}$ and $HS2_{LC}$ locus, since still some GFP positive cells were detectable (Fig. 4.36). This indicates that at least one more sequence, which is able to mediate DNA transfer, must be present on the linear chromosome. The amount of GFP expressing cells after the deletions is much lower, implying that the VirD1/VirD2 complex probably binds to this sequence with a lower affinity, potentially caused by a lower degree of homology to border/oriT sequences. Because of this lower degree of homology, it is difficult to predict the sequence responsible for this transfer. In this work amplification of 5 kbp sequences including the $HS1_{LC}$ or $HS2_{LC}$ locus still led to a low amount of GFP expressing plant cells (see 4.3.1). Thus, potentially these fragments harbour the sequence responsible for the remaining transfer of $HS1_{LC}$ and $HS2_{LC}$ in the oriT-like1 deletion strains.

5.6.1 Deletion of oriT-like1 led to a strain with a decreased *A*chrDNA transfer potential

With oriT-like1 a sequence with a high potential of transferring *A*chrDNA to plants was identified. Because its deletion dramatically reduced the unintended transfer of *A*chrDNA in the GFP labelled strains, the deletion was also conducted in the original unlabelled GV3101(pMP90) strain, resulting in GV3101(pMP90)ΔoriT-like1. This strain can be used for plant transformation with a reduced risk of simultaneously transferring *A*chrDNA fragments. Since the oriT-like1 sequence was replaced by *nptII*, the binary vector cannot be selected by kanamycin.

5.6.2 Conclusion and outlook: oriT-like1 mediates the transfer of $HS1_{LC}$ and $HS2_{LC}$

By inserting reporter genes into different regions of the bacterial genome, the transfer of *A*chrDNA was successfully visualized and the observations made by Ülker et al. (2008a) were confirmed. The results obtained by the deletion of oriT-like1

in $HS1_{LC}$ and $HS2_{LC}$::*pat*/*gfp*(pMP90) suggest that $HS1_{LC}$ and $HS2_{LC}$ are transferred together on one large *A*chrDNA fragment, which is initiated at the oriT-like1 sequence (Fig. 5.1). In order to obtain further evidence, the $HS1_{LC}$ locus could be labeled with *gfp* and the $HS2_{LC}$ locus with a fluorescent protein which emits light in a different wavelength (e.g. *dsRed*). After the *N. benthamiana* leaf infiltration assay the amount of cells expressing GFP, DsRed, or both should be assessed. If most of the leaf cells would express both reporters, it would be likely that both tagged loci would have been transferred together. If green and red fluorescenting cells would be mostly separated from each other, it would indicate that both loci would have been transported independently to plant cells.

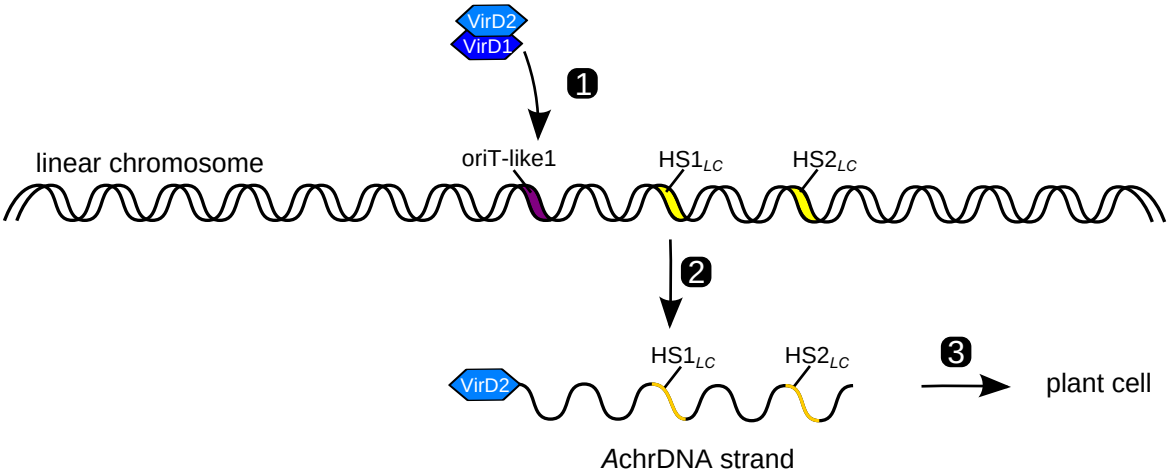


Figure 5.1: Model for the transfer of $HS1_{LC}$ and $HS2_{LC}$ by oriT-like1. 1: The VirD1/VirD2 complex binds to oriT-like1 and initiates mobilization of the single stranded *A*chrDNA strand. **2:** Mobilization of the *A*chrDNA strand is terminated at an unknown location. VirD2 stays attached at the 5' prime end of the *A*chrDNA strand. **3:** The *A*chrDNA strand is transferred to the plant cell.

5.7 Biological reasons for the transfer of *A*chrDNA

The fact that *A*chrDNA is transferred to plant cells imposes the question, if there is a biological reason for it. It might just be an artefact of T-DNA transfer. Alternatively, the transfer of *A*chrDNA might have a more important role for the bacterium.

The transfer of *A*chrDNA could be an artefact of the T-DNA transfer mechanism

An easy explanation would be that the transfer of *A*chrDNA is simply an artefact without further purpose. In this scenario, by chance the bacterial genome contains sequences which can be "misinterpreted" by the VirD1/VirD2 complex as T-DNA borders. The protein complex can tolerate variations from the consensus RB sequence to a certain degree and still mediates the generation of a T-strand (Dube et al., 2004; Rommens et al., 2005). The T-strand mobilization starting from oriT-like1 and RB-like2 is less effective, compared to the Ti plasmid RB. The occasional transfer of an *A*chrDNA to plants would probably not pose a great selective disadvantage for the bacterium, since it happens rarely and the "normal" T-DNA transfer is not influenced. Thus, these border-like sequences did not vanish from the bacterial chromosomes during evolution.

The transfer of *A*chrDNA could be intended for other bacteria

It would also be imaginable that *A*chrDNA fragments are not supposed to be transferred to plant cells but to other bacteria and thus might constitute genomic islands. Genomic islands are defined as large (chromosomal) DNA segments which can be transferred horizontally to other bacteria. They are usually flanked by 16 - 20 bp (almost) perfect direct repeats and often encode integrases or factors related to plasmid conjugation. They also carry genes which offer selective advantages for host bacteria, e.g. antibiotic resistance genes (Dobrindt et al., 2004; Juhas et al., 2009). Similar to plasmids, genomic islands can also be transferred between bacteria. Interestingly, in *A. tumefaciens* C58, only 3,000 bp upstream of oriT-like1, a tetracycline resistance gene is encoded (Atu4206) and 30 kbp downstream of $HS2_{LC}$, a chloramphenicol resistance gene. Thus, it might be the case that the hotspot region constitutes an ancient genomic island and the transfer to plant cells only happens "accidentally".

Several genes involved in *A. tumefaciens* pathogenicity are not located on the Ti-plasmid but on the chromosomes and the At-plasmid. The encoded proteins are mainly involved in attachment of the bacterium to the plant cell and in the defence against plant toxic compounds (Rudder et al., 2014). For many of these genes the exact function during plant infection is not known. Most of them are not located on chromosomal regions which are transferred to plant cells. However, in *A. tumefaciens* nopaline strain C58 one of these genes, called *acvB*, is located on the circular chro-

mosome between the $HS1_{CC}$ integration site and RB-like2. Thus, this gene is likely transferred to plant cells. *acvB* is essential for virulence of the bacterium (Wirawan et al., 1993). Interestingly, in nopaline type *A. tumefaciens* strains *acvB* is only located on the circular chromosome while in octopine type strains an additional homologue called *virJ* is located on the Ti-plasmid (Kalogeraki and Winans, 1995). Since this gene is important for virulence, nopaline strains rely on the chromosomal copy of this gene. It would be imaginable that *A. tumefaciens* found a way to transfer this gene to other bacteria which do not posses a version of it. By this, the number of bacteria which are able to induce tumour formation by an nopaline strain Ti-plasmid would be increased. In this scenario, the occasional transfer to plant cells would just constitute an artefact.

The transfer of *A*chrDNA could be involved in plant transformation

However, it is tempting to speculate that the transfer of *A*chrDNA to plants is an additional and so far unnoticed aspect of *A. tumefaciens* mediated plant transformation.

The genome of *A. tumefaciens* strain C58 encodes 153 complete ABC transporters. This number is higher than in any other sequenced organism and more than double the number found in any sequenced bacterium (Wood et al., 2001). The reason for this high amount is unknown. ABC transporters were also among the most frequently transferred protein classes from *A. tumefaciens* to plant cells (Ülker et al., 2008a). Even though this is statistically expected, the transporters might serve a purpose in the plant cell. ABC transporters are known to transport a wide range of substrates like for example mono- or oligosaccharides, amino acids, peptides and also opines (Davidson et al., 2008). The fact that *A. tumefaciens* transfers ABC transporter encoding genes from its chromosomes to plant cells might hind towards a role of these transporters in the transformation process.

For example, until today it is not known, how the by *A. tumefaciens* induced opines are exported from plant cells and become available for the bacterium (Flores-Mireles et al., 2012). Opines are a class of compounds which are normally unknown to plants. Therefore, it is unlikely that a plant transport system would recognize them as substrates. Early studies suggested that the T-DNA encoded *ons* gene is involved in this process (Messens et al., 1985). However, this observation could not be confirmed by later studies (Flores-Mireles et al., 2012). It might be that the via *A*chrDNA transferred ABC transporter genes are expressed in plant cells and facilitate the export of opines. This export mechanism might only be necessary in certain host species where

the export of opines is not given. The chromosomal encoded ABC transporters could substitute for a missing opine export mechanism in plants, similar to VirE3 which can substitute for a missing VIP1 protein (Lacroix et al., 2005) or VirF which can substitute for missing VBF of the proteasome machinery (Zaltsman et al., 2010). In this study, three bacterial chromosomal promoters were tested for their ability to control gene expression in plant cells. None of them led to transcription of a *gfp* gene. Since the *A*chrDNA fragments could be large, more promoters should be tested.

As described in section 1.1.3, the plant cell transfers the non-proteinogenic amino acid GABA (gamma-Aminobutyric acid) to *Agrobacterium*. In the bacterium, GABA represses the by quorum sensing activated conjugation of the Ti-plasmid. Thereby it decreases the number of pathogenic bacteria (Chevrot et al., 2006). In this study it was shown that the chromosomal region between the $HS1_{LC}$ and the $HS2_{LC}$ locus is transferred to plant cells. Interestingly, this region contains the *attK* gene. *attK* is a succinate semialdehyde dehydrogenase (Carlier et al., 2004). This class of enzymes is also involved in the catabolism of GABA (Bown and Shelp, 1997). Thus, if the transferred *attK* gene is expressed in the plant cell, it might be involved in the degradation of GABA in the plant cell. This way, *Agrobacterium* would prevent the export of GABA and thereby prevent repression of its quorum sensing system by the plant.

Alternatively, the transfer might interfere with the host's defence mechanism. Bacteria are known to transfer a wide range of so called effector proteins to plant cells via type III (sometimes type IV) secretion systems and thereby modulate the plant's defence response in different ways (Cascales and Christie, 2003; Grant et al., 2006). For example, the *Pseudomonas syringae* effectors AvrPto, AvrRpt2 and AvrRpm1 inhibit the plant's defence response which is normally triggered by the recognition of pathogen associated molecular patterns (PAMPs) by surface receptors (Hauck et al., 2003; Kim et al., 2005). Similarly, *A. thaliana* knock out mutants for such a receptor are hypersensitive to *Agrobacterium*-mediated transformation (Zipfel et al., 2006). It might be that by the transfer of *A*chrDNA, *A. tumefaciens* found an alternative way how to get effector proteins into the host. Instead of directly delivering the proteins through a secretion channel, it integrates effector encoding genes into the hosts genome. In many studies it has been shown that *Agrobacterium* and host plant are in a constant "arms race" (reviewed in Pitzschke (2013)). The transfer of *A*chrDNA to plant cells might constitute an additional way of *A. tumefaciens* to increase the plant's susceptibility for transformation.

In order to test this, the oriT-like1 sequence could be deleted from the genome

of an *A. tumefaciens* wild type strain. This strain could be used to induce tumour formation in different plants. Comparing the formation of tumours of this ΔoriT-like1 strain to one in which oriT-like1 was not deleted could lead to insights into the role of the transfer of *A*chrDNA.

List of Figures

References

Al Ghaithi, H. (2012). An attempt to characterize the environmental triggers of IS*426* activity in *Agrobacterium tumefaciens*. *Bachelor's thesis, University of Applied Sciences Bonn-Rhein-Sieg*.

Albright, L. M., Yanofsky, M. F., Leroux, B., Ma, D. Q., and Nester, E. W. (1987). Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. *Journal of bacteriology*, 169:1046–1055.

Alonso, J. M. and Ecker, J. R. (2006). Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nature reviews. Genetics*, 7:524–536.

Alonso, J. M., Stepanova, A. N., Leisse, T. J., Kim, C. J., Chen, H., Shinn, P., Stevenson, D. K., Zimmerman, J., Barajas, P., Cheuk, R., Gadrinab, C., Heller, C., Jeske, A., Koesema, E., Meyers, C. C., Parker, H., Prednis, L., Ansari, Y., Choy, N., Deen, H., Geralt, M., Hazari, N., Hom, E., Karnes, M., Mulholland, C., Ndubaku, R., Schmidt, I., Guzman, P., Aguilar-Henonin, L., Schmid, M., Weigel, D., Carter, D. E., Marchand, T., Risseeuw, E., Brogden, D., Zeko, A., Crosby, W. L., Berry, C. C., and Ecker, J. R. (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science (New York, N.Y.)*, 301(5633):653–657.

Altabe, S., Iñón de Iannino, N., de Mendoza, D., and Ugalde, R. A. (1990). Expression of the *Agrobacterium tumefaciens chvB* virulence region in Azospirillum spp. *Journal of bacteriology*, 172:2563–2567.

Alvarez-Martinez, C. E. and Christie, P. J. (2009). Biological diversity of prokaryotic type IV secretion systems. *Microbiology and molecular biology reviews*, 73(4):775–808.

Aly, K. A. and Baron, C. (2007). The VirB5 protein localizes to the T-pilus tips in *Agrobacterium tumefaciens*. *Microbiology*, 153:3766–3775.

Anand, A., Krichevsky, A., Schornack, S., Lahaye, T., Tzfira, T., Tang, Y., Citovsky, V., and Mysore, K. S. (2007). *Arabidopsis* VIRE2 INTERACTING PROTEIN2 is required for *Agrobacterium* T-DNA integration in plants. *The Plant cell*, 19:1695–1708.

Anderson, L. B., Hertzel, A. V., and Das, A. (1996). *Agrobacterium tumefaciens* VirB7 and VirB9 form a disulfide-linked protein complex. *Proceedings of the National Academy of Sciences of the United States of America*, 93:8889–8894.

Ankenbauer, R. G. and Nester, E. W. (1990). Sugar-mediated induction of *Agrobacterium tumefaciens* virulence genes: structural specificity and activities of monosaccharides. *Journal of bacteriology*, 172:6442–6446.

Arntzen, C., Plotkin, S., and Dodet, B. (2005). Plant-derived vaccines and antibodies: Potential and limitations. *Vaccine*, 23:1753–1756.

Atmakuri, K., Cascales, E., and Christie, P. J. (2004). Energetic components VirD4, VirB11 and VirB4 mediate early DNA transfer reactions required for bacterial type IV secretion. *Molecular microbiology*, 54:1199–1211.

Atmakuri, K., Ding, Z., and Christie, P. J. (2003). VirE2, a type IV secretion substrate, interacts with the VirD4 transfer protein at cell poles of *Agrobacterium tumefaciens*. *Molecular microbiology*, 49:1699–1713.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1987). *Current Protocols in Molecular Biology*, volume 1. Wiley Online Library.

Backert, S., Fronzes, R., and Waksman, G. (2008). VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? *Trends in Microbiology*, 16:409–413.

Bakó, L., Umeda, M., Tiburcio, A. F., Schell, J., and Koncz, C. (2003). The VirD2 pilot protein of *Agrobacterium*-transferred DNA interacts with the TATA box-binding protein and a nuclear protein kinase in plants. *Proceedings of the National Academy of Sciences of the United States of America*, 100:10108–10113.

Ballas, N. and Citovsky, V. (1997). Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. *Proceedings of the National Academy of Sciences of the United States of America*, 94:10723–10728.

Baron, C., Thorstenson, Y. R., and Zambryski, P. C. (1997). The lipoprotein VirB7 interacts with VirB9 in the membranes of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 179:1211–1218.

Bechtold, N., Ellis, J., and Pelletier, G. (1993). In planta *Agrobacterium* mediated gene transfer by infiltration of adult *Arabidopsis thaliana* plants. *Comptes Rendus de l'Academie des Sciences Serie III Sciences de la Vie*, 316:1194–1199.

Bechtold, N., Jaudeau, B., Jolivet, S., Maba, B., Vezon, D., Voisin, R., and Pelletier, G. (2000). The maternal chromosome set is the target of the T-DNA in the in planta transformation of *Arabidopsis thaliana*. *Genetics*, 155:1875–1887.

Beijersbergen, A., Dulk-Ras, A. D., Schilperoort, R. A., and Hooykaas, P. J. (1992). Conjugative Transfer by the Virulence System of *Agrobacterium tumefaciens*. *Science (New York, N.Y.)*, 256(5061):1324–1327.

Berger, B. R. and Christie, P. J. (1994). Genetic complementation analysis of the *Agrobacterium tumefaciens virB* operon: *virB2* through *virB11* are essential virulence genes. *Journal of bacteriology*, 176:3646–3660.

Bevan, M. (1984). Binary *Agrobacterium* vectors for plant transformation. *Nucleic acids research*, 12(22):8711–8721.

Bevan, M. W., Flavell, R. B., and Chilton, M.-D. (1983). A chimaeric antibiotic resistance gene as a selectable marker for plant cell transformation. *Nature*, 304(5922):184–187.

Bhattacharjee, S., Lee, L.-Y., Oltmanns, H., Cao, H., Veena, G., Cuperus, J., and Gelvin, S. B. (2008). IMPa-4, an *Arabidopsis* importin alpha isoform, is preferentially involved in *Agrobacterium*-mediated plant transformation. *The Plant cell*, 20:2661–2680.

Bouzar, H. and Jones, J. B. (2001). *Agrobacterium larrymoorei sp.* nov., a pathogen isolated from aerial tumours of *Ficus benjamina*. *International journal of systematic and evolutionary microbiology*, 51:1023–1026.

Bown, A. W. and Shelp, B. J. (1997). The Metabolism and Functions of [gamma]-Aminobutyric Acid. *Plant physiology*, 115(1):1–5.

Braun, A. C. (1947). Thermal studies on the factors responsible for tumor initiation in crown gall. *American journal of botany*, 34(4):234–240.

Braun, A. C. (1958). A Physiological Basis for Autonomous Growth of the Crown-Gall Tumor Cell. *Proceedings of the National Academy of Sciences of the United States of America*, 44(4):344–349.

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. *Journal of bacteriology*, 181(18):5758–5765.

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*. *Molecular plant-microbe interactions*, 11(1):57–63.

Brencic, A., Eberhard, A., and Winans, S. C. (2004a). Signal quenching, detoxification and mineralization of vir gene-inducing phenolics by the VirH2 protein of *Agrobacterium tumefaciens*. *Molecular microbiology*, 51:1103–1115.

Brencic, A., Xia, Q., and Winans, S. C. (2004b). VirA of *Agrobacterium tumefaciens* is an intradimer transphosphorylase and can actively block vir gene expression in the absence of phenolic signals. *Molecular microbiology*, 52:1349–1362.

Brocklehurst, K. R. and Morby, A. P. (2000). Metal-ion tolerance in *Escherichia coli*: analysis of transcriptional profiles by gene-array technology. *Microbiology*, 146 (Pt 9:2277–2282.

Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., Lepiniec, L., Caboche, M., and Lecharny, A. (2002). T-DNA integration into the *Arabidopsis* genome depends on sequences of pre-insertion sites. *EMBO reports*, 3:1152–1157.

Buchanan-Wollaston, V., Passiatore, J. E., and Cannon, F. (1987). The mob and oriT mobilization functions of a bacterial plasmid promote its transfer to plants. *Nature*, 328:172–175.

Bundock, P., den Dulk-Ras, A., Beijersbergen, A., and Hooykaas, P. J. (1995). Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *The EMBO journal*, 14:3206–3214.

Bundock, P. and Hooykaas, P. J. (1996). Integration of *Agrobacterium tumefaciens* T-DNA in the *Saccharomyces cerevisiae* genome by illegitimate recombination. *Proceedings of the National Academy of Sciences of the United States of America*, 93:15272–15275.

Cangelosi, G. A., Martinetti, G., Leigh, J. A., Lee, C. C., Thienes, C., and Nester, E. W. (1989). Role for *Agrobacterium tumefaciens* ChvA protein in export of beta-1,2-glucan. *Journal of bacteriology*, 171:1609–1615.

Carlier, A., Chevrot, R., Dessaux, Y., and Faure, D. (2004). The assimilation of gamma-butyrolactone in *Agrobacterium tumefaciens* C58 interferes with the accumulation of the N-acyl-homoserine lactone signal. *Molecular plant-microbe interactions*, 17(9):951–957.

Cascales, E. and Christie, P. J. (2003). The versatile bacterial type IV secretion systems. *Nature reviews. Microbiology*, 1:137–149.

Cascales, E. and Christie, P. J. (2004). Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science (New York, N.Y.)*, 304:1170–1173.

Chang, C. H. and Winans, S. C. (1992). Functional roles assigned to the periplasmic, linker, and receiver domains of the *Agrobacterium tumefaciens* VirA protein. *Journal of bacteriology*, 174:7033–7039.

Charlier, D., Piette, J., and Glansdorff, N. (1982). IS*3* can function as a mobile promoter in E. coli. *Nucleic acids research*, 10:5935–5948.

Chen, L., Chen, Y., Wood, D. W., and Nester, E. W. (2002). A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. *Journal of bacteriology*, 184(17):4838–4845.

Cheng, M., Fry, J. E., Pang, S., Zhou, H., Hironaka, C. M., Duncan, D. R., Conner, T. W., and Wan, Y. (1997). Genetic Transformation of Wheat Mediated by *Agrobacterium tumefaciens*. *Plant physiology*, 115:971–980.

Chevrot, R., Rosen, R., Haudecoeur, E., Cirou, A., Shelp, B. J., Ron, E., and Faure, D. (2006). GABA controls the level of quorum-sensing signal in *Agrobacterium tumefaciens*. *Proceedings of the National Academy of Sciences of the United States of America*, 103:7460–7464.

Chilton, M. D., Currier, T. C., Farrand, S. K., Bendich, A. J., Gordon, M. P., and Nester, E. W. (1974). *Agrobacterium tumefaciens* DNA and PS8 bacteriophage DNA not detected in crown gall tumors. *Proceedings of the National Academy of Sciences of the United States of America*, 71(9):3672–3676.

Chilton, M.-D., Drummond, M. H., Merlo, D. J., Sciaky, D., Montoya, A. L., Gordon, M. P., and Nester, E. W. (1977). Stable incorporation of plasmid DNA into higher plant cells: the molecular basis of crown gall tumorigenesis. *Cell*, 11:263–271.

Chilton, M.-D. M. and Que, Q. (2003). Targeted integration of T-DNA into the tobacco genome at double-stranded breaks: new insights on the mechanism of T-DNA integration. *Plant physiology*, 133:956–965.

Ciampi, M. S., Schmid, M. B., and Roth, J. R. (1982). Transposon Tn10 provides a promoter for transcription of adjacent sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 79:5016–5020.

Cirillo, J. D., Barletta, R. G., Bloom, B. R., and Jacobs, W. R. (1991). A novel transposon trap for mycobacteria: isolation and characterization of IS*1096*. *Journal of bacteriology*, 173:7772–7780.

Citovsky, V., Lee, L. Y., Vyas, S., Glick, E., Chen, M. H., Vainstein, A., Gafni, Y., Gelvin, S. B., and Tzfira, T. (2006). Subcellular Localization of Interacting Proteins by Bimolecular Fluorescence Complementation in Planta. *Journal of Molecular Biology*, 362:1120–1131.

Clough, S. J. and Bent, A. F. (1998). Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. *The Plant journal : for cell and molecular biology*, 16(6):735–743.

Conn, H. J. (1942). Validity of the Genus Alcaligenes. *Journal of bacteriology*, 44(3):353–360.

Costa, E. D., Chai, Y., and Winans, S. C. (2012). The quorum-sensing protein TraR of *Agrobacterium tumefaciens* is susceptible to intrinsic and TraM-mediated proteolytic instability. *Molecular Microbiology*, 84(5):807–815.

Costantino, P., Hooykaas, P. J., den Dulk-Ras, H., and Schilperoort, R. A. (1980). Tumor formation and rhizogenicity of *Agrobacterium rhizogenes* carrying Ti plasmids. *Gene*, 11:79–87.

Dang, T. A. and Christie, P. J. (1997). The VirB4 ATPase of *Agrobacterium tumefaciens* is a cytoplasmic membrane protein exposed at the periplasmic surface. *Journal of bacteriology*, 179:453–462.

Danhorn, T. and Fuqua, C. (2007). Biofilm formation by plant-associated bacteria. *Annual review of microbiology*, 61:401–422.

Davidson, A. L., Dassa, E., Orelle, C., and Chen, J. (2008). Structure, function, and evolution of bacterial ATP-binding cassette systems. *Microbiology and molecular biology reviews*, 72(2):317–364.

Dawson, R. J. P. and Locher, K. P. (2006). Structure of a bacterial multidrug ABC transporter. *Nature*, 443(7108):180–185.

De Neve, M., De Buck, S., Jacobs, A., Van Montagu, M., and Depicker, A. (1997). T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. *The Plant journal : for cell and molecular biology*, 11(1):15–29.

Dempsey, L. A. and Dubnau, D. A. (1989). Identification of plasmid and *Bacillus subtilis* chromosomal recombination sites used for pE194 integration. *Journal of bacteriology*, 171(5):2856–2865.

Deng, W., Chen, L., Wood, D. W., Metcalfe, T., Liang, X., Gordon, M. P., Comai, L., and Nester, E. W. (1998). *Agrobacterium* VirD2 protein interacts with plant host cyclophilins. *Proceedings of the National Academy of Sciences of the United States of America*, 95:7040–7045.

Desfeux, C., Clough, S. J., and Bent, A. F. (2000). Female reproductive tissues are the primary target of *Agrobacterium*-mediated transformation by the *Arabidopsis* floral-dip method. *Plant physiology*, 123:895–904.

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

Dobrindt, U., Hochhut, B., Hentschel, U., and Hacker, J. (2004). Genomic islands in pathogenic and environmental microorganisms. *Nature reviews. Microbiology*, 2(5):414–424.

Douglas, C. J., Halperin, W., and Nester, E. W. (1982). *Agrobacterium tumefaciens* mutants affected in attachment to plant cells. *Journal of bacteriology*, 152:1265–1275.

Dube, T., Kovalchuk, I., Hohn, B., and Thomson, J. A. (2004). *Agrobacterium tumefaciens*-mediated transformation of plants by the pTF-FC2 plasmid is efficient and strictly dependent on the MobA protein. *Plant molecular biology*, 55(4):531–539.

Duval-Valentin, G., Marty-Cointin, B., and Chandler, M. (2004). Requirement of IS*911* replication before integration defines a new bacterial transposition pathway. *The EMBO journal*, 23(19):3897–3906.

Eichenbaum, Z. and Livneh, Z. (1998). UV light induces IS*10* transposition in *Escherichia coli*. *Genetics*, 149:1173–1181.

Farrand, S. K., Van Berkum, P. B., and Oger, P. (2003). *Agrobacterium* is a definable genus of the family Rhizobiaceae. *International journal of systematic and evolutionary microbiology*, 53:1681–1687.

Filichkin, S. A. and Gelvin, S. B. (1993). Formation of a putative relaxation intermediate during T-DNA processing directed by the Agrobacterium tumefaciens VirD1,D2 endonuclease. *Molecular microbiology*, 8(5):915–926.

Flores-Mireles, A. L., Eberhard, A., and Winans, S. C. (2012). *Agrobacterium tumefaciens* can obtain sulphur from an opine that is synthesized by octopine synthase using S-methylmethionine as a substrate. *Molecular microbiology*, 84(5):845–856.

Fortin, C., Marquis, C., Nester, E. W., and Dion, P. (1993). Dynamic structure of *Agrobacterium tumefaciens* Ti plasmids. *Journal of bacteriology*, 175:4790–4799.

Fraley, R. T., Rogers, S. G., Horsch, R. B., Eichholtz, D. A., Flick, J. S., Fink, C. L., Hoffmann, N. L., and Sanders, P. R. (1985). The SEV System: A New Disarmed Ti Plasmid Vector System for Plant Transformation. *Nature biotechnology*, 3:629–635.

Fraley, R. T., Rogers, S. G., Horsch, R. B., Sanders, P. R., Flick, J. S., Adams, S. P., Bittner, M. L., Brand, L. A., Fink, C. L., Fry, J. S., Galluppi, G. R., Goldberg, S. B., Hoffmann, N. L., and Woo, S. C. (1983). Expression of bacterial genes in plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, 80(15):4803–4807.

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 : for cell and molecular biology*, 41:464–477.

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 (New York, N.Y.)*, 323:266–268.

Fuqua, W. C., Winans, S. C., and Greenberg, E. P. (1994). Quorum sensing in bacteria: the LuxR-LuxI family of cell density-responsive transcriptional regulators. *Journal of bacteriology*, 176:269–275.

Gao, R., Mukhopadhyay, A., Fang, F., and Lynn, D. G. (2006). Constitutive activation of two-component response regulators: characterization of VirG activation in *Agrobacterium tumefaciens*. *Journal of bacteriology*, 188:5204–5211.

Gelvin, S. B. (1998). *Agrobacterium* VirE2 proteins can form a complex with T strands in the plant cytoplasm. *Journal of bacteriology*, 180:4300–4302.

Gelvin, S. B. (2003). *Agrobacterium*-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiology and molecular biology reviews*, 67(1):16–37.

Gelvin, S. B. (2008). *Agrobacterium*-mediated DNA transfer, and then some. *Nature biotechnology*, 26(9):998–1000.

Gelvin, S. B. (2010). Plant proteins involved in *Agrobacterium*-mediated genetic transformation. *Annual review of phytopathol*, 48:45–68.

Gelvin, S. B. (2012). Traversing the Cell: *Agrobacterium* T-DNA's Journey to the Host Genome. *Frontiers in plant science*, 3:52.

Grant, S. R., Fisher, E. J., Chang, J. H., Mole, B. M., and Dangl, J. L. (2006). Subterfuge and manipulation: type III effector proteins of phytopathogenic bacteria. *Annual review of microbiology*, 60:425–449.

Gray, K. M. and Garey, J. R. (2001). The evolution of bacterial LuxI and LuxR quorum sensing regulators. *Microbiology*, 147:2379–2387.

Grove, J. I., Alandiyjany, M. N., and Delahay, R. M. (2013). Site-specific Relaxase Activity of a VirD2-like Protein Encoded within the tfs4 Genomic Island of *Helicobacter pylori*. *The Journal of biological chemistry*, 288(37):26385–26396.

Guilhot, C., Gicquel, B., Davies, J., and Martín, C. (1992). Isolation and analysis of IS*6120*, a new insertion sequence from *Mycobacterium smegmatis*. *Molecular microbiology*, 6:107–113.

Habeeb, L. F., Wang, L., and Winans, S. C. (1991). Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein. *Molecular plant-microbe interactions*, 4(4):379–385.

Hagblom, P., Korch, C., Jonsson, A. B., and Normark, S. (1986). Intragenic variation by site-specific recombination in the cryptic plasmid of *Neisseria gonorrhoeae*. *Journal of bacteriology*, 167(1):231–237.

Hagen, S. J., Son, M., Weiss, J. T., and Young, J. H. (2010). Bacterium in a box: sensing of quorum and environment by the LuxI/LuxR gene regulatory circuit. *Journal of biological physics*, 36:317–327.

Hanson, B., Engler, D., Moy, Y., Newman, B., Ralston, E., and Gutterson, N. (1999). A simple method to enrich an *Agrobacterium*-transformed population for plants containing only T-DNA sequences. *The Plant journal : for cell and molecular biology*, 19(6):727–734.

Hartl, D. L., Lozovskaya, E. R., and Lawrence, J. G. (1992). Nonautonomous transposable elements in prokaryotes and eukaryotes. *Genetica*, 86:47–53.

Hattori, Y., Iwata, K., Suzuki, K., Uraji, M., Ohta, N., Katoh, A., and Yoshida, K. (2001). Sequence characterization of the vir region of a nopaline type Ti plasmid, pTi-SAKURA. *Genes & Genetic Systems*, 76(2):121–130.

Hauck, P., Thilmony, R., and He, S. Y. (2003). A *Pseudomonas syringae* type III effector suppresses cell wall-based extracellular defense in susceptible *Arabidopsis* plants. *Proceedings of the National Academy of Sciences of the United States of America*, 100:8577–8582.

Haudecoeur, E., Tannières, M., Cirou, A., Raffoux, A., Dessaux, Y., and Faure, D. (2009). Different regulation and roles of lactonases AiiB and AttM in *Agrobacterium tumefaciens* C58. *Molecular plant-microbe interactions*, 22:529–537.

He, F., Nair, G. R., Soto, C. S., Chang, Y., Hsu, L., Ronzone, E., DeGrado, W. F., and Binns, A. N. (2009). Molecular basis of ChvE function in sugar binding, sugar utilization, and virulence in *Agrobacterium tumefaciens*. *Journal of bacteriology*, 191:5802–5813.

Hellens, R., Mullineaux, P., and Klee, H. (2000). Technical Focus:A guide to *Agrobacterium* binary Ti vectors. *Trends in plant science*, 5(10):446–451.

Herrera-Estrella, A., Chen, Z. M., Van Montagu, M., and Wang, K. (1988). VirD proteins of *Agrobacterium tumefaciens* are required for the formation of a covalent DNA–protein complex at the 5' terminus of T-strand molecules. *The EMBO journal*, 7(13):4055–4062.

Herrera-Estrella, L., Block, M. D., Messens, E., Hernalsteens, J. P., Montagu, M. V., and Schell, J. (1983). Chimeric genes as dominant selectable markers in plant cells. *The EMBO journal*, 2(6):987–995.

Hiei, Y., Ohta, S., Komari, T., and Kumashiro, T. (1994). Efficient transformation of rice (*Oryza sativa L.*) mediated by *Agrobacterium* and sequence analysis of the boundaries of the T-DNA. *The Plant journal : for cell and molecular biology*, 6:271–282.

Hobbs, S. L., Warkentin, T. D., and DeLong, C. M. (1993). Transgene copy number can be positively or negatively associated with transgene expression. *Plant molecular biology*, 21:17–26.

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.

Höppner, C., Liu, Z., Domke, N., Binns, A. N., and Baron, C. (2004). VirB1 orthologs from Brucella suis and pKM101 complement defects of the lytic transglycosylase required for efficient type IV secretion from *Agrobacterium tumefaciens*. *Journal of bacteriology*, 186:1415–1422.

Horsch, R. B. and Klee, H. J. (1986). Rapid assay of foreign gene expression in leaf discs transformed by *Agrobacterium tumefaciens*: Role of T-DNA borders in the transfer process. *Proceedings of the National Academy of Sciences of the United States of America*, 83(12):4428–4432.

Hwang, H.-H. and Gelvin, S. B. (2004). Plant proteins that interact with VirB2, the *Agrobacterium tumefaciens* pilin protein, mediate plant transformation. *The Plant cell*, 16:3148–3167.

Hwang, I., Li, P. L., Zhang, L., Piper, K. R., Cook, D. M., Tate, M. E., and Farrand, S. K. (1994). TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. *Proceedings of the National Academy of Sciences of the United States of America*, 91:4639–4643.

Ishida, Y., Saito, H., Ohta, S., Hiei, Y., Komari, T., and Kumashiro, T. (1996). High efficiency transformation of maize (*Zea mays L.*) mediated by *Agrobacterium tumefaciens*. *Nature biotechnology*, 14:745–750.

Itoh, Y., Watson, J. M., Haas, D., and Leisinger, T. (1984). Genetic and molecular characterization of the *Pseudomonas* plasmid pVS1. *Plasmid*, 11:206–220.

Jakubowski, S. J., Krishnamoorthy, V., Cascales, E., and Christie, P. J. (2004). *Agrobacterium tumefaciens* VirB6 domains direct the ordered export of a DNA substrate through a type IV secretion system. *Journal of Molecular Biology*, 341:961–977.

Jarchow, E., Grimsley, N. H., and Hohn, B. (1991). *virF*, the host-range-determining virulence gene of *Agrobacterium tumefaciens*, affects T-DNA transfer to *Zea mays*. *Proceedings of the National Academy of Sciences of the United States of America*, 88:10426–10430.

Jasper, F., Koncz, C., Schell, J., and Steinbiss, H. H. (1994). *Agrobacterium* T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proceedings of the National Academy of Sciences of the United States of America*, 91:694–698.

Jayaraman, R. (2011). Hypermutation and stress adaptation in bacteria. *Journal of genetics*, 90(2):383–391.

Jin, S., Roitsch, T., Ankenbauer, R. G., Gordon, M. P., and Nester, E. W. (1990). The VirA protein of *Agrobacterium tumefaciens* is autophosphorylated and is essential for vir gene regulation. *Journal of bacteriology*, 172:525–530.

Jones, A. L., Shirasu, K., and Kado, C. I. (1994). The product of the virB4 gene of *Agrobacterium tumefaciens* promotes accumulation of VirB3 protein. *Journal of bacteriology*, 176:5255–5261.

Jorgensen, R. A., Cluster, P. D., English, J., Que, Q., and Napoli, C. A. (1996). Chalcone synthase cosuppression phenotypes in petunia flowers: comparison of sense vs. antisense constructs and single-copy vs. complex T-DNA sequences. *Plant molecular biology*, 31(5):957–973.

Joubert, P., Beaupère, D., Wadouachi, A., Chateau, S., Sangwan, R. S., and Sangwan-Norreel, B. S. (2004). Effect of phenolic glycosides on *Agrobacterium tumefaciens virH* gene induction and plant transformation. *Journal of natural products*, 67:348–351.

Judd, P. K., Kumar, R. B., and Das, A. (2005). Spatial location and requirements for the assembly of the *Agrobacterium tumefaciens* type IV secretion apparatus. *Proceedings of the National Academy of Sciences of the United States of America*, 102:11498–11503.

Juhas, M., van der Meer, J. R., Gaillard, M., Harding, R. M., Hood, D. W., and Crook, D. W. (2009). Genomic islands: tools of bacterial horizontal gene transfer and evolution. *FEMS microbiology reviews*, 33(2):376–393.

Kalogeraki, V. S. and Winans, S. C. (1995). The octopine-type Ti plasmid pTiA6 of *Agrobacterium tumefaciens* contains a gene homologous to the chromosomal virulence gene *acvB*. *Journal of bacteriology*, 177:892–897.

Kerr, A. (1971). Acquisition of virulence by non-pathogenic isolates of *Agrobacterium radiobacter*. *Physiological Plant Pathology*, 1(3):241–246.

Kim, M. G., Da Cunha, L., McFall, A. J., Belkhadir, Y., DebRoy, S., Dangl, J. L., and Mackey, D. (2005). Two *Pseudomonas syringae* type III effectors inhibit RIN4-regulated basal defense in *Arabidopsis*. *Cell*, 121:749–759.

Kim, S.-I., Veena, G., 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 journal : for cell and molecular biology*, 51:779–791.

Kim, S.-R. and An, G. (2012). Bacterial transposons are co-transferred with T-DNA to rice chromosomes during *Agrobacterium*-mediated transformation. *Molecules and cells*, 33(6):583–589.

Komori, T., Imayama, T., Kato, N., Ishida, Y., Ueki, J., and Komari, T. (2007). Current status of binary vectors and superbinary vectors. *Plant physiology*, 145(4):1155–1160.

Koncz, C. and Schell, J. (1986). The promoter of TL-DNA gene 5 controls the tissue-specific expression of chimaeric genes carried by a novel type of *Agrobacterium* binary vector. *Molecular and General Genetics*, 204(3):383–396.

Kononov, M. E., Bassuner, B., and Gelvin, S. B. (1997). Integration of T-DNA binary vector âĂŸbackboneâĂŹ sequences into the tobacco genome: evidence for multiple complex patterns of integration. *The Plant journal : for cell and molecular biology*, 11(5):945–957.

Krishnamohan, A., Balaji, V., and Veluthambi, K. (2001). Efficient vir gene induction in *Agrobacterium tumefaciens* requires *virA*, *virG*, and vir box from the same Ti plasmid. *Journal of bacteriology*, 183:4079–4089.

Krysan, P. J. (1999). T-DNA as an Insertional Mutagen in *Arabidopsis*. *The Plant cell*, 11(12):2283–2290.

Kumar, R. B., Xie, Y. H., and Das, A. (2000). Subcellular localization of the *Agrobacterium tumefaciens* T-DNA transport pore proteins: VirB8 is essential for the assembly of the transport pore. *Molecular microbiology*, 36:608–617.

Kunik, T., Tzfira, T., Kapulnik, Y., Gafni, Y., Dingwall, C., and Citovsky, V. (2001). Genetic transformation of HeLa cells by *Agrobacterium*. *Proceedings of the National Academy of Sciences of the United States of America*, 98:1871–1876.

Lacroix, B., Loyter, A., and Citovsky, V. (2008). Association of the *Agrobacterium* T-DNA-protein complex with plant nucleosomes. *Proceedings of the National Academy of Sciences of the United States of America*, 105:15429–15434.

Lacroix, B., Vaidya, M., Tzfira, T., and Citovsky, V. (2005). The VirE3 protein of *Agrobacterium* mimics a host cell function required for plant genetic transformation. *The EMBO journal*, 24:428–437.

Lai, E.-M., Eisenbrandt, R., Kalkum, M., Lanka, E., and Kado, C. I. (2002). Biogenesis of T pili in *Agrobacterium tumefaciens* requires precise VirB2 propilin cleavage and cyclization. *Journal of bacteriology*, 184:327–330.

Lai, E. M. and Kado, C. I. (1998). Processed VirB2 is the major subunit of the promiscuous pilus of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 180:2711–2717.

Lee, L.-Y., Fang, M.-J., Kuang, L.-Y., and Gelvin, S. B. (2008). Vectors for multi-color bimolecular fluorescence complementation to investigate protein-protein interactions in living plant cells. *Plant methods*, 4:24.

Lee, L.-Y. and Gelvin, S. B. (2008). T-DNA binary vectors and systems. *Plant physiology*, 146(2):325–332.

Lee, L.-Y., Humara, J. M., and Gelvin, S. B. (2001). Novel Constructions to Enable the Integration of Genes into the *Agrobacterium tumefaciens* C58 Chromosome. *Molecular plant-microbe interactions*, 14(4):577–579.

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*. *Proceedings of the National Academy of Sciences of the United States of America*, 92:12245–12249.

Leloup, L., Lai, E. M., and Kado, C. I. (2002). Identification of a chromosomal tra-like region in *Agrobacterium tumefaciens*. *Molecular genetics and genomics*, 267(1):115–123.

Lessl, M. and Lanka, E. (1994). Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell*, 77:321–324.

Levy, S. B. (1992). Active efflux mechanisms for antimicrobial resistance. *Antimicrobial agents and chemotherapy*, 36(4):695–703.

Lewis, L. A., Astatke, M., Umekubo, P. T., Alvi, S., Saby, R., and Afrose, J. (2011). Soluble expression, purification and characterization of the full length IS*2* Transposase. *Mobile DNA*, 2:14.

Lewis, L. A., Cylin, E., Lee, H. K., Saby, R., Wong, W., and Grindley, N. D. F. (2004). The left end of IS*2*: a compromise between transpositional activity and an essential promoter function that regulates the transposition pathway. *Journal of bacteriology*, 186:858–865.

Lewis, L. A. and Grindley, N. D. (1997). Two abundant intramolecular transposition products, resulting from reactions initiated at a single end, suggest that IS*2* transposes by an unconventional pathway. *Molecular microbiology*, 25:517–529.

Li, J., Krichevsky, A., Vaidya, M., Tzfira, T., and Citovsky, V. (2005a). Uncoupling of the functions of the *Arabidopsis* VIP1 protein in transient and stable plant genetic transformation by *Agrobacterium*. *Proceedings of the National Academy of Sciences of the United States of America*, 102:5733–5738.

Li, J., Vaidya, M., White, C., Vainstein, A., Citovsky, V., and Tzfira, T. (2005b). Involvement of KU80 in T-DNA integration in plant cells. *Proceedings of the National Academy of Sciences of the United States of America*, 102:19231–19236.

Li, P. L., Everhart, D. M., and Farrand, S. K. (1998). Genetic and sequence analysis of the pTiC58 trb locus, encoding a mating-pair formation system related to members of the type IV secretion family. *Journal of bacteriology*, 180:6164–6172.

Li, Y., Rosso, M. G., Ülker, B., and Weisshaar, B. (2006). Analysis of T-DNA insertion site distribution patterns in *Arabidopsis thaliana* reveals special features of genes without insertions. *Genomics*, 87:645–652.

Llop, P., Murillo, J., Lastra, B., and Lopez, M. M. (2009). Recovery of nonpathogenic mutant bacteria from tumors caused by several *Agrobacterium tumefaciens* strains: a frequent event? *Applied and environmental microbiology*, 75(20):6504–6514.

Logemann, E., Birkenbihl, R. P., Ulker, B., and Somssich, I. E. (2006). An improved method for preparing *Agrobacterium* cells that simplifies the *Arabidopsis* transformation protocol. *Plant methods*, 2:16.

Loper, J. E. and Kado, C. I. (1979). Host range conferred by the virulence-specifying plasmid of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 139(2):591–596.

López de Felipe, F., Magni, C., de Mendoza, D., and López, P. (1996). Transcriptional activation of the citrate permease P gene of *Lactococcus lactis* biovar diacetylactis by an insertion sequence-like element present in plasmid pCIT264. *Molecular and General Genetics*, 250:428–436.

Loyter, A., Rosenbluh, J., Zakai, N., Li, J., Kozlovsky, S. V., Tzfira, T., and Citovsky, V. (2005). The plant VirE2 interacting protein 1. a molecular link between the *Agrobacterium* T-complex and the host cell chromatin? *Plant physiology*, 138(3):1318–1321.

Lu, J., den Dulk-Ras, A., Hooykaas, P. J. J., and Glover, J. N. M. (2009). *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon-helix-helix DNA-binding fold. *Proceedings of the National Academy of Sciences of the United States of America*, 106:9643–9648.

Luo, Z. Q. and Farrand, S. K. (1999). Cloning and characterization of a tetracycline resistance determinant present in *Agrobacterium tumefaciens* C58. *Journal of bacteriology*, 181:618–626.

Ma, J. K.-C., Chikwamba, R., Sparrow, P., Fischer, R., Mahoney, R., and Twyman, R. M. (2005). Plant-derived pharmaceuticals–the road forward. *Trends in plant science*, 10:580–585.

Magori, S. and Citovsky, V. (2011). *Agrobacterium* Counteracts Host-Induced Degradation of Its Effector F-Box Protein. *Science Signaling*, 4:ra69.

Mahillon, J. and Chandler, M. (1998). Insertion Sequences. *Microbiology and molecular biology reviews*, 62(3):725–774.

Mansfield, J., Genin, S., Magori, S., Citovsky, V., Sriariyanum, M., Ronald, P., Dow, M., Verdier, V., Beer, S. V., Machado, M. a., Toth, I., Salmond, G., and Foster, G. D. (2012). Top 10 plant pathogenic bacteria in molecular plant pathology. *Molecular plant pathology*, 13:614–629.

Marks, J. R., Lynch, T. J., Karlinsey, J. E., and Thomashow, M. F. (1987). *Agrobacterium tumefaciens* virulence locus *pscA* is related to the *Rhizobium meliloti* exoC locus. *Journal of bacteriology*, 169:5835–5837.

Martineau, B., Voelker, T. A., and Sanders, R. A. (1994). On Defining T-DNA. *The Plant cell*, 6(8):1032–1033.

Matthysse, A. G., Yarnall, H., Boles, S. B., and McMahan, S. (2000). A region of the *Agrobacterium tumefaciens* chromosome containing genes required for virulence and attachment to host cells. *Biochimica et biophysica acta*, 1490(1-2):208–212.

McCullen, C. A. and Binns, A. N. (2006). *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. *Annual review of cell and developmental biology*, 22:101–127.

McElver, J., Tzafrir, I., Aux, G., Rogers, R., Ashby, C., Smith, K., Thomas, C., Schetter, A., Zhou, Q., Cushman, M. A., Tossberg, J., Nickle, T., Levin, J. Z., Law, M., Meinke, D., and Patton, D. (2001). Insertional mutagenesis of genes required for seed development in *Arabidopsis thaliana*. *Genetics*, 159:1751–1763.

Melchers, L. S., Maroney, M. J., den Dulk-Ras, A., Thompson, D. V., van Vuuren, H. A., Schilperoort, R. A., and Hooykaas, P. J. (1990). Octopine and nopaline strains of *Agrobacterium tumefaciens* differ in virulence; molecular characterization of the virF locus. *Plant molecular biology*, 14:249–259.

Melchers, L. S., Regensburg-Tuïnk, T. J., Bourret, R. B., Sedee, N. J., Schilperoort, R. A., and Hooykaas, P. J. (1989). Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *The EMBO journal*, 8:1919–1925.

Messens, E., Lenaerts, A., Montagu, M., and Hedges, R. W. (1985). Genetic basis for opine secretion from crown gall tumour cells. *Molecular and General Genetics*, 199(2):344–348.

Meyer, P. and Saedler, H. (1996). Homology-dependent gene silencing in plants. *Annual review of plant physiology and plant molecular biology*, 47:23–48.

Mysore, K. S., Bassuner, B., Deng, X. B., Darbinian, N. S., Motchoulski, A., Ream, W., and Gelvin, S. B. (1998). Role of the *Agrobacterium tumefaciens* VirD2 protein in T-DNA transfer and integration. *Molecular plant-microbe interactions*, 11(7):668–683.

Nealson, K. H., Platt, T., and Hastings, J. W. (1970). Cellular control of the synthesis and activity of the bacterial luminescent system. *Journal of bacteriology*, 104:313–322.

Ochman, H., Gerber, A. S., and Hartl, D. L. (1988). Genetic Applications of an Inverse Polymerase Chain Reaction. *Genetics*, 120(3):621–623.

Oltmanns, H., Frame, B., Lee, L.-Y., Johnson, S., Li, B., Wang, K., and Gelvin, S. B. (2010). Generation of Backbone-Free, Low Transgene Copy Plants by Launching T-DNA from the *Agrobacterium* Chromosome. *Plant physiology*, 152(3):1158–1166.

Ophel, K. and Kerr, A. (1990). *Agrobacterium vitis sp. nov.* for Strains of *Agrobacterium* biovar 3 from Grapevines. *International journal of systematic and evolutionary microbiology*, 40(3):236–241.

Ortega, D., Raynal, M., Laudié, M., Llauro, C., Cooke, R., Devic, M., Genestier, S., Picard, G., Abad, P., Contard, P., Sarrobert, C., Nussaume, L., Bechtold, N., Horlow, C., Pelletier, G., and Delseny, M. (2002). Flanking sequence tags in *Arabidopsis thaliana* T-DNA insertion lines: A pilot study. *Comptes Rendus - Biologies*, 325:773–780.

Pacurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, C., and Bellini, C. (2011). *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. *Physiological and Molecular Plant Pathology*, 76(2):76–81.

Paine, J. A., Shipton, C. A., Chaggar, S., Howells, R. M., Kennedy, M. J., Vernon, G., Wright, S. Y., Hinchliffe, E., Adams, J. L., Silverstone, A. L., and Drake, R. (2005). Improving the nutritional value of Golden Rice through increased pro-vitamin A content. *Nature biotechnology*, 23:482–487.

Pansegrau, W. and Lanka, E. (1991). Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. *Nucleic acids research*, 19(12):3455.

Pansegrau, W., Schoumacher, F., Hohn, B., and Lanka, E. (1993). Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proceedings of the National Academy of Sciences of the United States of America*, 90(24):11538–11542.

Pazour, G. J. and Das, A. (1990). Characterization of the VirG binding site of *Agrobacterium tumefaciens*. *Nucleic acids research*, 18:6909–6913.

Peng, W. T., Lee, Y. W., and Nester, E. W. (1998). The phenolic recognition profiles of the *Agrobacterium tumefaciens* VirA protein are broadened by a high level of the sugar binding protein ChvE. *Journal of bacteriology*, 180:5632–5638.

Peralta, E. G., Hellmiss, R., and Ream, W. (1986). Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. *The EMBO journal*, 5:1137–1142.

Petersen, P. D., Lau, J., Ebert, B., Yang, F., Verhertbruggen, Y., Kim, J. S., Varanasi, P., Suttangkakul, A., Auer, M., Loqué, D., and Scheller, H. V. (2012). Engineering of plants with improved properties as biofuels feedstocks by vessel-specific complementation of xylan biosynthesis mutants. *Biotechnology for biofuels*, 5(1):84.

Petit, A., Delhaye, S., Tempe, J., and Morel, G. (1970). Recherches sur les guanidines des tissus de crown gall. mise en evidence d'une relation biochimique specifique entre le souches d'*Agrobacterium tumefaciens* et les tumeurs qu'elles induisent. *Physiologie végétale*, 8:205 – 213.

Pitzschke, A. (2013). *Agrobacterium* infection and plant defense-transformation success hangs by a thread. *Frontiers in plant science*, 4:519.

Planamente, S., Vigouroux, A., Mondy, S., Nicaise, M., Faure, D., and Moréra, S. (2010). A conserved mechanism of GABA binding and antagonism is revealed by structure-function analysis of the periplasmic binding protein Atu2422 in *Agrobacterium tumefaciens*. *The Journal of biological chemistry*, 285:30294–30303.

Platt, T. G., Bever, J. D., and Fuqua, C. (2012). A cooperative virulence plasmid imposes a high fitness cost under conditions that induce pathogenesis. *Proceedings. Biological sciences / The Royal Society*, 279:1691–1699.

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 research*, 15:557–571.

Polard, P. and Chandler, M. (1995). An in vivo transposase-catalyzed single-stranded DNA circularization reaction. *Genes & development*, 9:2846–2858.

Rabel, C., Grahn, A. M., Lurz, R., and Lanka, E. (2003). The VirB4 family of proposed traffic nucleoside triphosphatases: common motifs in plasmid RP4 TrbE are essential for conjugation and phage adsorption. *Journal of bacteriology*, 185:1045–1058.

Ramanathan, V. and Veluthambi, K. (1995). Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA. *Plant molecular biology*, 28(6):1149–1154.

Rashkova, S., Spudich, G. M., and Christie, P. J. (1997). Characterization of membrane and protein interaction determinants of the *Agrobacterium tumefaciens* VirB11 ATPase. *Journal of bacteriology*, 179:583–591.

Rawat, P., Kumar, S., Pental, D., and Burma, P. (2009). Inactivation of a transgene due to transposition of insertion sequence (IS*136*) of *Agrobacterium tumefaciens*. *Journal of Biosciences*, 34(2):199–202.

Regensburg-Tuïnk, A. J. and Hooykaas, P. J. (1993). Transgenic *N. glauca* plants expressing bacterial virulence gene virF are converted into hosts for nopaline strains of *A. tumefaciens*. *Nature*, 363:69–71.

Rodgers, P. B., Hamilton, W. D., and Adair, J. R. (1999). The therapeutic potential of plant-derived vaccines and antibodies. *Expert opinion on investigational drugs*, 8:211–227.

Rommens, C. M., Bougri, O., Yan, H., Humara, J. M., Owen, J., Swords, K., and Ye, J. (2005). Plant-derived transfer DNAs. *Plant physiology*, 139(3):1338–1349.

Rosso, M. G., Li, Y., Strizhov, N., Reiss, B., Dekker, K., and Weisshaar, B. (2003). An *Arabidopsis thaliana* T-DNA mutagenized population (GABI-Kat) for flanking sequence tag-based reverse genetics. *Plant molecular biology*, 53(1-2):247–259.

Rudder, S., Doohan, F., Creevey, C. J., Wendt, T., and Mullins, E. (2014). Genome sequence of *Ensifer adhaerens OV14* provides insights into its ability as a novel vector for the genetic transformation of plant genomes. *BMC genomics*, 15:268.

Salman, H., Abu-Arish, A., Oliel, S., Loyter, A., Klafter, J., Granek, R., and Elbaum, M. (2005). Nuclear localization signal peptides induce molecular delivery along microtubules. *Biophysical journal*, 89:2134–2145.

Salomon, S. and Puchta, H. (1998). Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. *The EMBO journal*, 17:6086–6095.

Samson, F., Brunaud, V., Balzergue, S., Dubreucq, B., Lepiniec, L., Pelletier, G., Caboche, M., and Lecharny, A. (2002). FLAGdb/FST: a database of mapped flanking insertion sites (FSTs) of *Arabidopsis thaliana* T-DNA transformants. *Nucleic acids research*, 30(1):94–97.

Scheeren-Groot, E. P., Rodenburg, K. W., den Dulk-Ras, A., Turk, S. C., and Hooykaas, P. J. (1994). Mutational analysis of the transcriptional activator VirG of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 176:6418–6426.

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. *The Journal of biological chemistry*, 270:1269–1276.

Schneeberger, R. G., Zhang, K., Tatarinova, T., Troukhan, M., Kwok, S. F., Drais, J., Klinger, K., Orejudos, F., Macy, K., Bhakta, A., Burns, J., Subramanian, G., Donson, J., Flavell, R., and Feldmann, K. A. (2005). *Agrobacterium* T-DNA integration in *Arabidopsis* is correlated with DNA sequence compositions that occur frequently in gene promoter regions. *Functional & integrative genomics*, 5:240–253.

Scordilis, G. E., Ree, H., and Lessie, T. G. (1987). Identification of transposable elements which activate gene expression in *Pseudomonas cepacia*. *Journal of bacteriology*, 169:8–13.

Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M., and Goff, S. A. (2002). A high-throughput *Arabidopsis* reverse genetics system. *The Plant cell*, 14:2985–2994.

Shaw, C. H. (1991). Swimming against the tide: chemotaxis in *Agrobacterium*. *BioEssays : news and reviews in molecular, cellular and developmental biology*, 13:25–29.

Shimoda, N., Toyoda-Yamamoto, A., Aoki, S., and Machida, Y. (1993). Genetic evidence for an interaction between the VirA sensor protein and the ChvE sugar-binding protein of *Agrobacterium*. *The Journal of biological chemistry*, 268:26552–26558.

Shri, M., Rai, A., Verma, P. K., Misra, P., Dubey, S., Kumar, S., Verma, S., Gautam, N., Tripathi, R. D., Trivedi, P. K., and Chakrabarty, D. (2013). An improved Agrobacterium-mediated transformation of recalcitrant indica rice (*Oryza sativa L.*) cultivars. *Protoplasma*, 250(2):631–636.

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. *Proceedings of the National Academy of Sciences of the United States of America*, 89:11837–11841.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. *Nucleic acids research*, 34(Database issue):D32–D36.

Simone, M., McCullen, C. A., Stahl, L. E., and Binns, A. N. (2001). The carboxy-terminus of VirE2 from *Agrobacterium tumefaciens* is required for its transport to host cells by the *virB*-encoded type IV transport system. *Molecular microbiology*, 41:1283–1293.

Smith, E. F. and Townsend, C. O. (1907). A plant-tumor of bacterial origin. *Science (New York, N.Y.)*, 25(643):671–673.

Solenberg, P. J. and Burgett, S. G. (1989). Method for selection of transposable DNA and characterization of a new insertion sequence, IS*493*, from *Streptomyces lividans*. *Journal of bacteriology*, 171:4807–4813.

Springer, P. S. (2000). Gene Traps: Tools for Plant Development and Genomics. *The Plant cell*, 12(7):1007–1020.

Spudich, G. M., Fernandez, D., Zhou, X. R., and Christie, P. J. (1996). Intermolecular disulfide bonds stabilize VirB7 homodimers and VirB7/VirB9 heterodimers during biogenesis of the *Agrobacterium tumefaciens* T-complex transport apparatus. *Proceedings of the National Academy of Sciences of the United States of America*, 93:7512–7517.

Stachel, S. E., Messens, E., Van Montagu, M., and Zambryski, P. (1985). Identification of the signal molecules produced by wounded plant cells that activate T-DNA transfer in *Agrobacterium tumefaciens*. *Nature*, 318(6047):624–629.

Stachel, S. E., Nester, E. W., and Zambryski, P. C. (1986). A plant cell factor induces *Agrobacterium tumefaciens* vir gene expression. *Proceedings of the National Academy of Sciences of the United States of America*, 83:379–383.

Stanford, W. L., Cohn, J. B., and Cordes, S. P. (2001). Gene-trap mutagenesis: past, present and beyond. *Nature reviews. Genetics*, 2(10):756–768.

Starr, M. and Weiss, J. (1943). Growth of phytopathogenic bacteria in a synthetic asparagin medium. *Phytopathology*, 33:314–318.

Stirnberg, A. (2011). Characterization of IS*426*, an active insertion sequence of *Agrobacterium tumefaciens*. *Master's thesis, University of Bonn*.

Strizhov, N., Li, Y., Rosso, M. G., Viehoever, P., Dekker, K. A., and Weisshaar, B. (2003). High-throughput generation of sequence indexes from T-DNA mutagenized *Arabidopsis thaliana* lines. *Biotechniques*, 35(6):1164–1168.

Szabados, L., Kovács, I., Oberschall, A., Abrahám, E., Kerekes, I., Zsigmond, L., Nagy, R., Alvarado, M., Krasovskaja, I., Gál, M., Berente, A., Rédei, G. P., Haim, A. B., and Koncz, C. (2002). Distribution of 1000 sequenced T-DNA tags in the *Arabidopsis* genome. *The Plant journal : for cell and molecular biology*, 32:233–242.

Tang, W., Newton, R. J., and Weidner, D. A. (2007). Genetic transformation and gene silencing mediated by multiple copies of a transgene in eastern white pine. *Journal of experimental botany*, 58(3):545–554.

Tao, Y., Rao, P. K., Bhattacharjee, S., and Gelvin, S. B. (2004). Expression of plant protein phosphatase 2C interferes with nuclear import of the *Agrobacterium* T-complex protein VirD2. *Proceedings of the National Academy of Sciences of the United States of America*, 101:5164–5169.

Thomas, D. R., Penney, C. A., Majumder, A., and Walmsley, A. M. (2011). Evolution of plant-made pharmaceuticals. *International journal of molecular sciences*, 12:3220–3236.

Thomashow, M. F., Panagopoulos, C. G., Gordon, M. P., and Nester, E. W. (1980). Host range of *Agrobacterium tumefaciens* is determined by the Ti plasmid. *Nature*, 283(5749):794–796.

Tingay, S., McElroy, D., Kalla, R., Fieg, S., Wang, M., Thornton, S., and Brettell, R. (1997). *Agrobacterium tumefaciens*-mediated barley transformation. *The Plant journal : for cell and molecular biology*, 11(6):1369–1376.

Tinland, B. and Hohn, B. (1995). Recombination between prokaryotic and eukaryotic DNA: integration of *Agrobacterium tumefaciens* T-DNA into the plant genome. *Genetic engineering*, 17:209–229.

Ton-Hoang, B., Bétermier, M., Polard, P., and Chandler, M. (1997). Assembly of a strong promoter following IS*911* circularization and the role of circles in transposition. *The EMBO journal*, 16:3357–3371.

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. *Journal of bacteriology*, 171:6845–6849.

Toro, N., Datta, A., Yanofsky, M., and Nester, E. (1988). Role of the overdrive sequence in T-DNA border cleavage in *Agrobacterium*. *Proceedings of the National Academy of Sciences of the United States of America*, 85:8558–8562.

Tzfira, T., Frankman, L. R., Vaidya, M., and Citovsky, V. (2003). Site-specific integration of *Agrobacterium tumefaciens* T-DNA via double-stranded intermediates. *Plant physiology*, 133:1011–1023.

Tzfira, T., Tian, G.-W., Lacroix, B., Vyas, S., Li, J., Leitner-Dagan, Y., Krichevsky, A., Taylor, T., Vainstein, A., and Citovsky, V. (2005). pSAT vectors: a modular series of plasmids for autofluorescent protein tagging and expression of multiple genes in plants. *Plant molecular biology*, 57:503–516.

Tzfira, T., Vaidya, M., and Citovsky, V. (2001). VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *The EMBO journal*, 20:3596–3607.

Tzfira, T., Vaidya, M., and Citovsky, V. (2002). Increasing plant susceptibility to *Agrobacterium* infection by overexpression of the *Arabidopsis* nuclear protein VIP1. *Proceedings of the National Academy of Sciences of the United States of America*, 99:10435–10440.

Tzfira, T., Vaidya, M., and Citovsky, V. (2004). Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. *Nature*, 431:87–92.

Ülker, B., Li, Y., Rosso, M. G., Logemann, E., Somssich, I. E., and Weisshaar, B. (2008a). T-DNA-mediated transfer of *Agrobacterium tumefaciens* chromosomal DNA into plants. *Nature biotechnology*, 26(9):1015–1017.

Ülker, B., Peiter, E., Dixon, D. P., Moffat, C., Capper, R., Bouché, N., Edwards, R., Sanders, D., Knight, H., and Knight, M. R. (2008b). Getting the most out of publicly available T-DNA insertion lines. *The Plant journal : for cell and molecular biology*, 56(4):665–77.

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

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. *Molecular plant-microbe interactions*, 22:1356–1365.

van Larebeke, N., Engler, G., Holsters, M., van Den Elsacker, S., Zaenen, I., Schilperoort, R. A., and Schell, J. (1974). Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature*, 252(5479):169–170.

Vanderleyden, J., Desair, J., Meirsman, D., Michiels, K., Van Gool, A. P., Chilton, D., and Jen, G. C. (1986). Nucleotide sequence of an insertion sequence (IS) element identified in the T-DNA region of a spontaneous variant of the Ti-plasmid pTiT37. *Nucleic acids research*, 14(16):6699–6709.

Vergunst, A. C., van Lier, M. C. M., den Dulk-Ras, A., Stüve, T. A. G., 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*. *Proceedings of the National Academy of Sciences of the United States of America*, 102:832–837.

Vogel, A. M. and Das, A. (1992). Mutational analysis of *Agrobacterium tumefaciens virD2*: tyrosine 29 is essential for endonuclease activity. *Journal of bacteriology*, 174:303–308.

Wang, K., Herrera-Estrella, L., Van Montagu, M., and Zambryski, P. (1984). Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell*, 38:455–462.

Ward, E. R. and Barnes, W. M. (1988). VirD2 Protein of *Agrobacterium tumefaciens* Very Tightly Linked to the 5' End of T-Strand DNA. *Science (New York, N.Y.)*, 242(4880):927–930.

Waters, V. L., Hirata, K. H., Pansegrau, W., Lanka, E., and Guiney, D. G. (1991). Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of *Agrobacterium* Ti plasmids. *Proceedings of the National Academy of Sciences of the United States of America*, 88(4):1456–1460.

Watson, B., Currier, T. C., Gordon, M. P., Chilton, M. D., and Nester, E. W. (1975). Plasmid required for virulence of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 123:255–264.

White, P. R. and Braun, A. C. (1942). A Cancerous Neoplasm of Plants. Autonomous Bacteria-Free Crown-Gall Tissue. *Cancer Research*, 2(9):597–617.

Winans, S. C. (1991). An *Agrobacterium* two-component regulatory system for the detection of chemicals released from plant wounds. *Molecular microbiology*, 5:2345–2350.

Wirawan, I. G., Kang, H. W., and Kojima, M. (1993). Isolation and characterization of a new chromosomal virulence gene of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 175(10):3208–3212.

Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Zhou, Y., Chen, L., Wood, G. E., Almeida, N. F., Woo, L., Chen, Y., Paulsen, I. T., Eisen, J. A., Karp, P. D., Bovee, D., Chapman, P., Clendenning, J., Deatherage, G., Gillet, W., Grant, C., Kutyavin, T., Levy, R., Li, M. J., McClelland, E., Palmieri, A., Raymond, C., Rouse, G., Saenphimmachak, C., Wu, Z., Romero, P., Gordon, D., Zhang, S., Yoo, H., Tao, Y., Biddle, P., Jung, M., Krespan, W., Perry, M., Gordon-Kamm, B., Liao, L., Kim, S., Hendrick, C., Zhao, Z. Y., Dolan, M., Chumley, F., Tingey, S. V., Tomb, J. F., Gordon, M. P., Olson, M. V., and Nester, E. W. (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science (New York, N.Y.)*, 294:2317–2323.

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. *Proceedings of the National Academy of Sciences of the United States of America*, 79:6322–6326.

Yanofsky, M., Lowe, B., Montoya, A., Rubin, R., Krul, W., Gordon, M., and Nester, E. (1985). Molecular and genetic analysis of factors controlling host range in *Agrobacterium tumefaciens*. *Molecular and General Genetics*, 201(2):237–246.

Yanofsky, M. F. and Nester, E. W. (1986). Molecular characterization of a host-range-determining locus from *Agrobacterium tumefaciens*. *Journal of bacteriology*, 168(1):244–250.

Ye, G. N., Stone, D., Pang, S. Z., Creely, W., Gonzalez, K., and Hinchee, M. (1999). *Arabidopsis* ovule is the target for *Agrobacterium* in planta vacuum infiltration transformation. *The Plant journal : for cell and molecular biology*, 19:249–257.

Ye, X., Al-Babili, S., Klöti, A., Zhang, J., Lucca, P., Beyer, P., and Potrykus, I. (2000). Engineering the provitamin A (beta-carotene) biosynthetic pathway into (carotenoid-free) rice endosperm. *Science (New York, N.Y.)*, 287(5451):303–305.

Yuan, Q., Carle, A., Gao, C., Sivanesan, D., Aly, K. A., Höppner, C., Krall, L., Domke, N., and Baron, C. (2005). Identification of the VirB4-VirB8-VirB5-VirB2 pilus assembly sequence of type IV secretion systems. *The Journal of biological chemistry*, 280:26349–26359.

Zahrl, D., Wagner, M., Bischof, K., Bayer, M., Zavecz, B., Beranek, A., Ruckenstuhl, C., Zarfel, G. E., and Koraimann, G. (2005). Peptidoglycan degradation by specialized lytic transglycosylases associated with type III and type IV secretion systems. *Microbiology*, 151:3455–3467.

Zaltsman, A., Krichevsky, A., Loyter, A., and Citovsky, V. (2010). *Agrobacterium* Induces Expression of a Host F-Box Protein Required for Tumorigenicity. *Cell Host and Microbe*, 7:197–209.

Zambryski, P., Joos, H., Genetello, C., Leemans, J., Montagu, M. V., and Schell, J. (1983). Ti plasmid vector for the introduction of DNA into plant cells without alteration of their normal regeneration capacity. *The EMBO journal*, 2:2143–2150.

Zhang, L., Murphy, P. J., Kerr, A., and Tate, M. E. (1993). *Agrobacterium* conjugation and gene regulation by N-acyl-L-homoserine lactones. *Nature*, 362(6419):446–448.

Zhang, R.-g., Pappas, K. M., Pappas, T., Brace, J. L., Miller, P. C., Oulmassov, T., Molyneaux, J. M., Anderson, J. C., Bashkin, J. K., Winans, S. C., and Joachimiak, A. (2002). Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. *Nature*, 417(6892):971–974.

Zhu, J. and Winans, S. C. (1999). Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. *Proceedings of the National Academy of Sciences of the United States of America*, 96(9):4832–4837.

Zhu, Y., Nam, J., Humara, J. M., Mysore, K. S., Lee, L.-Y., Cao, H., Valentine, L., Li, J., Kaiser, A. D., Kopecky, A. L., Hwang, H.-H., Bhattacharjee, S., Rao, P. K., Tzfira, T., Rajagopal, J., Yi, H., Veena, G., Yadav, B. S., Crane, Y. M., Lin, K., Larcher, Y., Gelvin, M. J. K., Knue, M., Ramos, C., Zhao, X., Davis, S. J., Kim, S.-I., Ranjith-Kumar, C. T., Choi, Y.-J., Hallan, V. K., Chattopadhyay, S., Sui, X., Ziemienowicz, A., Matthysse, A. G., Citovsky, V., Hohn, B., and Gelvin, S. B. (2003). Identification of *Arabidopsis* rat mutants. *Plant physiology*, 132:494–505.

Ziemienowicz, A., Tinland, B., Bryant, J., Gloeckler, V., and Hohn, B. (2000). Plant Enzymes but Not *Agrobacterium* VirD2 Mediate T-DNA Ligation In Vitro. *Molecular and cellular biology*, 20(17):6317–6322.

Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J. D. G., Boller, T., and Felix, G. (2006). Perception of the Bacterial PAMP EF-Tu by the Receptor EFR Restricts *Agrobacterium*-Mediated Transformation. *Cell*, 125:749–760.

Zupan, J., Hackworth, C. A., Aguilar, J., Ward, D., and Zambryski, P. (2007). VirB1* promotes T-pilus formation in the vir-Type IV secretion system of *Agrobacterium tumefaciens*. *Journal of bacteriology*, 189:6551–6563.

Danksagung

Herrn Prof. Dr. Diedrik Menzel und Herrn Prof. Dr. Lukas Schreiber danke ich für die Begutachtung dieser Arbeit.
Herrn Prof. Dr. Walter Witke und Frau Prof. Dr. Claudia Knief danke ich für die Bereitschaft, der Promotionskommission als fachnahe und fachfremde Mitglieder zur Verfügung zu stehen.

I want to thank Dr. Bekir Ülker for accepting me to be a member of his group, for offering the interesting project, for many fruitful discussions and cheerful musical excursus, and for his infectious joy of science.
I would like to thank all past members of the PME-Group for a good working atmosphere and for enduring my (excellent) taste in music.

Ich danke den anderen Arbeitsgruppen des IZMBs für die freundliche Bereitstellung ihrer Gerätschaften.
Besonders bedanken möchte ich mich bei Monika Polsakiewicz, der gesamten AG Menzel und insbesondere Claudia Heym für die Hilfsbereitschaft in allen Situationen des Laboralltags. Der AG Schreiber danke ich zudem für das stets bereitwillige Öffnen der Mikroskopraumtür.

Special thanks to Ahmet, Caro, Désirée, Laura, Hamed, Iulia, Julia, Nora and Pratibha for taking interest in my project and for their help and contributions during their time in our group.

Alex, Carl, Christian, David, Franzi, Julia und Reynel danke ich für die vielen fröhlichen Stunden (auch außerhalb des Labors) und für die zahllosen Auflüge zur Mensa (später Tibeter).

Ein oberspezieller Dank geht an Feli für eine unglaublich schöne Zeit, für Rat und Tat, fürs Aufbauen nach diversen Krisen und einfach generell fürs da Sein.
Meinen Eltern danke ich für ihre Untertützung und die Gewissheit, mich in allen Lebenslagen auf sie verlassen zu können.

Supplement

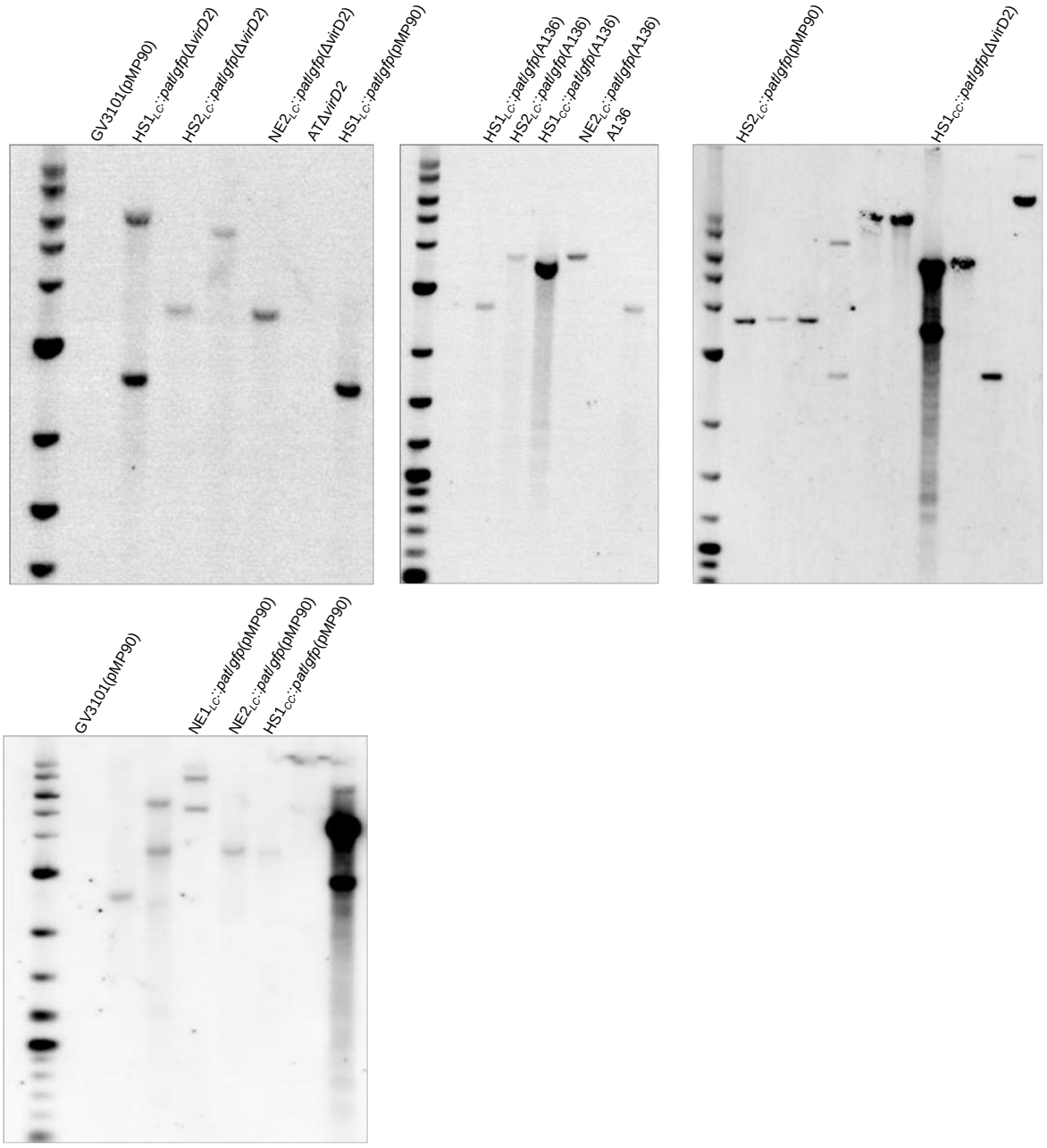


Figure S1: Unedited DNA blots. Bands shown in 4.2.4 are annotated.