Trans-kingdom DNA transfer: Comparing the Ti and RP4 systems

Wittleben, F.T.

Trans-kingdom DNA transfer

Comparing the Ti and RP4 systems

Felix. T. Wittleben

To Doris, Gabi and Thomas

Cover:	Front: 'Bacterial conjugation', graphic by F.T. Wittleben Back: 'Bacterial conjugation', graphic by F.T. Wittleben
Layout:	F.T. Wittleben
Printed by:	Ridderprint BV
ISBN:	978-94-6416-291-2

Trans-kingdom DNA transfer – Comparing the Ti and RP4 systems

Proefschrift

ter verkrijging van
de graad van Doctor aan de Universiteit Leiden,
op gezag van Rector Magnificus prof.mr. C.J.J.M. Stolker,
volgens besluit van het College voor Promoties
te verdedigen op woensdag 9 december 2020
klokke 13.45 uur

door

Felix Thomas Wittleben

geboren te Schwandorf, Duitsland
in 1984

Promotor:	Prof. dr. P.J.J Hooykaas
Co-promotor:	Dr. ir. E.J. van der Zaal

Promotiecommissie:	Prof. dr. G.P. van Wezel (voorzitter)
	Prof. dr. R. Offringa (secretaris)
	Dr. A.C. Vergunst
	Dr. J.A.L. van Kan
	Prof. dr. J. Memelink

Table of contents

Chapter 1

General introduction

Felix T. Wittleben

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Bacterial conjugation, basis for plant transformation by *Agrobacterium*

Bacteria have developed a myriad of different systems to interact with their environment and each other. Prominent amongst these are systems facilitating the exchange of DNA molecules, such as bacterial conjugation, which promote genetic adaption to the environment. This chapter will introduce the central concepts of bacterial conjugation and furthermore its evolution into a system for plant transformation by the bacterium *Agrobacterium tumefaciens*.

Exchange of genetic information between different organisms is known as horizontal gene transfer. It is a driver in the evolution of prokaryotic (and eukaryotic) organisms and occurs in three different ways: natural genetic transformation, viral-mediated transduction and conjugative transfer (Zatyka & Thomas, 1998). Horizontal gene transfer contributes to the emergence of genetic variation in related species, faster adaptation to environmental changes and stimulation of microbial cooperation (Smillie *et al.*, 2010). The advantage of horizontal gene transfer is the evolution of a bacterial genome without reliance on natural mutation; a freely available dynamic gene pool that can be rapidly shared, provides an immense adaptive potential for an organism (Harrison & Brockhurst, 2012).

The information exchange within horizontal gene transfer is driven by mobile genetic elements. There are different types of mobile genetic elements that can occur intra- or extra-chromosomal within the host organism: plasmids, bacteriophages and integrated conjugative elements (ICE) that include conjugative transposons (Frost *et. al.*, 2005). All elements use the host organism's endogenous replication machinery for propagation. Horizontal gene transfer by bacteriophages can occur via direct transfer of chromosomal or extra-chromosomal regions (e.g. plasmids), along its own DNA during its replication cycle (Canchaya *et al.*, 2003). ICE elements are present in the chromosome of a host cell, but are able to excise themselves and transfer between hosts using its own encoded transfer machinery (Johnson & Grossman, 2015). Both, bacteriophages and ICE will not be discussed in this dissertation. Plasmids are extra-chromosomal circular double-stranded DNA units, with some exceptions of linear double-stranded plasmids (Hinnebusch & Tilly, 1993), containing functional genetic modules organized in such a way to be self-replicating using the host endogenous replication machinery. Further, plasmids can contain genes distinctly different from the host genome, encoding accessory traits that can be utilized by the host organism (Frost *et al.*, 2005) and can be found in *Bacteria*, *Archaea*, and *Eukarya* (Phillips & Funnell, 2004). The focus of this dissertation is on bacterial plasmids. Besides extra-chromosomal existence, plasmids could integrate particular genes, large fragments of itself or completely into the host organism's chromosome (Carroll & Wong, 2018). Plasmids can exhibit different types of host ranges in which they can be active, with some plasmids showing a narrow host range with limitation to one or a few species or a broad host-range being able to be maintained in many different species (Jain & Srivastava, 2013). Plasmids, in general, can be grouped in three distinctive classes according to their mobility: conjugative or self-mobilizable/transmissible plasmids, mobilizable/transmissible plasmids that can be transferred only in the presence of conjugative plasmids, and non-mobilizable/transmissible plasmids (Garcillán-Barcia, Alvarado & de la Cruz, 2011). Conjugative and mobilizable plasmids contain a specific region called origin of transfer (*oriT*) allowing the plasmid to be transferred by a conjugative machinery into another host organism. Non-mobilizable plasmids can be transferred via natural transformation or bacteriophage-mediated transduction (Clark & Adelberg, 1962; Smillie *et al.*, 2010). Another form for a non-mobilizable plasmid to be transferred can be conduction, during which the non-mobilizable plasmid co-integrates with a conjugative plasmid. After transfer, the non-mobilizable plasmid is resolved from the co-integrate within the recipient cell (Clark & Warren, 1979). Each plasmid encodes an intricate system allowing for its own maintenance and propagation within the host organism. A partitioning system controls accurate segregation of plasmid copies into the new daughter cells of the original host organism after cell division

(Frost *et al.*, 2005). There is further a strict regulation of the amount of copies present within the cell, enforced by an encoded negative regulation mechanism. Different plasmids can exhibit a high or a low copy number (Del Solar & Espinosa, 2002). Each plasmid has a specific mechanism for replication control, and different plasmids can contain the same replication mechanism. The replication mechanism acts on as specific region within the plasmid called origin of replication (*oriR*/*ori*), or synonymously origin of vegetative replication (*oriV*). If two plasmids contain the same *oriV*, they can in principle, depending on their copy number, not be stably maintained in the same host cell without external selection (Novick, 1987), as both would not be discriminated by the partitioning machinery. However, a more recent analysis demonstrated that co-existence (of two plasmids with the same *oriV*) could last much longer than previously expected (Velappan *et al.*, 2007). Nevertheless, plasmids can be classified by their incompatibility of being maintained in the same cell at the same time. This trait of incompatibility (Inc) was and is still widely used to distinguish and classify different plasmids (Couturier *et al.*, 1988), and is defined for plasmids of *Enterobacteriaceae*, *Pseudonomdas, Agrobacterium*, and Gram-positive *Staphylococci* (Frost *et al.*, 2005). With the advent of rapid sequencing methods plasmids can now be grouped based on molecular differences in replication-, vegetative maintenance- or conjugative systems (Garcillán-Barcia, Alvarado & de la Cruz, 2011; Shintani, Sanchez & Kimbara, 2015; Orlek *et al.*, 2017). In this dissertation, the following plasmids are discussed and employed: the IncP conjugative broad host range plasmid RP4 (Datta & Hedges 1971), the IncW conjugative broad host range plasmid R388 (Fernández-López *et al.*, 2006), the IncQ mobilizable broad host range plasmid RSF1010 (Rawlings & Tietze, 2001) and IncRh1 Ti (tumor inducing) plasmids of *A. tumefaciens* (Hooykaas *et al.*, 1980).

A key element of conjugative DNA transfer systems is the type-IV secretion system (T4SS) machinery (Guglielmini, de la Cruz & Rocha, 2013). In this dissertation, the focus will be on plasmid-encoded conjugative systems. The T4SS may have evolved from a protein secretion machinery into an apparatus equipped (also) for DNA transfer. Proteins translocated may have a variety of functions in the recipient cell, for instance may be involved in the recircularization of the transferred single-stranded plasmid DNA and in the conversion into a double-stranded form, but may also have functions influencing the growth of the recipient cell (Guglielmini, de la Cruz & Rocha, 2013).

To summarize and provide more detail, the three elements of horizontal gene transfer occurring between organisms are:

1. Natural transformation: where freely available genetic information in the form of DNA molecules can be taken up by naturally occurring competent bacterial cells. Natural transformation is a phenomenon observed in many bacterial species. The percentage of naturally competent cells inside a population can vary from a minuscule fraction to almost the entire group (Dubnau, 1999; Bakkali, 2013). The natural transformation process differs from Gram-positive to Gram-negative bacteria (Chen & Dubnau, 2004). In Gram-negative bacteria with two cell membranes, the captured DNA is transported from the outer membrane into the periplasmic space and then further into the cell. Captured double-stranded DNA is usually converted into single strand molecules, whereby one strand is taken up and the other complementary strand is degraded. If the new DNA is not a plasmid that can replicate, the new genes are either lost or integrated into the genome by homologous recombination (Chen & Dubnau, 2004).

2. Transduction: the direct introduction of genetic information by a viral vector into the organism. In this manner, bacteriophages may be an essential part of spreading genetic material throughout the bacterial community within their reach. A limiting factor in transduction is the ability of bacteriophages to infect only a particular set of host strains within a species (Smillie *et al.*, 2010).

3. Conjugative transfer: the exchange of genetic information between cells during direct physical contact. Conjugation is the imperative mode of terrestrial-based horizontal gene transfer and is more prevalent than viral transduction (Bushman, 2002; Volkova *et al.*, 2014; Nazarin, Tran & Boedicker, 2018). An extensive network analysis study of 111 bacterial genomes and hundreds of thousands of environmental DNA samples of plasmids and viruses from various databases provided evidence of the importance of plasmid driven exchange (Halary *et al.*, 2010). Horizontal gene transfer also occurs in marine-based environments, but little is known about prevalence between conjugation and transduction (Hermansson & Lindberg, 1994; Sobecky & Hazen, 2009) The exchanged genetic material is almost invariably mediated by conjugative plasmids, extra-chromosomal DNA elements in the possession of conjugative transfer (*tra*) genes as well as genes regulating its own replication. Conjugative elements can provide the host with specialized traits to settle in specific niches that could previously not be colonized (Kado, 1998). The initial cell-to-cell (donor to recipient) contact can be established by the conjugative pilus structure (Willetts & Wilkins, 1984). For a successful conjugational process, cell-to-cell contact is required for unidirectional exchange of genetic material (Achtman, Morelli & Schwuchow 1978). During conjugation, single-stranded DNA is transferred between bacterial cells (Guglielmini, de la Cruz & Rocha, 2013). There is an exception with the Gram-positive bacterium *Streptomyces spp*., as double-stranded DNA has been found to be transferred in a chromosomal segregation type transfer mechanism in this organism (Possoz *et al.*, 2001; Vogelmann *et al.*, 2011).

Conjugation

Lederberg and Tatum (1946) first described the process of bacterial conjugation in the Gram-negative *Escherichia coli* K12. The conjugative F-plasmid of *E. coli* K12 then became the model system for bacterial conjugation. For simplification and categorization, the general process of bacterial conjugation can be broken down into four phases: DNA substrate processing and substrate recruitment to the transfer apparatus in the donor, the actual translocation itself by the T4SS, re-circularization of the transferred single-stranded plasmid DNA and conversion into double-stranded DNA in the recipient (Grohmann, Muth & Espinosa, 2003; Thomas & Nielsen, 2005; Alvarez-Martinez & Christie 2009; Cabezon *et al.*, 2015). The fourth phase will not be discussed in more detailed as the fate of the transferred DNA depends on the type of DNA transferred as well as on the type of recipient cell. These phases are orchestrated by three discernable systems:

1. The DNA transfer and replication system (**Dtr**), mediating the enzymatic processing of DNA substrate to be transferred into the donor. The primary enzyme of the Dtr is the relaxase, which forms a nick (single-strand break) at a specific *nic* site within the *oriT*. Auxiliary proteins assist in this step.

2. The coupling protein (**CP**), linking the processing (Dtr) and transport (Mpf) of the substrate during conjugation.

3. The mating pair formation system (**Mpf**) creating and mediating a physical bridge between two cells, that allows for substrate transfer. Mpf systems are large multi-protein complexes that create a *trans*-membrane-spanning channel (Schröder & Lanka, 2005). The

actual physical cell-to-cell contact that is necessary for substrate transfer is still a poorly understood process. A cell junction is created during this event, and specific signals are triggered to initiate the transfer process (Samuels, Lanka & Davies, 2000; Lang *et al.*, 2011; Bhatty, Laverde Gomez & Christie, 2013). Nowadays the Mpf is often referred to as the T4SS. Although the T4SS is best known from its role in conjugation, it has also evolved and specific T4SS have been found to play a role in transformation (secretion and uptake of DNA molecules from the environment) and also in virulence, by facilitating the transfer of virulence proteins and sometimes also DNA molecule into eukaryotic host cells (plant, animal and human) contributing to disease development (Christie, 2001; Llosa *et al.*, 2002; Cabezon *et al.*, 2015).

Mobilizable plasmids do not encode and Mpf transport system, but may still be transmitted by use of the Mpf encoded by another plasmid. Such mobilizable plasmids encode their own Dtr system that is mostly referred to as Mob system, and may or may not encode their own CP (Szpirer, Faelen & Couturier, 2000; Smillie *et al.*, 2010).

To summarize these three pillars of conjugation, a simplified overview is given in **figure 1**. All components acting together make the transfer of DNA to a host cell possible. In some instances, the T4SS only transfers effector proteins without DNA transfer (Cascales & Christie, 2003).

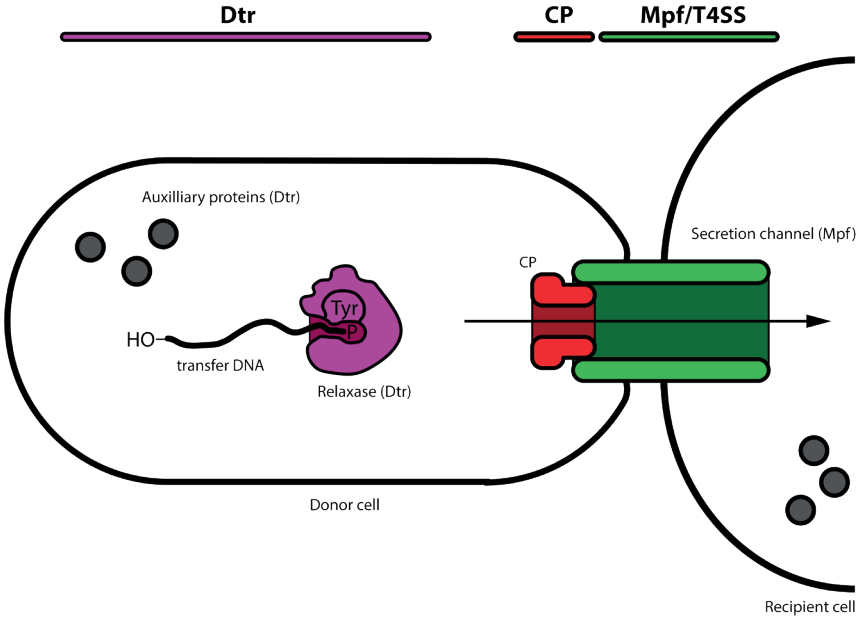


Figure 1. The three elements of bacterial conjugation in a simplified representation. The designated transfer DNA can be either a full-length plasmid or a specific DNA sequence on a plasmid. The DNA transfer and replication system (Dtr) and its auxiliary proteins process the transfer DNA. A relaxosome complex binds to a specific double-stranded DNA sequence and initiates nicking of one of the DNA strands via the relaxase (purple). The relaxase binds during this process covalently to the 5'-end of the transfer DNA, and a single-strand still bound to the relaxase is displaced. The single-strand associates via the translocation signal of the relaxase with the coupling protein (CP; red). The CP thus recognizes the specific substrate and initiates the transfer via the type-IV secretion system (T4SS) or mating pair formation (Mpf; green) channel into the recipient cell. Besides the relaxase, other proteins may also be recognized by the CP and transported via the T4SS into the recipient cell, aiding in further DNA processing, for example in the conversion from the single-stranded into a double-stranded form.

For easier understanding of the nomenclature, the system established by Guglielmini, de la Cruz & Rocha (2013) will be used throughout this dissertation. Proteins and genetic elements will be noted by **GI_{GE}**, referring first to the protein name or gene identification (**GI**), and then the genetic element (**GE**) where it is derived from in subscript. For example, $VirD2_{Ti}$ is the relaxase VirD2 protein encoded in the virulence system of the Ti plasmid of *A. tumefaciens*, or $traI_{RP4}$ the gene encoding the relaxase TraI on the conjugational plasmid RP4. Proteins are noted in roman letters with the first letter in capital. Genetic elements will be given in italicized letters with lower case first letters. For example, the protein $VirD2_{Ti}$ is encoded in the gene $virD2_{Ti}$. Plasmid names are always given in roman letters and are not italicized. For consistency, the abbreviation of the type-IV secretion system will be T4SS.

Agrobacterium tumefaciens - plant pathogen

This dissertation studies gene and protein transfer from *A. tumefaciens*, a Gram-negative bacterium of the rhizosphere (**Figure 2**). It is the causative agent of crown gall disease, a formation of neoplastic growth in a variety of plant species (De Cleene & De Ley, 1976). Specific DNA is transferred into plant cells, which in turn are transformed by the DNA that further leads to tumor growth. This *trans*-kingdom gene transfer is carried out by a set of virulence (Vir_{Ti}) proteins able to transport oncogenes, encoded on transfer-DNA (T-DNA) into host cells. Integration into the host cells genome leads to stable transformation. This process resembles bacterial conjugation (Lessl & Lanka, 1994). The induced neoplastic plant growth produces a vast amount of specific compounds called opines that can be utilized by *A. tumefaciens.* The virulence genes (vir_{Ti}), as well as the opine catabolic genes, are located in the tumor-inducing plasmid (Ti plasmid; **Figure 3A**). The most studied strain is the pathogenic *A. tumefaciens* strain C58, on which many laboratory strains are based. It has a genomic size of around 5 Mbp, split on a circular and linear chromosome. It harbors the cryptic plasmid pAtC58, sized around 550 kbp, and the nopaline-type Ti plasmid pTiC58 with an approximate size of 220 kbp (Goodner *et al.*, 2001).

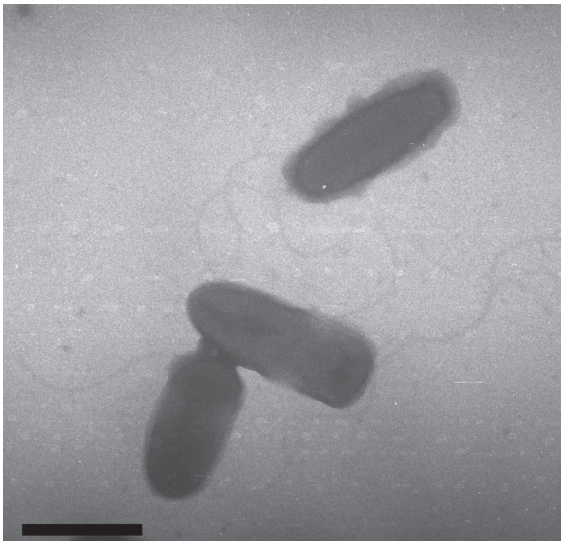


Figure 2. Electron micrograph of cells of *Agrobacterium tumefaciens* strain LBA1100. Bacterial cells were negatively stained with 3% ammonium molybdate. The cell is approximately 1.2 μm in length. The bar = 1 μm. Magnification is 15000×.

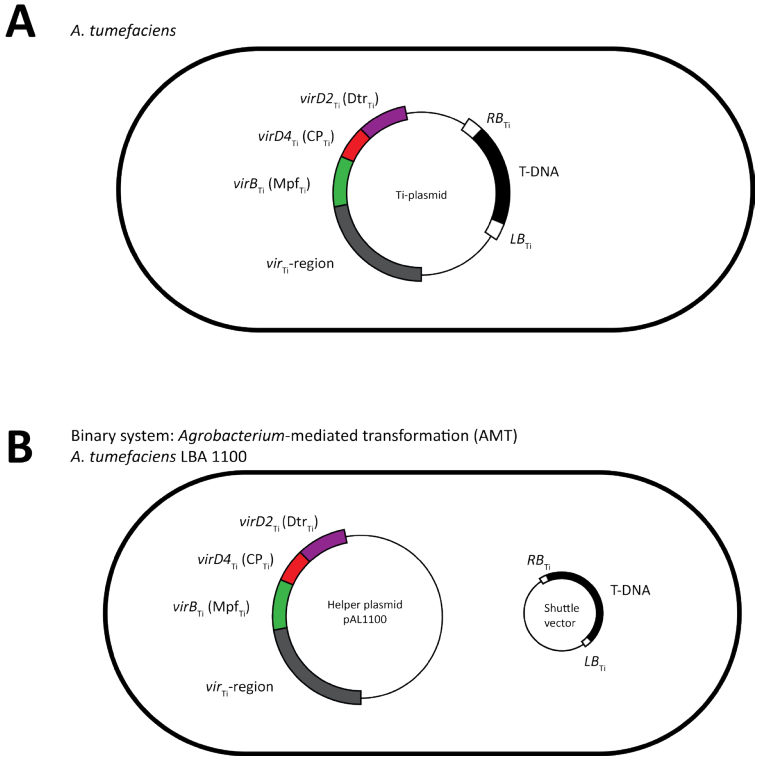


Figure 3. General representation of the *A. tumefaciens*-mediated transformation system. The graphic is a general overview as shown in **figure 1**, including Dtr with relaxase (purple), CP (red), T4SS/Mpf channel (green). **A:** The Ti plasmid in the wild-type configuration. All parts essential for plant transformation are present on the Ti plasmid: the transferred T-DNA and vir_{Ti}-region, which encodes for all virulence proteins involved in transfer. **B:** The binary system, widely used for AMT, separates the vir_{Ti}-region including Dtr, Mpf and CP on the helper plasmid from the T-DNA that is provided on a separate plasmid (shuttle vector). The figure is not a true spatial representation.

The cryptic plasmid pAtC58 is not essential for transfer but has a positive effect on it, is self-transmissible and contains genes for metabolic regulation (Kado, 1998; Nair *et al.*, 2003). The pTiC58 belongs to the incompatibility group IncRh1 (Hooykaas *et al.*, 1980). Within Ti plasmids, several functional regions have been recognized: replication, virulence (vir_{Ti} region), opine uptake and catabolism, self-conjugation (tra_{Ti}/trb_{Ti} region), and the region designated to be transferred to the recipient cells (T-DNA region). Additionally, the Ti plasmid contains the $oriV_{Ti}$ for plasmid maintenance within *Agrobacterium*. Ti plasmids, in general, have a host range limited to the family of *Rhizobiaceae* that includes the genus *Agrobacterium* (Hooykaas & Beijersbergen, 1994). The tra_{Ti}/trb_{Ti} region, which mediates the transfer of the Ti plasmid in-between *Agrobacteria*, shares similarities with the conjugative systems of the RP4, RSF1010 and F plasmids (Farrand, Hwang & Cook, 1996). All these features are likely to reflect the evolutionary history of the Ti plasmid (Alt-Mörbe *et al.*, 1996).

To infect and transform plants, *A. tumefaciens* must become virulent by expressing its virulence proteins. These are encoded in the vir_{Ti} region containing 8-10 operons that are together essential for tumorigenesis (Hooykaas & Beijersbergen, 1994). The release of sugars and phenolic compounds present in the sap of wounded plant cells and stable low pH more acidic level (Brencic & Winans, 2005), caused by the apoplastic fluid release, can be sensed by the two-component system $VirA_{Ti}/VirG_{Ti}$ detecting these physiological signals. Within

laboratory settings, the phenolic compound acetosyringone is used to stimulate virulence of *Agrobacteria*. Phenolic compounds are bound by the *trans*-membrane-spanning histidine kinase $VirA_{Ti}$. The chromosomally encoded virulence protein $ChvE_{Ti}$ binds to released sugar molecules and enhances the induction signal of $VirA_{Ti}$ (Hu *et al.*, 2013; Gao & Lynn, 2005). Both $ChvE_{Ti}$ and $VirA_{Ti}$ can detect an environmental change to acidic conditions, at around a pH of 5.5, to induce virulence (Rogoswky *et al.*, 1987; Melchers *et al.*, 1989; Liu *et al.*, 1993; Turk *et al.*, 1994; Gao & Lynn, 2005; Hu *et al.*, 2013). When $VirA_{Ti}$ is activated, it auto-phosphorylates and in return phosphorylates the cytoplasmic transcriptional activator $VirG_{Ti}$ (Jin *et al.*, 1990). Active $VirG_{Ti}$ binds to vir_{Ti}-box regions on the Ti plasmid and induces vir_{Ti} gene expression (Gao & Lynn, 2005). In the plant wound, *Agrobacterium* attaches to the plant cells (Matthysse, 2014). The induced $VirB_{Ti}$ proteins form a core channel-structure ($T4SS_{Ti}$) spanning through the membrane of *A. tumefaciens* contacting with the host cell, and together with the CP_{Ti} $VirD4_{Ti}$ facilitating the transfer of genetic material and auxiliary/effector proteins into the recipient cell (Cascales & Christie, 2004; Schröder & Lanka, 2005; Aguilar *et al.*, 2011). The T-DNA, the transfer vehicle for tumor-inducing genes (auxin & cytokinin) as well as opine synthesis genes, is flanked by *cis*-acting recognition sequences comparable in function to *oriT* sequences, called right and left border (RB_{Ti} and LB_{Ti} respectively; **Figure 3A**; Bundock *et al.*, 1999). The relaxase $VirD2_{Ti}$ recognizes the border sequences marking the T-DNA region on the T-region present on the Ti plasmid. $VirD2_{Ti}$ nicks one of the DNA strands at the border sequences and attaches to it covalently at its 5'-end. Several auxiliary proteins aid $VirD2_{Ti}$ in binding and nicking the border sequences. The nicking function of $VirD2_{Ti}$ is aided when $VirD2_{Ti}$ is associated with $VirD1_{Ti}$ (Relić *et al.*, 1998), and binding of $VirD2_{Ti}$ to the designated DNA sequences is aided by $VirC1_{Ti}$ and $VirC2_{Ti}$ (Toro *et al.*, 1988). Ensuing DNA synthesis at the 3'-end of the nick site restores the original T-DNA region, and helps to displace the single-stranded T-strand. After displacement, the T-strand bound to $VirD2_{Ti}$ associates with the CP_{Ti} $VirD4_{Ti}$, which then facilitates the transfer of the T-strand complex via the $T4SS_{Ti}$, comprised of $VirB_{Ti}$ proteins, into the host cell (Vergunst *et al.*, 2000). Inside the host cell, the T-strand bound to $VirD2_{Ti}$ is forming a so-called T-complex with $VirE2_{Ti}$. The effector protein $VirE2_{Ti}$, another transferred virulence protein, is able to bind along the transferred single-stranded T-strand inside the host cell (Citovsky, de Vos & Zambryski 1988). Nuclear localization signals are present in $VirD2_{Ti}$ (Rossi, Hohn & Tinland, 1993) and $VirE2_{Ti}$ (Zupan, Citovsky & Zambryski, 1996), that guide the T-complex into the host cell nucleus. In the nucleus, host endogenous DNA processing systems convert the single-stranded T-strand into a double-stranded form (Liang & Tzfira, 2013). The T-DNA can be integrated into the host cell genome (Tzfira *et al.*, 2004), creating stable transformation and tumor formation. Transient expression of the T-DNA, when not integrated, is also possible (Narasimhulu *et al.*, 1996). In biotechnological research, disarmed (no tumorigenic T-DNA) *Agrobacterium* can be used to create transgenic plants (Barton *et al.*, 1983). The discovery that vir_{Ti} genes located *in trans* on a separate plasmid-vector still allowed for T-strand transfer (Hoekema *et al.*, 1983), prompted the development of the binary vector system (**Figure 3B**). The binary vector system is now routinely used for *Agrobacterium*-mediated transformation (AMT) of plants. AMT is also possible, within laboratory conditions, to transform yeast and fungi. (Bundock *et al.*, 1995; de Groot *et al.*, 1998). That the *Agrobacterium* virulence system is evolutionary related to bacterial conjugation systems became apparent, when it was shown that it could still mediated conjugative plasmid transfer from *Agrobacterium* to other *Agrobacterium* and *E. coli* (Beijersbergen *et al.*, 1992).

The broad host range conjugative plasmid RP4

The RP4 plasmid was initially identified to be an agent of antibiotic drug resistance spread between bacteria (Saunders & Grinsted, 1972). It was found in *Pseudomonas aeruginosa*

isolated from a clinical sample in Birmingham (UK) at the Birmingham Accident Hospital (Ingram, Richmond & Skyes, 1973). Other conjugative plasmids R18, R68, RK2 and RP1 isolated in the same hospital in the same period were identified to be identical to RP4, and are also referred to as Birmingham IncPα plasmid (Bryan, van den Elzen & Tseng, 1972; Saunders & Grinsted, 1972; Pansegrau *et al.*, 1994). The plasmid RP4 is part of the IncPα incompatibility group of conjugative plasmids together with plasmids such as R702 (Hedges & Jacob, 1974). RP4 is a broad range host plasmid, able to be maintained for instance in *E. coli* and *A. tumefaciens,* conferring antibiotic resistance to its hosts (Datta & Hedges, 1971; Levin *et al.*, 1976). It has a size of 60 kbp and contains around 74 genes, of which only 60 have been functionally identified. It is a self-transmissible plasmid with conjugative transfer encoded within the *tra1*$_{RP4}$ region (Dtr$_{RP4}$, CP$_{RP4}$) and the *tra2*$_{RP4}$ region (Mpf). The Mpf$_{RP4}$ system used by RP4 is a T4SS$_{RP4}$. The *oriT*$_{RP4}$ is located in the *tra1*$_{RP4}$ region. The *oriV*$_{RP4}$ and replication genes, regulating plasmid maintenance, are situated around 6000 bp upstream of the *tra1*$_{RP4}$ region (Pansegrau *et al.*, 1994). After its discovery and realization of its strong promiscuity, the conjugative system of RP4 was adapted for use in genetic engineering (**Figure 4**; Jacob & Grinter, 1975; Ditta *et al.*, 1980; Paget *et al.*, 1990; Babic, Guérout & Mazel, 2008).

The conjugative system encoded by plasmid RP4 and the core of the virulence system of the Ti plasmid share a close structural and functional homology, and are likely derived from one ancestral conjugative system (Lessl, Pansegrau & Lanka, 1992; Pansegrau *et al.*, 1994; Lessl & Lanka, 1994; Alt-Mörbe *et al.*, 1996; Hamilton *et al.*, 2000). Each conjugational system is specialized for each host organism or its function. The Ti plasmid, using a T4SS$_{Ti}$ can transport and integrate DNA into plant cells as well as transfer auxiliary proteins aiding in this process. The RP4 plasmid, on the other hand, is highly effective in moving itself as well as other transmissible plasmids to a multitude of Gram-positive and -negative species. The function of the different elements of both conjugative systems will be discussed in detail and compared to each other in context below.

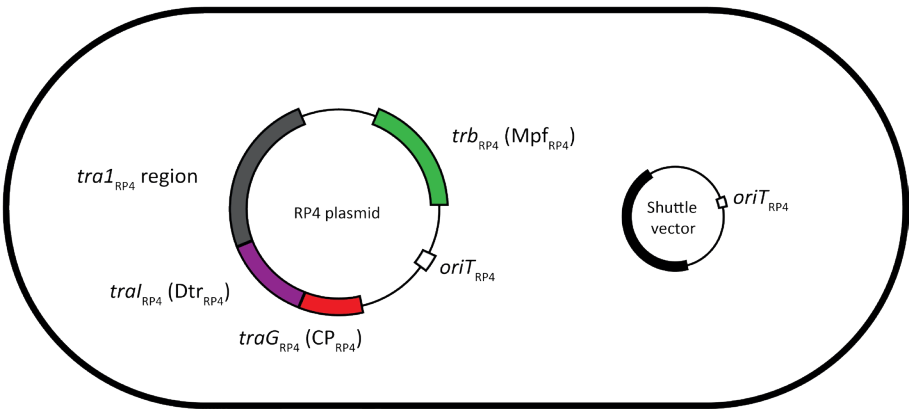


Figure 4. General representation of the RP4 plasmid-based transfer system. This is a general overview as shown in **figure 1**, including Dtr with relaxase (purple), CP (red), the T4SS/Mpf channel (green). The plasmid itself, and a shuttle vector to be transferred, both contain an *oriT*$_{RP4}$ that is recognized by TraI$_{RP4}$. The Dtr, CP and Mpf machinery will transfer both RP4 and the shuttle plasmid into a recipient cell. The plasmid RP4 can be maintained in *A. tumefaciens* as well as in *E. coli*. The figure is not a true spatial representation.

DNA transfer and recognition system (Dtr)

Regardless of the further details, the transferred DNA must contain at least one recognition site for the Dtr system. This *cis*-acting sequence, the *oriT*, is essential for conjugative transfer and contains the site where the relaxase forms a nick (Fuqua & Winans, 1996). The relaxase requires the assistance of auxiliary proteins able to open the double-stranded DNA sequence at the *oriT* recognition site to provide access of the relaxase to the nicking site.

The Dtr_{RP4} system of the RP4 plasmid consists of the relaxase $TraI_{RP4}$ and its auxiliary proteins $TraH_{RP4}$, $TraJ_{RP4}$ and $TraK_{RP4}$. Plasmids that can be mobilized by the conjugative RP4 plasmid contain a specific $oriT_{RP4}$ (Waters *et al.*, 1991). The $oriT_{RP4}$ is recognized by the relaxase $TraI_{RP4}$ and by the auxiliary proteins $TraH_{RP4,}$ $TraJ_{RP4}$ and $TraK_{RP4}$, collectively forming the relaxosome complex. All these proteins are encoded in the Dtr_{RP4} operon in the $tra1_{RP4}$ region of RP4 (Ziegelin *et al.*, 1991). The $oriT_{RP4}$ contains 19 bp inverted repeat sequences that are recognized by $TraJ_{RP4}$, recruiting the relaxase $TraI_{RP4}$ to bind to its nicking site (Ziegelin, Fürste & Lanka, 1989). The binding site for $TraI_{RP4}$ is called *sri*, for $TraJ_{RP4}$ *srj*, and for $TraK_{RP4}$ *srk* (Pansegrau & Lanka, 1996). $TraH_{RP4}$ can stabilize the binding of $TraJ_{RP4}$ and $TraI_{RP4}$ as a kind of scaffold protein and does not bind DNA by itself (Pansegrau & Lanka, 1996). $TraK_{RP4}$ is thought to unwind superhelical DNA for easier access of $TraI_{RP4}$ to the $oriT_{RP4}$. The DNA strand is wrapped around a multimeric complex of $TraK_{RP4}$ that can recognize *srk*, a site 180 bp downstream of the $oriT_{RP4}$-nic_{RP4} site (Ziegelin *et al.*, 1992). $TraK_{RP4}$ is essential for the conjugative process *in vivo* (Fürste *et al.*, 1989). The relaxases are the key enzymes in DNA processing. Once recruited to the transfer DNA recognition site, they cut the nic_{RP4} site in the recognition sequence, creating a single-strand DNA break in one of the strands (Vogel & Das, 1992; Pansegrau *et al.*, 1990; Pansegrau *et al.*, 1993). The DNA backbone is cleaved by a nucleophilic attack on the phosphate backbone of the DNA strand, started by the hydroxyl group of the tyrosine residue within the active center of the relaxase. The cleavage leads to a nick and the covalent attachment of the 5'-end to the relaxase. As a reverse reaction, the relaxase can again ligate the 5'- and 3'-end of the DNA strand of the $oriT_{RP4}$ without additional energy. This so-called transesterification process could play an important biological role in single-stranded DNA re-circularization in the recipient host cell (Becker & Meyer, 2012). The essential tyrosine of $TraI_{RP4}$ is the Tyr22 residue (Pansegrau *et al.*, 1993).

The Ti plasmid contains two distinct Dtr_{Ti} systems, one for conjugative plasmid transfer of the Ti plasmid, encoded in the tra_{Ti} and trb_{Ti} regions, and one for T-DNA transfer encoded in the vir_{Ti} regions (Christie & Gordon, 2014). This study focused on the Dtr_{Ti} system involved in T-DNA transfer. This Dtr_{Ti} system consists of the relaxase $VirD2_{Ti}$, and its auxiliary proteins $VirD1_{Ti}$, $VirC1_{Ti}$ and $VirC2_{Ti}$. They are encoded in the $virD_{Ti}$ and $virC_{Ti}$ operon. The T-DNA is defined by two 25 bp imperfect direct repeats, the right-border (RB_{Ti}) and the left-border (LB_{Ti}) sequences. The border sequences contain the nicking sites nic_{Ti} for the relaxase $VirD2_{Ti}$. Within the extensively studied octopine Ti plasmids, the octopine RB_{Ti} is flanked by an upstream enhancer domain. The enhancer is also called overdrive (OD_{Ti}) aiding to increase the nicking by the relaxase (Toro *et al.*, 1988). The OD_{Ti} domain is recognized by $VirC1_{Ti}$ (Toro *et al.*, 1989) and together with $VirC2_{Ti}$ they facilitate the nicking reaction of $VirD2_{Ti}$ and $VirD1_{Ti}$ (Lu *et al.*, 2009). $VirC2_{Ti}$ contains a conserved ribbon-helix-helix (RHH) DNA binding domain to initiate the contact with the recognition site (Lu *et al.*, 2009). The $VirD2_{Ti}$ nicking of the border sequences is facilitated by the essential tyrosine 29 (Tyr29) in its active center (Vogel & Das, 1992). After DNA nicking, $VirC1_{Ti}$ seems involved in helping the T-strand, coupled to $VirD2_{Ti}$, to be recruited to cellular membrane and associate to the coupling protein $VirD4_{Ti}$ for the subsequent transfer (Atmakuri *et al.*, 2007).

The *nic* sites in RB_{Ti}, $LB_{Ti,}$ as well as in the $oriT_{RP4}$ show similarity and share the common relaxase binding site YATCCTG*Y, with the nicking site marked with an asterisk (Pansegrau & Lanka, 1991). In addition, the relaxases $TraI_{RP4}$ and $VirD2_{Ti}$ are closely related, and belong to the same P-type group of relaxase-*nic* systems (Pansegrau & Lanka, 1991), which is clearly distinct from the Dtr systems of other plasmids such as IncF, IncW and IncQ plasmids (Lanka & Wilkins, 1995).

Coupling protein (CP)

The translocation of the nucleo-protein complex via the T4SS is determined by the CP (**Figure 1**). The CP particulates the specificity of the conjugational system in what precisely is recognized and transported. Although CP are formally not part of the multi-protein complex of the transport channel, they are sometimes described as if being a T4SS protein. The CP is not required for the assembly of the T4SS and the pilus structure (Hamilton *et al.*, 2000). CPs range typically in size of around 600-700 amino acid residues (AA) and possess a minimum of two predicted N-terminal *trans*-membrane domains, with an intervening periplasmic loop of 30-50 residues. The CP contains three domains to ensure functionality: an N-terminal part for association with the membrane, a nucleotide-binding domain, and a C-terminal all-alpha domain (Alvarez-Martinez & Christie, 2009). All CPs have a nucleotide-binding domain with a conserved Walker motif, involved in ATP hydrolysis. The ATPase function could drive DNA transport in conjunction with two other ATPases $VirB4_{Ti}$ and $VirB11_{Ti}$ associated with the $T4SS_{Ti}$ (Llosa *et al.*, 2002; Alvarez-Martinez & Christie, 2009; Cascales *et al.*, 2013). The CP is the first factor contacted by the T-strand/$VirD2_{Ti}$ complex (Cascales & Christie, 2004). The mode of action of CP is not known, but it is thought that that the DNA/protein complex is moved through the middle of a hexameric ring structure comprised of multiple $VirD4_{Ti}$ proteins present at the inner membrane structure of the $T4SS_{Ti}$ channel (Alvarez-Martinez & Christie, 2009). This so-called 'shoot and pump' model for CP mediated substrate transfer via the T4SS seems the most likely pathway (Llosa *et al.*, 2004; Atmakuri, Cascales & Christie, 2004; Cascales *et al.*, 2013). The 'shoot and pump' model refers to the act of the T4SS initially 'shooting' the relaxase bound to the transfer DNA across several membranes and then 'pumping' the rest of the DNA via the CP into the recipient cell (Llosa *et al.*, 2004). The pathway for the transfer of single proteins is still unclear. Although there is substrate selection of the transport system, a yet unknown cell-to-cell contact derived activation signal via the CP is required to mediate the direct transfer of the DNA into the recipient cell (Cascales *et al.*, 2013).

The CP of the RP4 conjugative system is $TraG_{RP4}$, encoded in the $tra1_{RP4}$ region (Ziegelin *et al.*, 1991; Pansegrau *et al.*, 1994), while the CP of the Ti plasmid virulence system is $VirD4_{Ti}$, encoded in the $virD_{Ti}$ operon (Porter *et al.*, 1987).

In certain cases, CP of different conjugative systems can be exchanged and continue their function within other systems. The CP_{RP4} $TraG_{RP4}$ and CP_{R388} $TrwB_{R388}$ can substitute each other's function, but with reduced the efficiency in between plasmids IncW R388 and IncPα RP4. In chimeric systems of both plasmids, the RSF1010 plasmid can be translocated (Cabezon *et al.*, 1997). Within *Agrobacterium*, transfer of the plasmid RSF1010 is dependent on the CP_{Ti} $VirD4_{Ti}$for transfer as it lacks its own CP (Beijersbergen *et al.*, 1992), but transfer of CloDF13 is possible in absence of $VirD4_{Ti}$ as CloDF13 encodes it's on $CP_{CloDF13}$ (Escudero *et al.*, 2003).

Mating pair formation (Mpf) and architecture of the T4SS

After the DNA substrate is successfully processed and associated with the CP, via the covalently bound relaxase, the final step is the transfer from the donor to the recipient cell. The transfer is ensured by the T4SS, a transmembrane spanning protein channel. The structure is essential for the successful transport of the nucleoprotein-complex. Along the transmembrane spanning T4SS an extracellular pilus structure is present, visible by microscopic observation. This pilus structure may be involved in making the initial cell-to-cell contact, but is not necessary for actual DNA transfer.

The T4SS is one of the most abundant secretion systems present in bacteria. The core elements of the T4SS are highly conserved (Segal *et al.*, 1999). The most studied T4SS are those encoded by the conjugative transfer genes of the conjugative F and RP4 plasmids, and the virulence *virB*$_{Ti}$ genes of *A. tumefaciens*. To illustrate the architecture of the T4SS, first the T4SS$_{RP4}$ of the RP4 plasmid is briefly discussed followed by the more detailed functional analysis of the T4SS$_{Ti}$ of *A. tumefaciens*. Other T4SS, their functions and more detailed experimental findings, are excellently reviewed by Alvarez-Martinez & Christie (2009) and Fronzes, Christie & Waksman (2009).

The T4SS$_{RP4}$ channel of the conjugative plasmid RP4 is made up of the Trb$_{RP4}$ proteins TrbB$_{RP4}$-TrbL$_{RP4}$ encoded in the *trb*$_{RP4}$ operon of the *tra2*$_{RP4}$ region, except for TrbF$_{RP4}$, which is encoded in the *tra1*$_{RP4}$ region (Pansegrau *et al.*, 1994). The elements of T4SS$_{RP4}$ are largely homologous in function to elements of the T4SS$_{Ti}$ encoded by the *virB*$_{Ti}$ operon (Grahn *et al.*, 2000): TrbC$_{RP4}$ (VirB2$_{Ti}$), TrbD$_{RP4}$ (VirB3$_{Ti}$), TrbE$_{RP4}$ (VirB4$_{Ti}$), TrbF$_{RP4}$/TrbJ$_{RP4}$ (VirB5$_{Ti}$), TrbL$_{RP4}$ (VirB6$_{Ti}$), TrbH$_{RP4}$ and TrbK$_{RP4}$ (both possibly VirB7$_{Ti}$), TrbG$_{RP4}$ (VirB9$_{Ti}$), TrbI$_{RP4}$ (VirB10$_{Ti}$), and TrbB$_{RP4}$ (VirB11$_{Ti}$). There are no readily discernable functional homologues for the function of VirB1$_{Ti}$ and VirB8$_{Ti}$ with RP4 proteins (Fernandez *et al.*, 1996; Eisenbrandt *et al.*, 1999; Lawley *et al.*, 2003; Souza *et al.*, 2012). The extracellular pilus structure of RP4 is composed of TrbC$_{RP4}$ and creates a rigid and brittle structure (Haase *et al.*, 1995; Eisenbrandt *et al.*, 1999; Grahn *et al.*, 2000). The Ti plasmid T4SS$_{Ti}$ similarly encodes a pilin subunit VirB2$_{Ti}$ that is used for building a T-pilus on the cell surface.

The T4SS$_{Ti}$ of *A. tumefaciens* is encoded by the *virB*$_{Ti}$ region on the Ti plasmid. The channel structure is made up of 11 Vir$_{Ti}$ proteins (Fronzes, Christie & Waksman, 2009), that can be divided into energy using components (VirB4$_{Ti}$ and VirB11$_{Ti}$) responsible for the physical transfer of the substrate or for morphogenesis of the T-pilus structure (Sagulenko *et al.*, 2001), components of the inner membrane pore (VirB3$_{Ti}$, VirB6$_{Ti}$, VirB8$_{Ti}$ and VirB10$_{Ti}$), and proteins forming the outer membrane complex including the extracellular protruding T-pilus structure (VirB2$_{Ti}$, VirB5$_{Ti}$, VirB7$_{Ti}$ and VirB9$_{Ti}$). The T4SS$_{Ti}$ pilus associated protein VirB1$_{Ti}$ is not part of the structure, but is a transglycosylase that is supposed to aid in the T4SS$_{Ti}$ channel formation as well as the T-pilus biogenesis by opening the cell wall (Zupan *et al.*, 2007). Besides this simplified positional classification of the different Vir$_{Ti}$ proteins, the actual architecture of the transport channel can be deduced using structural data obtained for the closely related T4SS$_{pKM101}$ from the IncN plasmid pKM101 plasmid (Pohlman, Genetti & Winans, 1994; Mortelmans, 2006; Fronzes *et al.*, 2009; Chandran *et al.*, 2009; Rivera-Calzada *et al.*, 2013). Using near atomic scale cryo-electron microscopy a self-assembled core complex (VirB7$_{pKM101}$, VirB9$_{pKM101}$, and VirB10$_{pKM101}$) was observed that is made up of an inner layer and an outer layer, reaching from the inner membrane to the outer membrane, respectively. This core complex is thought to be the central structure of the T4SS$_{pKM101}$, where all other Vir$_{pKM101}$ proteins assemble around, possibly acting as a scaffold. It has an estimated size of 185 Å in diameter and height. The outer layer of the core complex is made up of the full-length VirB7$_{pKM101}$ and the C-terminal parts of VirB9$_{pKM101}$ and VirB10$_{pKM101}$. It is shaped as a helical barrel made up of several copies of the C-terminal part of VirB10$_{pKM101}$. VirB7$_{pKM101}$

anchors the C-terminus of $VirB9_{pKM101}$ into the inner leaflet of this barrel structure (Fronzes *et al.*, 2009; Chandran *et al.*, 2009). The outer layer connects to the distinct inner layer with small and thin linkers. The channel exhibits a constriction at the height of the linker region in between both layers. The diameter of the channel structure varies from 55 Å at its base to a 10 Å constriction at its tip (Rivera-Calzada *et al.*, 2013). These size changes also indicate that a conformational change of the structure is necessary for the substrate transport. It is thought that upon activation $VirB10_{Ti}$ undergoes a conformational change (Cascales & Christie, 2004), and mediates the full conformational change of the pore (Rivera-Calzada *et al.*, 2013). $VirB10_{Ti}$ is not involved in direct contact with the substrate, but indirectly orchestrates its movement through the $T4SS_{Ti}$. The ATPase proteins are located at the cytoplasmic side of the inner membrane, allowing for interaction with the T-complex for transport (Jakubowski *et al.*, 2009). **Figure 5** shows a simplified version of the channel structure.

Besides the transmembrane spanning channel structure, the $VirB_{Ti}/VirD4_{Ti}$ $T4SS_{Ti}$ further contains a structure called the T-pilus (**Figure 5**). The T-pilus forms an extra-cellular protruding structure of $VirB2_{Ti}$, with $VirB5_{Ti}$ forming the tip. It is thought that the T-pilus is not involved in the actual T-DNA transfer but in the interaction with the recipient cell (Christie, 2004; Aly & Baron, 2007; Grohman *et al.*, 2018). The extra-cellular T-pilus structure is long and semi-rigid and seems to be more flexible than the RP4 pilus (Eisenbrandt *et al.*, 1999; Lai & Kado, 2002).

The interaction path of the T-strand/$VirD2_{Ti}$ complex within the $T4SS_{Ti}$ was measured by using a specific transfer DNA immunoprecipitation (TrIP) assay (Cascales & Christie, 2004). After processed T-strands interact with the coupling protein $VirD4_{Ti}$, the second ATPase $VirB11_{Ti}$ binds to the single-strand DNA molecule. The subsequent steps during the transfer process are not fully understood yet. It is proposed that T-strands interact further with $VirB7_{Ti}$, $VirB8_{Ti}$ and $VirB9_{Ti}$ (Atmakuri, Cascales & Christie, 2004). As mentioned before, $VirB10_{Ti}$ guides the conformational change of the channel, but so far, this has only been measured indirectly. A specific mutation in $VirB10_{Ti}$ and its homolog $TraF_{pKM101}$ showed a 'gating' defect of the T4SS, allowing unregulated protein leakage in the extracellular space (Banta *et al.*, 2011).

From a broader perspective, the T4SS can be differentiated into three groups according to their function (Christie and Vogel, 2000; Cascales & Christie, 2003), without taking the phylogenetic distribution of all different T4SS into account (Guglielmini, de la Cruz & Rocha, 2013). All three groups can be found in Gram-negative bacteria (Alvarez-Martinez & Christie, 2009). First, the group participating in conjugative systems, for example, the conjugative plasmids F, R388 and RP4. This group of T4SS is the only one also present in Gram-positive bacteria as well as in archaea (Grohmann, Muth & Espinosa, 2003; Alvarez-Martinez & Christie, 2009). Secondly, the DNA uptake and release systems, that either import free DNA from the surrounding environment and provide for a natural competence, as it occurs for example in *Helicobacter pylori*, or excrete DNA to its environment, as *Neisseria gonorrhoea* is capable of. The third category consists of effector translocator systems, playing a profound role in the pathogenicity and infection of many microorganisms (Cascales & Christie, 2003). The T4SSs of *Helicobacter pylori* and of *Legionella pneumophila* belong to this group. The $T4SS_{Ti}$ encoded by the *A. tumefaciens* virulence system is special, as besides DNA, also a set of effector proteins are translocated (not bound at the T-complex), attributing to the bacterial virulence to infect and transform plant tissue. Nevertheless, the fundamental process of T-DNA transfer by $T4SS_{Ti}$ is similar to that by the true conjugative group, with the difference that effector proteins are translocated independently of DNA transfer. This grouping does not provide a complete and definite categorization of the different T4SS, but manages to divide the vast number of different systems into three defined by function. An earlier classification of the T4SS grouped the systems into a general IVA class with systems

similar to the *A. tumefaciens* $VirB_{Ti}/VirD4_{Ti}$ $T4SS_{Ti}$ (that included conjugative and effector translocator systems), class IVB with systems similar to the T4SS of IncI plasmids and the *Legionella pneumophila* Dot/Icm system, and the class 'others' for unclassified systems (Christie & Vogel, 2000; Christie *et al.*, 2005). A more recent study based on a phylogenic analysis of T4SS proposed to create a new classification based on the highly conserved VirB4 protein present in most T4SS (Guglielmini, de la Cruz & Rocha, 2013).

T4SS are able to transfer effector proteins with or without DNA transfer (Christie, 2004). In general, protein translocation requires a full T4SS and a CP as well as cell-to-cell contact. There is one exception with the *Bordetella pertussis* toxin translocation. The pathogenic *B. pertussis* contains a T4SS named Ptl, that has no $VirD4_{Ti}$ homolog, hence does not contain a CP. The pertussis toxin is transferred through the inner membrane via the general secretory (*sec*) pathway and from there it is translocated via the $T4SS_{Ptl}$ (Locht, Coutte & Mielcarek, 2011; Christie, Whitaker & González-Rivera, 2014). Toxin secretion does not require cell-to-cell contact; the pertussis toxin is rather secreted in the vicinity of host cells and taken up by cells due to an intrinsic property of the toxin.

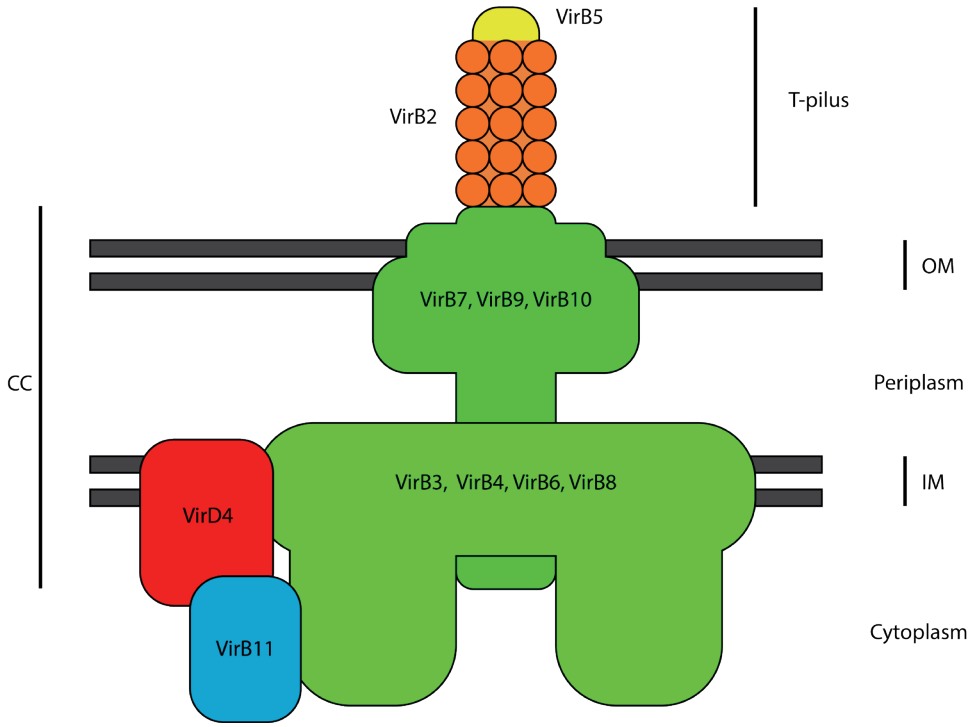


Figure 5. Simplified structure of a VirB/VirD4-type T4SS channel and pilus. This overview depicts the structural assembly across a Gram-negative cell wall spanning both across the inner membrane (IM) and outer membrane (OM). The core complex (CC) reaches across both membranes and comprises of VirB7, VirB9 and VirB10 on top and the central chamber complex of VirB3, VirB4, VirB6 and VirB8. Within the inner membrane, the CP VirD4 is connecting with an ATPase VirB11 to the main complex ATPase VirB4. The extracellular pilus structure is assembled from VirB2 and VirB5 aided by VirB1 (not shown as it is not part of the final assembled structure). The figure is not a true spatial representation.

Trans-cellular transport

The physical transfer of the conjugational substrate in between donor and recipient cell is still not fully determined. The initial cell-to-cell contact is concerted by the extracellular conjugative pilus structure (Willetts & Wilkins, 1984). However, the function of the extracellular conjugative pilus during the transfer is not clear. There are two different hypotheses for the transfer. In both cases, a mating group, an accumulation of cells of donor and recipient cells, must form first. The proximity of the cells then triggers the conjugative process. In one case, it is thought that two cells attached to each other, form a mating-bridge/pore for transfer of DNA and proteins by fusion of both membranes. In the other case, it is thought that the extracellular pilus structure is responsible for the DNA and protein transfer between two not necessarily touching cells. Experiments showed that conjugation could occur when cells were physically separated. DNA transfer, mediated by the F plasmid, was shown to be successful when cells were not interacting via their membranes. Those experiments led to the conclusion that the DNA must have been transferred via the extracellular F-pilus structure (Ou & Anderson, 1970; Achtman, Morelli & Schwuchow, 1978; Willetts & Wilkins, 1984; Harrington & Rogerson, 1990). On the other hand, a clear association of mating pairs via membrane-to-membrane seemed to be the pre-requisite for a successful DNA transfer using a T4SS of the F plasmid and RP4 plasmid (Panicker & Minkley, 1985; Samuels, Lanka & Davies, 2000). This indicated that DNA transfer was not performed via the pilus structure. Further evidence was provided by direct visualization of transformed cells using light microscopy. This experiment showed that a direct cell-to-cell contact was necessary for successful DNA transfer mediated by IncPβ R751 T4SS (Lawley *et al.*, 2002). This could be different for the $VirB2_{Ti}$/$VirB5_{Ti}$ Ti plasmid T-pilus as it contains an inner cylindrical lumen (Haase *et al.*, 1995; Eisenbrandt *et al.*, 1999; Fronzes, Christie & Waksman, 2009) that could allow for substrate translocation. So far, contrary to the F-pilus function as discusses above, DNA transfer via the Ti- and RP4-pilus structure has not been experimentally proven yet. While an early idea considered the formation of a mating bridge/pore by pore-forming proteins of the Mpf system to be the conjugational system (Willetts & Wilkins, 1984), it is now evident that the transmembrane spanning core complex of the VirB/virD4-like T4SS (for Tra_{RP4} as well as for the $VirB_{Ti}$) must be the mating pore able to deliver the T-complex into the recipient cell (Grohmann *et al.*, 2018). It is still unknown how the membrane of the recipient cell is manipulated for the conjugative transfer. So far, no evidence is available showing that a T4SS complex protrudes inside the recipient cell membrane.

Bacterial conjugation in biotechnology and implication for future research

With the background of increasing occurrences of human pathogens acquiring multi-resistance against antibiotic treatment, there is not only a need to research antibiotics but also a need for understanding the spread of antibiotic resistance through conjugational systems. Bacterial conjugation research was very active during the 1970s and 1980s, but stagnated in recent years. Understanding the interaction of individual parts of T4SS allows for identifying suitable targets to optimize the process, or if needed to efficiently disrupt the process. Furthermore, there can be more straightforward biotechnological applications of conjugative systems. Conjugative systems allow for a speedy transfer of large quantities of genetic information into a recipient cell. However, most methods need to be tailor-made for each particular task. Incompatibility of donor and recipient cell and material to be transferred commonly occurs. To find solutions to those problems, easy to handle and easy to combine conjugational systems are essential for efficient future applications.

Outline of this dissertation

Agrobacterium-mediated transformation is a robust system for plant transformation in academia and industry. Up to date, only a few improvements have been made to this methodology, for example adjustments of cultivation conditions, identification of high virulence mutations of *A. tumefaciens* or use of site-specific nucleases to achieve targeted insertions. A goal in plant transformation would be the use of protein therapy, enabling genomic modification without integration of foreign DNA, by just transferring effector molecules into a recipient cell. Protein therapy could be achieved for example by transferring non-integrative T-strands, from modified T-DNA, which will transiently lead to the expression of desired proteins. **Chapter 2** explores for the first time a direct modification of T-DNA termini, aimed to influence the transfer and integration of the T-DNA. A mutation in one of the T-DNA end-modifications was discovered, that influenced T-strand end processing within the recipient cells after transfer via the $T4SS_{Ti}$. In plant cells, this mutation facilitated faster protein expression, and in yeast, the mutation negatively influenced T-DNA integration.

In order to study single components of AMT more thoroughly, it would be advantageous to find an easy to handle model organism for high-throughput screening. Commonly, plant or yeast cells are used for AMT research, but all showed to be cumbersome and inefficient. *Streptomyces*, on the other hand, a Gram-positive bacterium from the same biome as *Agrobacterium*, shares a similar size, cultivation conditions as well as growth rates. Those attributes could be beneficial to create a new screening platform for AMT research. **Chapter 3** explores this model organism as a potential recipient for AMT and compared transfer by AMT with RP4-mediated transformation. Studying these two conjugative systems with their specific functions side by side in the same model organism is highly favorable for in-depth comparative analysis. Unfortunately, AMT to *Streptomyces* was not successful within the tested conditions, challenging previously published results. Nevertheless, for the first time, an RP4-mediated transformation from *A. tumefaciens* to *Streptomyces* was described. RP4-mediated transformation between *E. coli* strains and between *E. coli* and *Streptomyces* has been described previously, and is routinely used for biotechnological application. RP4-mediated transfer of genetic information between *Agrobacterium* and *Streptomyces* has not been reported yet in literature.

The plasmid RP4 was shown to be a very robust and capable translocation system for large DNA sequences. The encoded $T4SS_{RP4}$ is in comparison to the $T4SS_{Ti}$ of the Ti plasmid easier to handle in laboratory conditions. **Chapter 4** describes the characterization of the RP4 relaxase translocation signal. The signal is further tested for possible applications of protein and DNA transport. Here, for the first time, protein only transport from RP4 to recipient cells was shown; indicating that RP4 could be further modified and used efficiently in biotechnological applications.

The hybridization of both conjugational systems of the Ti plasmid and RP4 plasmid is described in **Chapter 5**. This segment ventures in the creation of a modular conjugational Dtr system that could be easily adapted for biotechnological adaptions. The hybridization of the relaxases was studied, as well as the exchange and use of the channel proteins between both systems and modification of the recognition sequence for DNA transfer. Single modules of the hybrid relaxases have been created and tested to be able to be transferred via T4SSs. Further, a fully functional operon expressing essential auxiliary proteins as well as the $VirD2_{Ti}$-$TraI_{RP4}$ hybrid relaxes was created. No transfer was detected so far, but a limited proof of principle for the modular conjugational system could be delivered.

Acknowledgements

This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Achtman M, Morelli G & Schwuchow S (1978) Cell-cell interactions in conjugating *Escherichia coli*: role of F pili and fate of mating aggregates. *J. Bacteriol.* **135:** 1053–1061

Aguilar J, Cameron TA, Zupan J & Zambryski P (2011) Membrane and core periplasmic *Agrobacterium tumefaciens* virulence Type IV secretion system components localize to multiple sites around the bacterial perimeter during lateral attachment to plant cells. *MBio* **2:** e00218-11

Alt-Mörbe J, Stryker JL, Fuqua C, Li PL, Farrand SK & Winans SC (1996) The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid *vir* genes. *J. Bacteriol.* **178:** 4248–4257

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* **73:** 775–808

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

Atmakuri K, Cascales E, Burton OT, Banta LM & Christie PJ (2007) *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *EMBO J.* **26:** 2540–2551

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

Babic A, Guérout A-M & Mazel D (2008) Construction of an improved RP4 (RK2)-based conjugative system. *Res. Microbiol.* **159:** 545–549

Bakkali M (2013) Could DNA uptake be a side effect of bacterial adhesion and twitching motility? *Arch. Microbiol.* **195:** 279–289

Balzer D, Pansegrau W & Lanka E (1994) Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.* **176:** 4285–4295

Banta LM, Kerr JE, Cascales E, Giuliano ME, Bailey ME, McKay C, Chandran V, Waksman G & Christie PJ (2011) An *Agrobacterium* VirB10 mutation conferring a type IV secretion system gating defect. *J. Bacteriol.* **193:** 2566–2574

Barton KA, Binns AN, Matzke AJM & Chilton MD (1983) Regeneration of intact tobacco plants containing full length copies of genetically engineered T-DNA, and transmission of T-DNA to R1 progeny. *Cell* **32:** 1033–1043

Becker EC & Meyer R (2012) Origin and fate of the 3' ends of single-stranded DNA generated by conjugal transfer of plasmid R1162. *J. Bacteriol.* **194:** 5368–5376

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Bhatty M, Laverde Gomez JA & Christie PJ (2013) The expanding bacterial type IV secretion lexicon. *Res. Microbiol.* **164:** 620–639

Brencic A & Winans SC (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol. Mol. Biol. Rev.* **69:** 155–194

Bryan LE, Elzen HMVD & Tseng JT (1972) Transferable drug resistance in *Pseudomonas aeruginosa*. *Antimicrob. Agents Chemother.* **1:** 22–29

Bundock P, Mróczek K, Winkler AA, Steensma HY & Hooykaas PJJ (1999) T-DNA from *Agrobacterium tumefaciens* as an efficient tool for gene targeting in *Kluyveromyces lactis*. *Mol. Gen. Genet.* **261:** 115–121

Bushman F (2002) Lateral DNA Transfer: Mechanisms and Consequences Cold Spring Harbor Laboratory Press

Byrd DR & Matson SW (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol. Microbiol.* **25:** 1011–1022

Cabezon E, Sastre JI & de la Cruz F (1997) Genetic evidence of a coupling role for the TraG protein family in bacterial conjugation. *Mol. Gen. Genet.* **254:** 400–406

Cabezon E, Ripoll-Rozada J, Pena A, de la Cruz F & Arechaga I (2014) Towards an integrated model of bacterial conjugation. *FEMS Microbiol. Rev.* **39:** 81–95

Canchaya C, Fournous G, Chibani-Chennoufi S, Dillmann M-L & Brüssow H (2003) Phage as agents of lateral gene transfer. *Curr. Opin. Microbiol.* **6:** 417–424

Carroll AC & Wong A (2018) Plasmid persistence: costs, benefits, and the plasmid paradox. *Can. J. Microbiol.* **64:** 293–304

Cascales E, Atmakuri K, Sarkar MK & Christie PJ (2013) DNA substrate-induced activation of the *Agrobacterium* VirB/VirD4 type IV secretion system. *J. Bacteriol.* **195:** 2691–2704

Cascales E & Christie PJ (2003) The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* **1:** 137–149

Cascales E & Christie PJ (2004) Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* **304:** 1170–1173

Chandran V, Fronzes R, Duquerroy S, Cronin N, Navaza J & Waksman G (2009) Structure of the outer membrane complex of a type IV secretion system. *Nature* **462:** 1011–1015

Chen I & Dubnau D (2004) DNA uptake during bacterial transformation. *Nat. Rev. Microbiol.* **2:** 241–249

Christie PJ (2001) Type IV secretion: intercellular transfer of macromolecules by systems ancestrally related to conjugation machines. *Mol. Microbiol.* **40:** 294–305

Christie PJ (2004) Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. *Biochim Biophys Acta* **1694:** 219–234

Christie PJ, Atmakuri K, Krishnamoorthy V, Jakubowski S & Cascales E (2005) Biogenesis, architecture, and function of bacterial type IV secretion systems. *Annu. Rev. Microbiol.* **59:** 451–485

Christie PJ & Gordon JE (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Christie PJ & Vogel JP (2000) Bacterial type IV secretion: conjugation systems adapted to deliver effector molecules to host cells. *Trends Microbiol.* **8:** 354–360

Citovsky V, de Vos G & Zambryski P (1988) Single-stranded DNA binding protein encoded by the *virE* locus of *Agrobacterium tumefaciens*. *Science.* **240:** 501–504

Christie PJ, Whitaker N & González-Rivera C (2014) Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843:** 1578–1591

Clark AJ & Adelberg EA (1962) Bacterial conjugation. *Annu. Rev. Microbiol.* **16:** 289–319

Clark AJ & Warren GJ (1979) Conjugal transmission of plasmids. *Annu. Rev. Genet.* **13:** 99–125

Costa TRD, Ilangovan A, Ukleja M, Redzej A, Santini JM, Smith TK, Egelman EH & Waksman G (2016) Structure of the bacterial sex F pilus reveals an assembly of a stoichiometric protein-phospholipid complex. *Cell* **166:** 1436–1444

Couturier M, Bex F, Bergquist PL & Maas WK (1988) Identification and classification of bacterial plasmids. *Microbiol. Rev.* **52:** 375–395

Datta N & Hedges RW (1972) Host ranges of R factors. *J. Gen. Microbiol.* **70:** 453–460

De Cleene M & De Ley J (1976) The host range of crown gall. *Bot. Rev.* **42:** 389–466

Del Solar G & Espinosa M (2002) Plasmid copy number control: an ever-growing story. *Mol. Microbiol.* **37:** 492–500

Ditta G, Stanfield S, Corbin D & Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **77:** 7347–3751

Dubnau D (1999) DNA uptake in bacteria. *Annu. Rev. Microbiol.* **53:** 217-244

Eisenbrandt R, Kalkum M, Lai EM, Lurz R, Kado CI & Lanka E (1999) Conjugative pili of IncP plasmids, and the Ti plasmid T pilus are composed of cyclic subunits. *J. Biol. Chem.* **274:** 22548–22555

Escudero J, den Dulk-Ras AD, Regensburg-Tuïnk TJG & Hooykaas PJJ (2003) VirD4-independent transformation by CloDF13 evidences an unknown factor required for the genetic colonization of plants via *Agrobacterium*. *Mol. Microbiol.* **47:** 891–901

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Fernández-López R, Garcillán-Barcia MP, Revilla C, Lázaro M, Vielva L & de la Cruz F (2006) Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. *FEMS Microbiol. Rev.* **30:** 942–966

Fernandez D, Spudich GM, Zhou XR & Christie PJ (1996) The *Agrobacterium tumefaciens* VirB7 lipoprotein is required for stabilization of VirB proteins during assembly of the T-complex transport apparatus. *J. Bacteriol.* **178:** 3168–3176

Fronzes R, Christie PJ & Waksman G (2009) The structural biology of type IV secretion systems. *Nat. Rev. Microbiol.* **7:** 703–714

Fronzes R, Schäfer E, Wang L, Saibil HR, Orlova E V & Waksman G (2009) Structure of a type IV secretion system core complex. *Science* **323:** 266–268

Frost LS, Leplae R, Summers AO & Toussaint A (2005) Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* **3:** 722–732

Fuqua C & Winans SC (1996) Conserved *cis*-acting promoter elements are required for density-dependent transcription of *Agrobacterium tumefaciens* conjugal transfer genes. *Journals Bacteriol.* **178:** 435–440

Fürste JP, Pansegrau W, Ziegelin G, Kröger M & Lanka E (1989) Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 1771–1775

Gao R & Lynn DG (2005) Environmental pH sensing: resolving the VirA/VirG two-component system inputs for *Agrobacterium* pathogenesis. *J. Bacteriol.* **187:** 2182–2189

Garcillán-Barcia MP, Alvarado A & De la Cruz F (2011) Identification of bacterial plasmids based on mobility and plasmid population biology. *FEMS Microbiol. Rev.* **35:** 936–956

Goodner B, Hinkle G, Gattung S, Miller N, Blanchard M, Qurollo B, Goldman BS, Cao Y, Askenazi M, Halling C, Mullin L, Houmiel K, Gordon J, Vaudin M, Iartchouk O, Epp A, Liu F, Wollam C, Allinger M, Doughty D, *et al.* (2001) Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science* **294:** 2323–2328

Grahn AM, Haase J, Bamford DH & Lanka E (2000) Components of the RP4 conjugative transfer apparatus form an envelope structure bridging inner and outer membranes of donor cells: implications for related macromolecule transport systems. *J. Bacteriol.* **182:** 1564–1574

Grohmann E, Christie PJ, Waksman G & Backert S (2018) Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* **107:** 455–471

Grohmann E, Muth G & Espinosa M (2003) Conjugative plasmid transfer in Gram-positive bacteria. *Microbiol. Mol. Biol. Rev.* **67:** 277–301

Guglielmini J, de la Cruz F & Rocha EPC (2013) Evolution of conjugation and type IV secretion systems. *Mol. Biol. Evol.* **30:** 315–331

Haase J, Lurz R, Grahn AM, Bamford DH & Lanka E (1995) Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. *J. Bacteriol.* **177:** 4779–4791

Halary S, Leigh JW, Cheaib B, Lopez P & Bapteste E (2010) Network analyses structure genetic diversity in independent genetic worlds. *Proc. Natl. Acad. Sci. U.S.A.* **107:** 127–132

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, Von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Harrington LC & Rogerson AC (1990) The F pilus of *Escherichia coli* appears to support stable DNA transfer in the absence of wall-to-wall contact between cells. *J. Bacteriol.* **172:** 7263–7264

Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* **20:** 262–267

Hedges RW & Jacob AE (1974) Transposition of ampicillin resistance from RP4 to other replicons. *Mol. Gen. Genet.* **132:** 31–40

Hermansson M & Linberg C (1994) Gene transfer in the marine environment. *FEMS Microbiol. Ecol.* **15:** 47–54

Hinnebusch J & Tilly K (1993) Linear plasmids and chromosomes in bacteria. *Mol. Microbiol.* **10:** 917–922

Hoekema A, Hirsch PR, Hooykaas PJJ & Schilperoort RA (1983) A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti plasmid. *Nature* **303:** 179–180

Hooykaas PJJ & Beijersbergen AGM (1994) The virulence system of *Agrobacterium tumefaciens*. *Annu. Rev. Phytopathol.* **32:** 157–181

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Hu X, Zhao J, DeGrado WF & Binns AN (2013) *Agrobacterium tumefaciens* recognizes its host environment using ChvE to bind diverse plant sugars as virulence signals. *Proc. Natl. Acad. Sci. U.S.A.* **110:** 678–683

Ingram LC, Richmond MH & Sykes RB (1973) Molecular characterization of the R factors implicated in the carbenicillin resistance of a sequence of *Pseudomonas aeruginosa* strains isolated from burns. *Antimicrob. Agents Chemother.* **3:** 279–288

Jacob AE & Grinter NJ (1975) Plasmid RP4 as a vector replicon in genetic engineering. **255:** 504-506

Jain A & Srivastava P (2013) Broad host range plasmids. *FEMS Microbiol. Lett.* **348:** 87–96

Jakubowski SJ, Kerr JE, Garza I, Krishnamoorthy V, Bayliss R, Waksman G & Christie PJ (2009) *Agrobacterium* VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Mol. Microbiol.* **71:** 779–794

Jin SG, Prusti RK, Roitsch T, Ankenbauer RG & Nester EW (1990) Phosphorylation of the VirG protein of *Agrobacterium tumefaciens* by the autophosphorylated VirA protein: essential role in biological activity of VirG. *J. Bacteriol.* **172:** 4945–4950

Johnson CM & Grossman AD (2015) Integrative and conjugative elements (ICEs): what they do and how they work. *Annu. Rev. Genet.* **49:** 577–601

Kado CI (1998) Origin and evolution of plasmids. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* **73:** 117–126

Lai EM & Kado CI (2000) The T-pilus of *Agrobacterium tumefaciens*. *Trends Microbiol.* **8:** 361–369

Lang S, Kirchberger PC, Gruber CJ, Redzej A, Raffl S, Zellnig G, Zangger K & Zechner EL (2011) An activation domain of plasmid R1 TraI protein delineates stages of gene transfer initiation. *Mol. Microbiol.* **82:** 1071–1085

Lanka E & Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64:** 141–169

Lawley TD, Klimke WA, Gubbins MJ & Frost LS (2003) F factor conjugation is a true type IV secretion system. *FEMS Microbiol. Lett.* **224:** 1–15

Lawley TD, Gordon GS, Wright A & Taylor DE (2002) Bacterial conjugative transfer: visualization of successful mating pairs and plasmid establishment in live *Escherichia coli*. *Mol. Microbiol.* **44:** 947–956

Lederberg J & Tatum EL (1946) Gene recombination in *Escherichia coli*. *Nature* **158:** 558

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

Lessl M, Pansegrau W & Lanka E (1992) Relationship of DNA-transfer-systems: essential transfer of plasmids RP4, Ti an F share common sequences. *Nucleic Acids Res.* **20:** 6099–6100

Levin RA, Farrand SK, Gordon MP & Nester EW (1976) Conjugation in *Agrobacterium tumefaciens* in the absence of plant tissue. *J. Bacteriol.* **127:** 1331–1336

Liang Z & Tzfira T (2013) *In vivo* formation of double-stranded T-DNA molecules by T-strand priming. *Nat. Commun.* **4:** 1–8

Liu C-N, Steck TR, Habeck LL, Meyer JA & Gelvin SB (1993) Multiple copies of *virG* allow induction of *Agrobacterium tumefaciens vir* genes and T-DNA processing at alkaline pH. *Mol. Plant-Microbe Interact.* **6:** 144–156

Llosa M, Gomis-Rüth FX, Coll M & de la Cruz F (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol. Microbiol.* **45:** 1–8

Locht C, Coutte L & Mielcarek N (2011) The ins and outs of pertussis toxin. *FEBS J.* **278:** 4668–4682

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

Matthysse AG (2014) Attachment of *Agrobacterium* to plant surfaces. *Front. Plant Sci.* **5:** 1–8

Melchers LS, Regensburg-Tuïnk TJ, Bourret RB, Sedee NJ, Schilperoort RA & Hooykaas PJJ (1989) Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *EMBO J.* **8:** 1919–1925

Mortelmans K (2006) Isolation of plasmid pKM101 in the Stocker laboratory. *Mutat. Res.* **612:** 151–164

Nair GR, Liu Z & Binns AN (2003) Reexamining the role of the accessory plasmid pAtC58 in the virulence of *Agrobacterium tumefaciens* strain C58. *Plant Physiol.* **133:** 989–999

Narasimhulu SB, Deng XB, Sarria R & Gelvin SB (1996) Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* **8:** 873–886

Nazarian P, Tran F & Boedicker JQ (2018) Modeling multispecies gene flow dynamics reveals the unique roles of different horizontal gene transfer mechanisms. *Front. Microbiol.* **9:** 1–11

Novick RP (1987) Plasmid incompatibility. *Microbiol. Rev.* **51:** 381–395

Orlek A, Stoesser N, Anjum MF, Doumith M, Ellington MJ, Peto T, Crook D, Woodford N, Walker AS, Phan H & Sheppard AE (2017) Plasmid classification in an era of whole-genome sequencing: application in studies of antibiotic resistance epidemiology. *Front. Microbiol.* **8:** 182

Ou JT & Anderson TF (1970) Role of pili in bacterial conjugation. *J. Bacteriol.* **102:** 648–654

Paget MS, Chamberlin L, Atrih A, Foster SJ & Buttner MJ (1999) Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **181:** 204–211

Panicker MM & Minkley EG (1985) DNA transfer occurs during a cell-surface contact stage of F-sex factor-mediated bacterial conjugation. *J. Bacteriol.* **162:** 584–590

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

Pansegrau W, Schoumacher F, Hohn B & Lanka E (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proc. Natl. Acad.* **90:** 11538–11542

Pansegrau W, Balzer D, Kruft V, Lurz R & Lanka E (1990) In vitro assembly of relaxosomes at the transfer origin of plasmid RP4. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 6555–6559

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

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncPα plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Phillips GJ & Funnell BE eds. (2004) Plasmid biology. American Society of Microbiology

Pohlman RF, Genetti HD & Winans SC (1994) Common ancestry between IncN conjugal transfer genes and macromolecular export systems of plant and animal pathogens. *Mol. Microbiol.* **14:** 655–668

Porter SG, Yanofsky MF & Nester EW (1987) Molecular characterization of the *virD* operon from *Agrobacterium tumefaciens*. *Nucleic Acids Res.* **15:** 7503–7517

Possoz C, Ribard C, Gagnat J, Pernodet JL & Guérineau M (2001) The integrative element pSAM2 from *Streptomyces*: kinetics and mode of conjugal transfer. *Mol. Microbiol.* **42:** 159–166

Rawlings DE & Tietze E (2001) Comparative biology of IncQ and IncQ-like plasmids. *Microbiol. Mol. Biol. Rev.* **65:** 481–496

Relić B, Andjelković M, Rossi L, Nagamine Y & Hohn B (1998) Interaction of the DNA modifying proteins VirD1 and VirD2 of *Agrobacterium tumefaciens*: analysis by subcellular localization in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **95:** 9105–9110

Rivera-Calzada A, Fronzes R, Savva CG, Chandran V, Lian PW, Laeremans T, Pardon E, Steyaert J, Remaut H, Waksman G & Orlova E V (2013) Structure of a bacterial type IV secretion core complex at subnanometre resolution. *EMBO J.* **32:** 1195–1204

Rogowsky PM, Close TJ, Chimera JA, Shaw JJ & Kado CI (1987) Regulation of the *vir* genes of *Agrobacterium tumefaciens* plasmid pTiC58. *J. Bacteriol.* **169:** 5101–5112

Rossi L, Hohn B & Tinland B (1993) The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plants. *Mol. Gen. Genet.* **239:** 345–353

Sagulenko E, Sagulenko V, Chen J & Christie PJ (2001) Role of *Agrobacterium* VirB11 ATPase in T-pilus assembly and substrate selection. *J. Bacteriol.* **183:** 5813–5825

Samuels AL, Lanka E & Davies JE (2000) Conjugative junctions in RP4-mediated mating of *Escherichia coli*. *J. Bacteriol.* **182:** 2709–2715

Saunders JR & Grinsted J (1972) Properties of RP4, an R factor which originated in *Pseudomonas aeruginosa* S8. *J. Bacteriol.* **112:** 690–696

Schröder G & Lanka E (2005) The mating pair formation system of conjugative plasmids - a versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54:** 1–25

Segal G, Russo JJ & Shuman HA (1999) Relationships between a new type IV secretion system and the icm/dot virulence system of *Legionella pneumophila*. *Mol. Microbiol.* **34:** 799–809

Shintani M, Sanchez ZK & Kimbara K (2015) Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Front. Microbiol.* **6:** 242

Smillie C, Garcillán-Barcia MP, Francia M V., Rocha EPC & de la Cruz F (2010) Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74:** 434–452

Sobecky PA & Hazen TH (2009) Horizontal gene transfer and mobile genetic elements in marine systems. In *Horizontal gene transfer: genomes in flux*, Gogarten MB Gogarten JP & Olendzenski LC (eds) pp 435–453. Totowa, NJ: Humana Press

Souza RC, del Rosario Quispe Saji G, Costa MOC, Netto DS, Lima NCB, Klein CC, Vasconcelos ATR & Nicolás MF (2012) AtlasT4SS: a curated database for type IV secretion systems. *BMC Microbiol.* **12:** 172

Szpirer CY, Faelen M & Couturier M (2000) Interaction between the RP4 coupling protein TraG and the pBHR1 mobilization protein Mob. *Mol. Microbiol.* **37:** 1283–1292

Thomas CM & Nielsen KM (2005) Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* **3:** 711–721

Thomas J & Hecht DW (2007) Interaction of *Bacteroides fragilis* pLV22a relaxase and transfer DNA with *Escherichia coli* RP4-TraG coupling protein. *Mol. Microbiol.* **66:** 948–960

Toro N, Datta A, Yanofsky M & Nester E (1988) Role of the overdrive sequence in T-DNA border cleavage in *Agrobacterium*. *Proc. Natl. Acad. Sci. U.S.A.* **85:** 8558–8562

Toro N, Datta A, Carmi OA, Young C, Prusti RK & Nester EW (1989) The *Agrobacterium tumefaciens virC1* gene product binds to overdrive, a T-DNA transfer enhancer. *J. Bacteriol.* **171:** 6845–6849

Turk SCHJ, van Lange RP, Regensburg-Tuïnk TJG & Hooykaas PJJ (1994) Localization of the VirA domain involved in acetosyringone-mediated *vir* gene induction in *Agrobacterium tumefaciens*. *Plant Mol. Biol.* **25:** 899–907

Tzfira T, Li J, Lacroix B & Citovsky V (2004) *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet.* **20:** 375–383

van Kregten M, Lindhout BI, Hooykaas PJJ & van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Mol. Plant. Microbe. Interact.* **22:** 1356–1365

Velappan N, Sblattero D, Chasteen L, Pavlik P & Bradbury ARM (2007) Plasmid incompatibility: more compatible than previously thought? *Protein Eng. Des. Sel.* **20:** 309–313

Vergunst AC, Schrammeijer B, den Dulk-Ras AD, de Vlaam CM, Regensburg-Tuïnk TJ & Hooykaas PJJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* **290:** 979–982

Vogel AM & Das A (1992) Mutational analysis of *Agrobacterium tumefaciens* VirD2: tyrosine 29 is essential for endonuclease activity. *J. Bacteriol.* **174:** 303–308

Vogelmann J, Ammelburg M, Finger C, Guezguez J, Linke D, Flötenmeyer M, Stierhof Y-D, Wohlleben W & Muth G (2011) Conjugal plasmid transfer in *Streptomyces* resembles bacterial chromosome segregation by FtsK/SpoIIIE. *EMBO J.* **30:** 2246–2254

Volkova VV, Lu Z, Besser T & Gröhn YT (2014) Modeling the infection dynamics of bacteriophages in enteric *Escherichia coli*: estimating the contribution of transduction to antimicrobial gene spread. *Appl. Environ. Microbiol.* **80:** 4350–4362

Waters VL, Hirata KH, Pansegrau W, Lanka E & Guiney DG (1991) Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of *Agrobacterium* Ti plasmids. *Proc. Natl. Acad. Sci. U.S.A.* **88:** 1456–1460

Willetts N & Wilkins B (1984) Processing of plasmid DNA during bacterial conjugation. *Microbiol. Rev.* **48:** 24–41

Zatyka M & Thomas CM (1998) Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiol. Rev.* **21:** 291–319

Ziegelin G, Fürste JP & Lanka E (1989) TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. *J. Biol. Chem.* **264:** 11989–11994

Ziegelin G, Pansegrau W, Lurz R & Lanka E (1992) TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin. *J. Biol. Chem.* **267:** 17279–17286

Ziegelin G, Pansegrau W, Strack B, Balzer D, Kröger M, Kruft V & Lanka E (1991) Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4. *DNA Seq.* **1:** 303–327

Zupan JR, Citovsky V & Zambryski P (1996) *Agrobacterium* VirE2 protein mediates nuclear uptake of single-stranded DNA in plant cells. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 2392–2397

Zupan JR, Hackworth CA, Aguilar J, Ward D & Zambryski P (2007) VirB1 promotes T-pilus formation in the *vir*-type-IV secretion system of *Agrobacterium tumefaciens*. *J. Bacteriol.* **189:** 6551–6563

Chapter 2

Terminal T-DNA modifications modulate protein expression and genomic integration

Felix T. Wittleben, Charlotte C. van der Velde, Roy van Gerven,
Krishna A.K. Gopie, Paul J.J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Abstract

Agrobacterium-mediated transformation (AMT) is a powerful method used widely in research and biotechnology. This study aimed to test direct modification of the AMT substrate, the single-stranded form of the transferred DNA (T-DNA) called the T-strand, to increase T-DNA transient expression within the recipient cell. T-DNA constructs were designed with modifications on the 3'- and/or 5'-terminal end to yield perfect hairpin (HP) structures, or a tRNA-binding site. The effects of these modifications on double-strand T-DNA formation, transient expression and integration were compared in relation to constructs lacking such modifications. The modifications leading to HP structures were introduced at the 5'- and/or the 3'-terminal end, flanking the remains of the RB_{Ti} and LB_{Ti} sequences, respectively. The tRNA-binding site was introduced at the 3'-end. Further, a T-DNA construct was created which should lead to T-strand hairpin structures on the 3'-end as well as on the 5'-end to theoretically form a closed DNA entity for transient expression, by possibly evading integration into the recipient genome. However, experiments showed firstly that T-DNA with two perfect HPs on both termini did not exhibit increased transient expression and were still able to stably transform yeast. Secondly, as a consequence of a fortuitously obtained mutation in the 3´-end hairpin an imperfect HP was created, which led to an increased *gus* reporter gene expression in plant cells within 48-72 h. Furthermore, it was found that this mutation decreased stable transformation, via homologous recombination, in yeast cells after AMT. These results show for the first time that a single border mutation can affect gene expression and integration of T-DNA in recipient cells.

Introduction

Agrobacterium tumefaciens uses a specifically adapted conjugational system to transform plant cells. Concisely, *Agrobacterium* harbors a tumor-inducing plasmid (Ti plasmid) containing a specific DNA sequence that can be transferred to plant cells, designated the transferred-DNA or in short T-DNA. The Ti plasmid further encodes the virulence proteins (Vir_{Ti}) responsible for T-DNA processing, including the formation of a single-stranded copy of the T-DNA (T-strand), and transfer thereof via a type-IV secretion system ($T4SS_{Ti}$) to recipient plant cells. Inside the recipient cells, the T-strand moves to the nucleus and integrates at random positions in the plant genome. Translocated effector virulence proteins assist in the transfer process. AMT is possible not only to plant cells, but also to yeast and fungi (Bundock *et al.*, 1995; de Groot *et al.*, 1998).

The T-DNA is defined by the *cis*-acting right-border (RB_{Ti}) and left-border (LB_{Ti}) imperfect 21-25 base pair (bp) repeat sequences (Yadav *et al.*, 1982). Supported by auxiliary proteins, these border sequences are recognized by the relaxase protein $VirD2_{Ti}$, which creates a single-stranded DNA break within each border sequence and thereby becomes covalently bound to the remaining 5'-ends. Subsequently the T-strand with $VirD2_{Ti}$ bound at the 5'-end of the RB_{Ti} sequence is displaced and transported by the $T4SS_{Ti}$ encoded by the $virB_{Ti}$ genes to plant cell recipients (for reviews: Herrera-Estrella *et al.*, 1988; Gelvin, 2003). Inside the cytosol of the recipient plant cell, the single-stranded DNA binding effector protein $VirE2_{Ti}$, also transported by the $T4SS_{Ti}$, is thought to attach to the T-strand, forming a so-called (hypothetical) T-complex (Howard & Citovsky, 1990). $VirE2_{Ti}$ potentially shields the T-strand from endonucleolytic attacks (Rossi, Hohn & Tinland, 1996), while at the same time $VirD2_{Ti}$ protects the RB_{Ti} end from exonucleolytic attacks (Dürrenberger *et al.*, 1989; Jasper *et al.*, 1994). Nuclear import is mediated by the nuclear localization signal (NLS) present on $VirD2_{Ti}$ (Ballas & Citovsky, 1997). Once inside the nucleus, the T-strand is prepared for integration. Most likely, the T-strand needs to be stripped off all attached effector proteins prior to integration, with the effector protein $VirF_{Ti}$ possibly facilitating the regulation of protein removal (Zaltsman *et al.*, 2013). So far, the integration process of the T-strand is not yet fully understood and described, as the integration process is dependent on host factors in the recipient cell. Evidence suggests that it is likely driven by double-strand break repair pathways (DSBR) in both yeast and plants (Salomon & Puchta, 1998; van Attikum, Bundock & Hooykaas, 2001; Chilton & Que, 2003). T-DNA integrates preferentially via the homologous recombination (HR) pathway in yeast, and via the non-homologous end joining (NHEJ) pathway in plants (Offringa *et al.*, 1990; Salomon & Puchta, 1998; van Attikum, Bundock & Hooykaas, 2001). In yeast, the factors of classical NHEJ (*yKu70*, *yKu80*, *yLig4*) were found to be essential for T-DNA integration (van Attikum, Bundock & Hooykaas, 2001). Upon integration within the recipient cell genome, the T-strand has to become a double-stranded DNA structure or designated T-DNA, allowing T-DNA located genes to be transcribed and expressed. However, little experimental evidence is available about the precise timing of the double-strand T-DNA formation in recipient cells, nor about the precise mechanism. A model describes a possibility of T-strand priming via primase functions or via random DNA/RNA fragment priming occurring in the nucleus of the recipient cell shortly after transfer and prior to integration (Liang & Tzfira, 2013). Screening of T-DNA integration sites in plants indicated that for the majority of insertion events the T-strands must have been rendered double-stranded prior to integration (Kleinboelting *et al.*, 2015). Recently, however it was found that the Pol-θ helicase-polymerase was essential for T-DNA integration, suggesting that the single-stranded T-strand may be captured directly for integration (van Kregten *et al.*, 2016); by a process of Pol-θ-mediated end joining (TMEJ). In yeast T-DNA integration, however, can occur by classical NHEJ (van Attikum, Bundock & Hooykaas, 2001) suggesting that here the double-stranded form of T-DNA is the integrated entity. The double-strand formation step could be the limiting factor determining the integration efficiency of T-DNA,

but the actual time point of priming, before or after the complete displacement of associated $VirE2_{Ti}$ proteins is unknown. In addition, no evidence exists about the time point of the removal of the covalently 5´-end bound $VirD2_{Ti}$.

It was found that T-DNA was expressed shortly after delivery, but expression diminishes over time (Janssen & Gardner, 1989). This has been interpreted as a consequence of temporary expression of unintegrated T-strands that have become double-stranded. Transfer of T-DNA molecules expressing the Cre recombinase provided direct evidence for the expression of non-integrative T-DNA molecules, as in cells eventually lacking integrated T-DNA, still Cre-mediated recombination events were detected (Vergunst & Hooykaas 1998). Before T-DNA located genes can be expressed, either in a transient manner or after integration, the transported T-strand must be of course converted to a double-stranded form within the recipient cell. The expression of proteins from T-DNA in plant cells via AMT was further optimized as a routine methodology, e.g. for antibody production, whereas the integration mechanisms of the used expression T-DNA constructs into the recipient cell genome were not investigated further (Kapila *et al.*, 1997; Menassa, Ahmad & Joensuu, 2012; Shah, Almaghrabi & Bohlmann, 2013). The fate of T-strands, which have become double-stranded T-DNA entities that are transcribed but not integrated in the genome of the recipient cell, is poorly understood. It is known that these non-integrative double-stranded T-DNA units can exhibit expression activity of encoded genes during their existence in the nucleus (Narasimhulu *et al.*, 1996; Singer *et al.*, 2012; Park *et al.*, 2015). In yeast, plasmids with only one RB_{Ti} border sequence can be transferred, linearized and re-circularized in the recipient cell to re-form a functional plasmid (Bundock *et al.*, 1995; Rolloos *et al.*, 2014). Also in plants, nuclear-localized T-DNA can form T-circles (Singer *et al.*, 2012) or be present as rather hard to detect linear double-stranded T-DNA units (Narasimhulu *et al.*, 1996). T-circles do not possess open ends that are immediately susceptible to exonuclease degradation via endogenous host factors and were shown to be detectable inside recipient cells (Singer *et al.*, 2012).

This study aimed to directly modify intrinsic properties of T-strand molecules, in this case allowing the 3'-terminal ends to more swiftly prime the synthesis of double-stranded structures, which should then lead to earlier and/or enhanced transient T-DNA expression. To facilitate double-strand formation either hairpin (HP) structures were inserted that could fold back on the T-strand or a tRNA-binding site was inserted at the 3'-end. The presence of HP structures, also at the 5'-end, might also prevent genomic integration. It was found that one particular 3´-end HP structure led to an increased reporter gene expression in plants and decreased integration in yeast. Remarkably, the cause of the detected increase was discovered to be a mutation within the intended 3'-end HP structure, but the mode of function of this end modification could not be fully explained. Nevertheless, this study shows for the first time compelling evidence that end-modification of T-DNA could lead to higher levels of transgene expression within the recipient cell.

Results

A novel approach to influence post-transfer T-DNA gene expression and T-DNA processing within the recipient cell genome was investigated. T-DNA termini were adapted to see the influence of HP structures on both ends on double-strand T-DNA formation and T-DNA integration. Further, the effect of an RNA-priming site, to stimulate double-strand formation on the 3'-terminus, was tested. One of the modifications was the introduction of HP structures on the RB_{Ti} (5`-terminus/end) and LB_{Ti} (3´-terminus/end). After nicking and single-strand displacement of the T-strand, such modified 3´- or 5´-termini were supposedly able to fold back and create a HP.

Theoretically, when a HP folds back perfectly at the LB_{Ti} (3´-terminus), the resulting double-stranded loop structure provides for a free 3'-end, that can readily be used for chain elongation and double-strand formation along the complete T-strand. In nature, folded HPs creating secondary DNA structures are common in initiating replication due to direct interaction with regulatory proteins (Bikard *et al.*, 2010). In this case, the folded HP can be viewed as a stem-loop mimicking a branched structure initiating the endogenous double-strand formation within the recipient cell (Pearson *et al.*, 1996; Bikard *et al.*, 2010). On the contrary, a HP forming at the RB_{Ti} (5´-terminus) would only generate a looped structure with a free 5'-end phosphate, which by itself is no substrate for chain elongation, but which might finally be covalently connected to a growing 3'-end by ligation. Furthermore, any existing HP structure could shield the T-strand from an exonucleolytic attack within the recipient cell. In order to gain an overview about the potential effects of HP formation on T-DNA metabolism, T-DNA molecules with several different HP variants were created. One version contained HPs on both sides to create potentially a linear T-strand without open ends. Constructs with such double-HP insertions could theoretically be used as vehicles for swiftly generating double-stranded T-DNA molecules in recipient cells that can be transiently expressed, but are not prone to integration into the recipient genome. In another version, instead of a HP, a tRNA-binding site was inserted adjacent to the LB_{Ti} border sequence (3'-end). The tRNA-binding site is supposed to initiate the direct priming of double-strand DNA formation via the annealed 3'-end of an endogenous free tRNA molecule within the recipient cells. Such a priming strategy using a tRNA-priming site for site-specific strand initiation naturally occurs within the fungal Ty1 retrotransposon, but also within viral replication units in eukaryotic host cells (Friant *et al.*, 1996; Burnett & McHenry, 1997).

The T-DNA variations were introduced at the distal ends of T-DNA sequences aimed to be functioning in yeast or plant cells. In order to create a comparable construct with the possibility to study genomic integration in both recipient organisms the following modifications had to be made. For yeast transformation, the T-DNA contained sequences homologous to the *PDC6* locus to mediate genomic integration via HR. Homology is essential in yeast, to mediate efficient genomic integration. A *KanMX* cassette was inserted in between the *PDC6* segment of homology, to determine successful yeast transformation via antibiotic selection (**Figure 1**). A conventional co-cultivation method between *A. tumefaciens* and yeast was used for T-DNA transfer, modified from Bundock *et al.* (1995). For detection of transfer in plants, a *gusA* cassette was inserted in between the *PDC6* sites, at the same location as the *KanMX* cassette (**Figure 1**). The selection cassettes would only be transcribed and functionally expressed if rendered double-stranded after being successfully transferred. Expression of the *gusA* cassette, in plant cells, can be detected by histochemical staining (X-Gluc staining) or an enzymatic assay (MUG assay). The GUS detection would be indicative of double-strand formation either before or after integration. Leaves of *Nicotiana benthamiana* were directly agroinfiltrated (Sakalis, 2013) and after AMT *in planta* GUS expression was detected using a sensitive MUG assay.

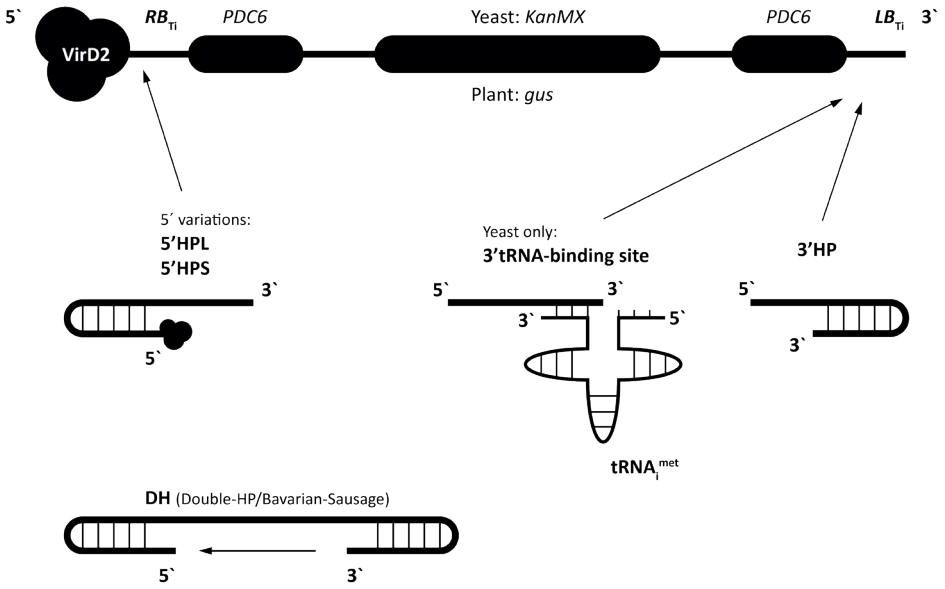


Figure 1. T-DNA construct to study the effect of terminal border sequence modifications. The T-DNA of plasmid p14Y was used to insert different hairpin (HP) variants on the 5´- and 3´-ends, without interfering with the original border sequences. **5'HPL**: 5´-end long HP; **5'HPS**: 5`-end short HP; **3'HP**: 3´-end HP; **DH**: 5´-end and 3´-end double-HP construct harboring both 5'HPL and 3'HP; **3'tRNA-binding site**: a tRNA-binding site for RNA priming of the T-DNA strand within the recipient cell (yeast only for this study). The marker was either for use in plants (*gus*) or yeast (*KanMX*). For transformation of yeast by HR, sequences homologous to the *PDC6* locus were inserted internally to the border sequences and flanking the *KanMX* marker. In plants, these *PDC6* sequences share no known homology with the genomic DNA.

Obviously, HP structures might be formed as soon as T-strand molecules are generated, not only after transfer of T-strands to recipient cells, but also already inside *Agrobacterium* cells, when single-strand formation ensues. Therefore, when interpreting the read out of the experiments by assessing reporter gene expression, it has to be realized that HP effects located within *Agrobacterium* and HP effects located within the recipient cell can be intertwined. For instance, positive effects on reporter gene expression due to 3'-end HP formation in recipient cells might be nullified by negative effects on T-strand transfer due to precocious double-strand formation in *Agrobacterium*.

Construction of the T-DNA terminal modifications

A multipurpose construct was created (**Figure 1**) to accommodate the different modifications and variations of T-DNA border structures, as well as the traits required. HP structures were based on a hinge derived from the *Saccharomyces cerevisiae* initiator methionine tRNA ($tRNA_i^{met}$) anticodon-loop with the sequence CUCAUAA (Friant *et al.*, 1996), reading CTCATAA as DNA sequence. The HP sequences were introduced via oligonucleotide insertion at the desired location. The theoretical folding of the HP sequences is depicted in **figure 2A**.

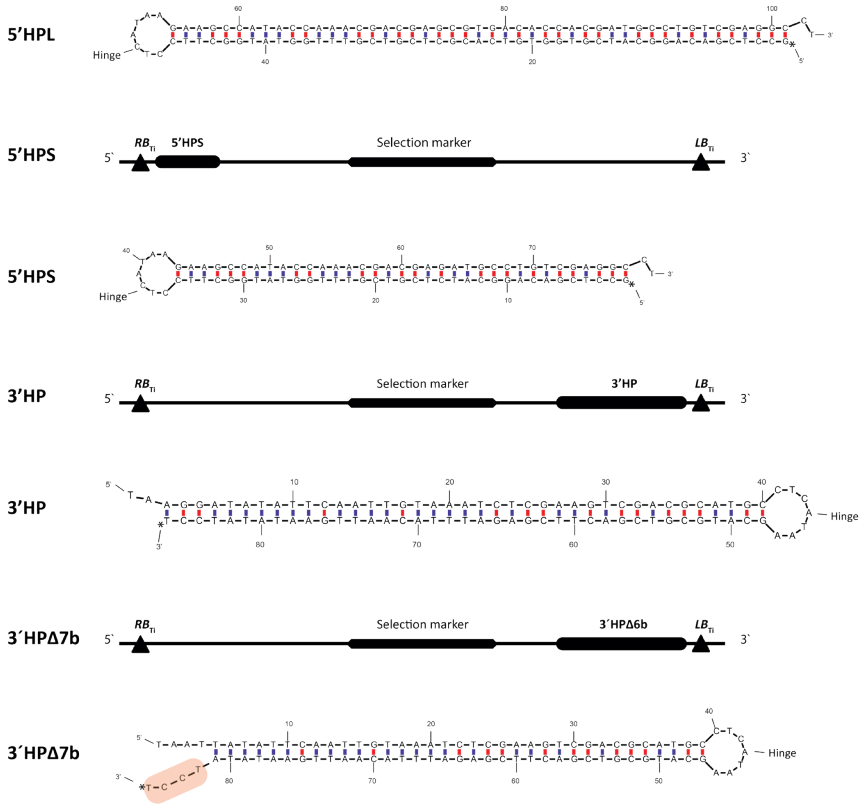
For the 3´-terminus HP structure (**3'HP**) the hinge sequence was located directly upstream of the octopine LB_{Ti} (Rolloos, 2013). When the LB_{Ti} would be nicked, the displaced single-strand could theoretically fold back on itself via the hinge sequence, since the 38 base (b) sequence located between the hinge sequence and the LB_{Ti} nick site is again present in a reverse complementary sequence upstream of the hinge. This should then generate a short double-stranded structure with a free 3'-end for priming of DNA synthesis, starting from this HP fold (**Figure 2A**), and yielding a full-length double-stranded structure after DNA synthesis.

To see the influence of a 5´-terminal HP sequence, different structures were introduced downstream of the octopine RB_{Ti} sequence (Rolloos, 2013). For this HP, it was important not to recreate another RB_{Ti} sequence at the reverse complementary sequence downstream of the hinge. A large HP structure was created (**5'HPL**) with a 47 b sequence between the RB_{Ti} 5´-nick site and the hinge, which was present as well as a reverse complementary sequence behind the hinge (**Figure 2A**). A shorter HP version (**5'HPS**) containing a 35 b sequence was created too (**Figure 2A**).

To see the effect of HP structures on both termini of the T-DNA, a double-HP was created (**DH**). This encompassed the 5´HPL and the 3´HP sequence together in one T-DNA molecule. This construct was theoretically capable forming a complete closed molecule after initiating double-strand formation via the 3'HP, and ligating back to the 5'HPL. Theoretically, the double-strand formation could occur in the recipient cell nucleus, using the endogenous DNA replication machinery. The double-strand synthesis primed via the 3'HP should meet with the 5'-end of the 5´HP, and after ligation of the remaining nick a double-strand T-DNA entity without open ends would be created (**Figure 1**). The encoded gene of interest could be then expressed without prior integration into the recipient cell genome. Further, this entity would be protected from immediate exonucleolytic degradation.

Another mode of priming to initiate double-strand formation of the T-strand was the introduction of a tRNA-binding site (**Figure 1**). To explore this priming method, the $tRNA_i^{met}$-binding sequence (Friant *et al.*, 1998) of the yeast retrotransposon Ty1 was inserted at the 3´-terminus of the T-DNA upstream of the LB_{Ti} nicking site. When the T-strand would be introduced into the recipient cell nucleus, the freely available abundant $tRNA_i^{met}$ could bind at the sequence and create an RNA-primer for double-strand synthesis. The RNA-priming-site was not tested in plants.

To study the effect of the HP in plants, the HP variations were introduced in a modified vector. The explanation and nomenclature of the created constructs is as follows: the base plasmid used was p14-PDC6KX, a binary vector containing T-DNA with a *PDC6* homology sequence for integration into a yeast genome with an embedded *KanMX* resistance cassette between an octopine RB_{Ti} and an octopine LB_{Ti} sequence. For experiments in plants, the reporter protein GUS expression cassette, containing the strong plant specific *CaMV 35S* promoter, was introduced into the construct. This created the plasmid line p14P (P = plants). The GUS cassette was inserted in such an orientation that the T-strand contained the *gusA* as non-coding sequence, inhibiting any potential direct single-strand expression within plant cells. The plant vectors contained several HP modifications and were called as follows: for the HP on the 5'-terminus (large and small version; **Figure 2A**) of the T-DNA: p14P5'HPL and p14P5'HPS. For the HPs on the 3'terminus (**Figure 2A**) of the T-DNA: p14P3'HP. For the double-HP (**Figure 2A**), a combination of the 5'- and 3'-terminus HPs: p14PDH. All plasmids used are also listed in **table 3**.



B

```
                                                                                  Hinge
5'HPL     5´-  --CAATTACAACGGTATATATCCT*GCCTCGACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCCTCATAAGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTCGAGGCCT--  -3'
5´HPS     5´-  --CAATTACAACGGTATATATCCT*GCCTCGACAGGCATCTCGTCGTTTGGTATGGCTTCCTCATAAGAAGCCATACCAAACGACGAGATGCCTGTCGAGGCCT--  -3'
                  RB Ti                                                           Hinge

3'HP      5´-  --TAATTAAAGGATATATTCAATTGTAAATCTCGAAGTCGACGCATGCCTCATAAGCATGCGTCGACTTCGAGATTTACAATTGAATATATCCT*GCC----  -3'
3´HPΔ7b   5´-  --TAAT-------TATATTCAATTGTAAATCTCGAAGTCGACGCATGCCTCATAAGCATGCGTCGACTTCGAGATTTACAATTGAATATATCCT*GCC----  -3'
                                                  Hinge                                  LB Ti
```

Figure 2. A: Folding of the terminal HP insertions; a graphic representation of the T-DNA is followed by the predicted folding of the HP sequence. 5'HPL: 5´-end long HP; 5'HPS: 5`-end short HP; 3'HP: 3´-end HP; 3´HPΔ7b: a 7 b deletion mutant of 3'HP creating a fold with a 4 b overhang at the 3´-end of the T-DNA. The 3´HPΔ7b base overhang is marked with a red box. The asterisk marks nick between bases created by the relaxase. The sequence of the folding hinge CTCATAA represents the DNA version of the methionine tRNA anticodon loop (Friant *et al.*, 1996). For the HP folding only the actual displaced single-stranded T-strand is depicted. The DNA folding models were created using the web server Mfold (Zuker, 2003) **B:** Sequence comparison of 5'HPL, 5'HPS, 3'HP and 3´HPΔ7b single T-strands. Only the actual displaced single-strand is depicted. The folding hinge is shown in bold. The asterisk marks the nick created by the relaxase. In case of the RB_{Ti}, VirD2$_{Ti}$ will bind to the G base (next to the asterisk), and in case of the LB_{Ti} VirD2$_{Ti}$ will nick the phosphate backbone next to the T base. The bold italic sections at both ends indicate where the 5´-end can snap back to the 3'-sequence in case of the RB_{Ti} and the 3´-end can snap back to the 5'-sequence at the LB_{Ti}. The part of the T-DNA that is before the nicking site and not part of the T-strand but still part of the border sequence is shown in grey. The two minus bars indicate continuation of the T-DNA sequence. The 3´HPΔ7b base overhang is marked with a red box. For 3´HPΔ7b the base deletion is indicated with 7 thick bars within the sequence.

To study the effect of the HP in yeast, the HP variations were introduced in a modified vector. The explanation and nomenclature of the created constructs is as follows: the base plasmid used was p14. For experiments with yeast, the antibiotic *KanMX* selection cassette, containing the yeast specific *TEF1* promoter, was introduced into the construct. The construct did not contain a yeast specific origin of replication. This created the plasmid line p14Y (Y = yeast). The yeast vectors containing the HP modifications used were called as follows: for the HP on the 5'-terminus (large and small version; **Figure 2A**) of the T-DNA: p14Y5'HPL and p14Y5'HPS. For the HPs on the 3'-terminus (**Figure 2A**) of the T-DNA: p14Y3'HP. And for the double-HP, a combination of the 5'- and 3'-terminus HPs: p14YDH. For the 3'-terminus t-RNA binding site on the T-DNA: p14YSctRNA. All plasmids used are also listed in **table 3.**

A mutation within the 3′-end HP T-DNA showed increased GUS expression in *N. benthamiana*

Sequencing of the HP T-DNA emerged to be problematic, due to the difficult nature of the sequences with recurrent repeats and HP structures. This led to the testing of multiple clones of each HP, using AMT (agroinfiltration) on *N. benthamiana* leaves followed by enzymatic assays for GUS expression (X-Gluc staining and MUG analysis). After this initial testing, one clone with a 3'-end HP modification showed a higher effect in GUS expression than comparable HP constructs in initial screenings (data not shown). Further analysis revealed a discrepancy with its 3′-end HP sequence. It appeared to contain a deletion of 7 b that will lead to a subsequent shortening of the fold sequence (**Figures 2B, S2A and S2B**). The deletion was directly situated at the 5′-upstream site of the homologous sequence for the HP to fold back (**Figure 2B**). Re-sequencing of the particular clone only resulted in incomplete sequencing due to many repetitive sequences and a possible difficult structure. With the analysis of the sequencing reads (**Figures S2A and S2B**), it seemed that a possible 4 b overhang could occur, instead of a complete HP folding. DNA folding modeling of this mutated sequence led to the conclusion that a 4 b overhang could occur (**Figure 2A**). This HP mutation construct was renamed into **3′HPΔ7b** (3′: location on the T-DNA; HP: hairpin; Δ: deletion mutant; 7b: 7 bases deleted). This mutation was then tested further and used in this study. Plasmids containing this mutation within the HP were named p14Y3′HPΔ7b (for yeast selection) and p14P3′HPΔ7b (for plant selection). When this mutation was used in combination with a 5′-end HP insertion this was called **DHΔ7b** (DH: double-HP; Δ: deletion mutant; 7b: 7 base deletion), and the plasmids were named p14YDHΔ7b (for yeast selection) and p14PDHΔ7b (for plant selection).

Agroinfiltration of the single 3′-end 3'HP into leaves of *N. benthamiana* showed a significant increase in relative GUS expression with the 3′HPΔ7b mutation compared to two independent sub-clones with complete folding HP constructs, as seen in **figure 3**.

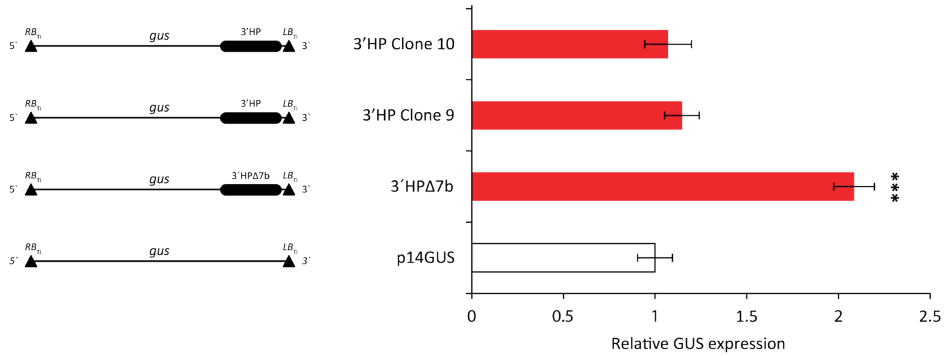


Figure 3. Relative GUS expression by the 3´HPΔ7b HP mutant compared to complete 3'HP constructs. Relative GUS expression after *N. benthamiana* transformation with T-DNA constructs after 48 h. The constructs 3´HPΔ7b (n = 7) and 3'HP Clone 9 (n = 7) and Clone 10 (n = 7) were all normalized to the baseline expression of p14GUS (n = 19), which was set at 1. Error bars depicted are SEM. There was a statistically significant difference between GUS expression constructs as determined by one-way ANOVA ($F(3,33) = 16.571$, $p < 0.0005$, $\omega^2 = 0.55$). The data was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > 0.05$) and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = 0.366$). A Tukey-Kramer *post hoc* test showed that 3´HPΔ7b was significantly higher expressed compared to 3'HP Clone 9 ($p < 0.0005$), 3'HP Clone 10 ($p < 0.0005$) and p14GUS ($p < 0.0005$). There was no statistically significant difference between p14GUS and 3'HP Clone 9 ($p = 0.818$) and p14GUS with 3'HP Clone 10 ($p = 0.972$). Plasmids used were p14P (p14GUS), p143'HP (3'HP, two independent sub-clones, Clone 9 and Clone10), and p14P3´HPΔ7b (3´HPΔ7b). *** for $p < 0.0005$ compared to the p14GUS construct.

The GUS expression level was assessed by histochemical staining and colorimetric analysis (X-Gluc staining and MUG analysis). After 48 h, a statistically significant increase of relative GUS expression was measured. The same effect was observed when the 3´HPΔ7b mutation was joined with the 5´-end 5'HPL to form a double-HP construct DHΔ7b. The DHΔ7b evoked the same statistically significant increase of relative GUS expression, compared to the two other independent double-HP constructs with complete HP sequences (**Figure 4**).

The effect of the 3´HPΔ7b mutation could be replicated when compared with a construct without HPs, with a perfect 5´-end 5'HPL or also in combination as double-HP DHΔ7b, with the 3´HPΔ7b fused with a 5´-end 5'HPL. In this comparative analysis, the mutation retained the observed statistically significant effect of increased GUS expression (**Figure 5**). In this experiment, the 5´-end HP 5'HPL had the same level of GUS expression compared to the baseline T-DNA GUS expression after 48 h (**Figure 5**).

T-DNA equipped with the DHΔ7b shows a rapid increase of GUS expression

When the double-HP construct DHΔ7b, containing the 3´HPΔ7b mutation combined with the 5´-end 5'HPL, was introduced into plant leaf tissue, the difference in GUS expression was most pronounced between 24 h and 48 h. Compared to the control expression of the GUS cassette on a construct without HP, the 5'HPL was not significantly lower initially, but between the measured time points of 24 and 48 h, a strong increase in GUS expression was detected. The increased expression level was detected even at a 72 h and a 10 d time-point after agroinfiltration (**Figure 6**). In comparison, the expression level of all pooled control constructs for each time point, with only an RB_{Ti} and LB_{Ti}, showed no increase of expression during the measured time interval (**Figure 6**).

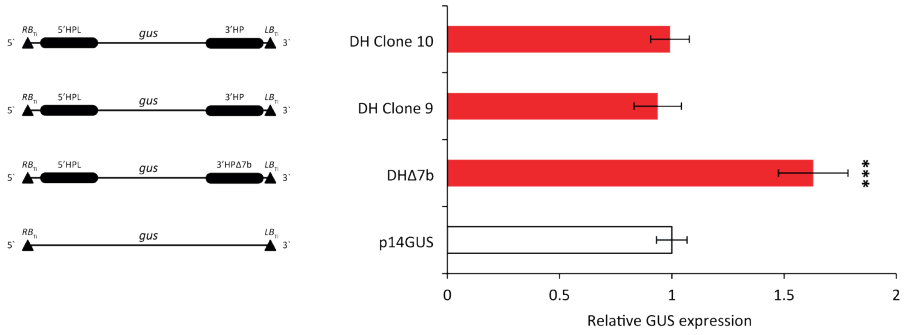


Figure 4. Relative GUS expression by the DHΔ7b HP mutant compared to complete DH constructs. All clones of the 3´-end HP (3'HP) experiment were combined with the 5´-end HP (5'HPL) and transferred into *N. benthamiana* leaves using agroinfiltration. The constructs DHΔ7b (n = 7) and DH Clone 9 (n = 6) and Clone 10 (n = 6) were all normalized to the baseline expression of p14GUS (n = 19) set at 1. Time for co-cultivation was 48 h. Error bars depicted are SEM. There was a statistically significant difference between GUS expression constructs as determined by one-way ANOVA ($F(3,34) = 8.767$, $p < 0.0005$, $\omega^2 = 0.38$). The data was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > 0.05$) and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = 0.785$). A Tukey-Kramer *post hoc* test showed that DHΔ7b was significantly higher expressed compared to DH Clone 9 ($p = 0.001$) and DH Clone 10 ($p = 0.003$) and p14GUS ($p < 0.0005$). There was no statistically significant difference between p14GUS and DH Clone 9 ($p = 0.971$) and p14GUS with DH Clone 10 ($p = 1.000$). Plasmids used were p14P (p14GUS), p14PDH (two independent clones #9 and #10), and p14PDHΔ7b (DHΔ7b). *** for $p < 0.0005$ compared to the p14GUS construct.

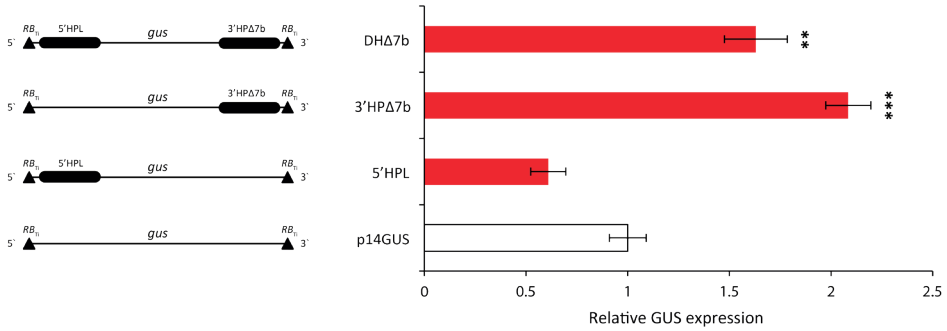


Figure 5. Increased GUS expression if T-DNA was combined with the 3´HPΔ7b HP mutant. When the 3´HPΔ7b is added to the T-DNA construct or combined with a 5'HPL HP, an increased GUS expression in *N. benthamiana* leaves was detected. For this analysis the data of DHΔ7b of **figure 4**, and 3´HPΔ7b of **figure 3** was used. The 5´-end HP 5'HPL alone was not able to create this effect. The data sets of 5´HP (n = 6), 3´HPΔ7b (n = 7) and DHΔ7b (n = 7) were normalized to the p14GUS expression (n = 20) that was set as 1. Time for co-cultivation was 48 h. The error bar depicted is SEM. There was a statistically significant difference between GUS expression constructs as determined by one-way ANOVA ($F(3,36) = 22.165$, $p< 0.0005$, $\omega^2 = 0.61$). The data was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > 0.05$) and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = 0.225$). A Tukey-Kramer *post hoc* test showed that DHΔ7b was significantly higher expressed compared to 5'HPL ($p < 0.0005$) and p14GUS ($p = 0.003$). There was no significant difference between DHΔ7b and 3´HPΔ7b ($p = 0.129$). The 3´HPΔ7b construct was also significantly higher expressed compared to 5'HPL ($p < 0.0005$) and p14GUS ($p < 0.0005$). There was no difference in expression between 5'HPL and p14GUS ($p = 0.138$). Plasmids used were p14P (p14GUS), p14P5'HPL (5'HPL), p14P3´HPΔ7b (3´HPΔ7b) and p14PDHΔ7b (DHΔ7b). ** for $p < 0.005$ and *** for $p < 0.0005$ compared to the p14GUS construct.

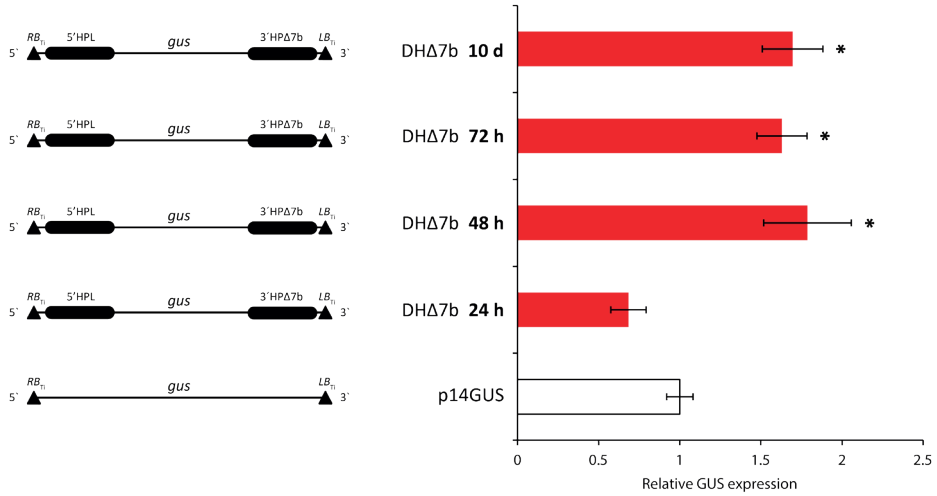


Figure 6. Temporal effect of DHΔ7b HP mutant on relative GUS expression in *N. benthamiana* leaves. Time-points for GUS measurements were taken for 24 h, 48 h, 72 h, and 10 d post infiltration. The initial peak of the DHΔ7b was seen in between 24-48 h. Elevated levels were detected until 10 d. The time-point data sets, (each n = 6), were normalized to the p14GUS expression (n = 24) which was set as 1. Error bars are SEM. There was a statistically significant difference between GUS expression constructs as determined by one-way ANOVA ($F(4,45) = 9.726$, $p < 0.0005$, $\omega^2 = 0.41$). The data was normally distributed for each group, as assessed by Shapiro-Wilk test ($p > 0.05$) and there was homogeneity of variances, as assessed by Levene's test for equality of variances ($p = 0.368$). A Tukey-Kramer *post hoc* test showed that the DHΔ7b construct was not higher expressed compared to p14GUS after 24 h ($p = 0.510$). Though DHΔ7b was expressed higher significantly after 48 h ($p = 0.02$), 72 h ($p = 0.013$) and 10 d ($p = 0.009$). Plasmids used were p14P (p14GUS) and p14PDHΔ7b (DHΔ7b). * for $p < 0.05$ compared to the p14GUS construct.

To study the long-term effect of the DHΔ7b in plant cells, a T-DNA co-transformation tumor assay was carried out in *Nicotiana glauca*. The constructs were introduced into *A. tumefaciens* LBA1010, containing oncogenic T-DNA. During AMT, the oncogenic T-DNA is transferred at the same time as the introduced T-DNA construct. During the subsequent tumor formation, the modified T-DNA could also be expressed, either after integration into the genome, or remaining as a non-integrated copy. A decrease in staining with one of the HP constructs, when compared to a control vector without HPs, could indicate possible transient expression of GUS. After *N. glauca* stem infection, tumors were allowed to develop for 21 d before sampling. X-Gluc staining revealed pronounced heterogeneous blue staining of the tumor tissue (**Figure 7**). It was unclear if a decrease of X-Gluc staining intensity occurred with the HP constructs compared to the control p14GUS construct. The 3´HPΔ7b mutant or 3'HP (no mutation) showed no visible difference in GUS expression between the constructs. In addition, no difference in GUS expression was detected with the DH and its mutant 3´DHΔ7b. None of the constructs showed a detectable decrease of X-Gluc staining after 21 d. The tumor size did not differ significantly between the wild-type control, using the tumorigenic *A. tumefaciens* LBA1010 alone without HP constructs, compared to the strains containing vectors with HPs during the co-transformation. No X-Gluc staining was detected in the tumors created by the control strain wild-type *A. tumefaciens* LBA1010 (**Figure 7**). This indicated that the DHΔ7b T-DNA potentially integrates after transformation.

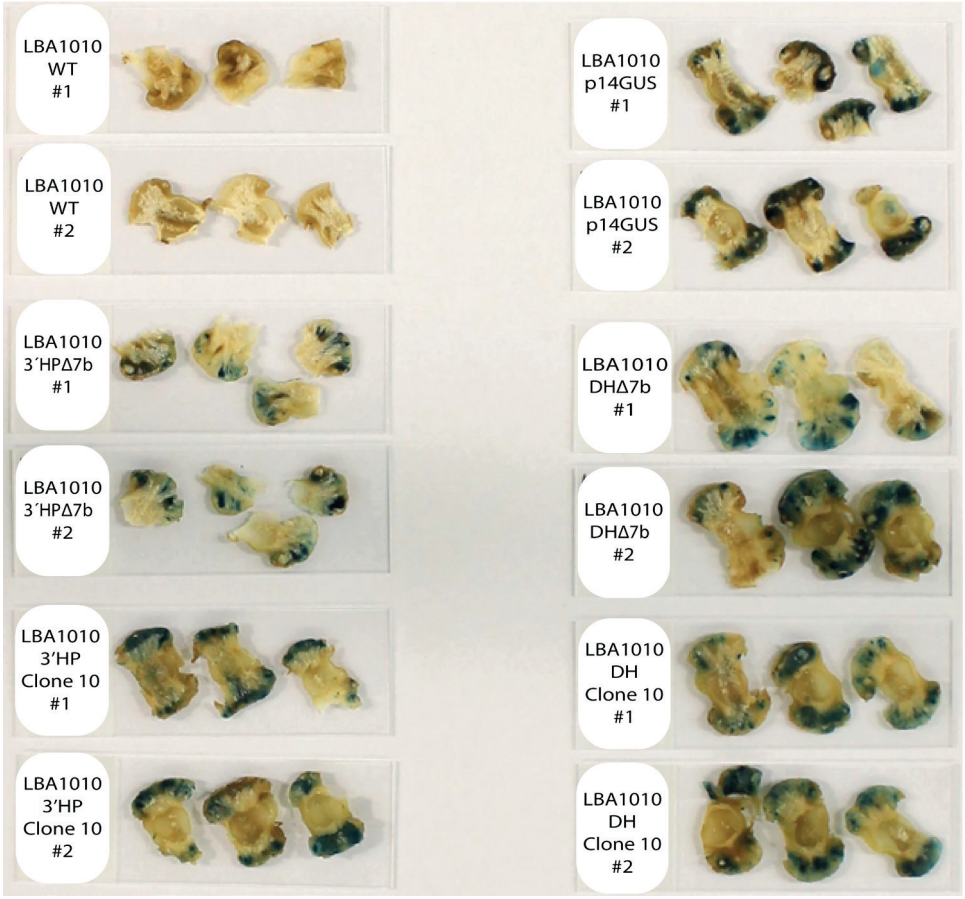


Figure 7. Long-term GUS expression assay in *N. glauca* tumors with HP T-DNA constructs. Tumors were induced by *A. tumefaciens* LBA1010 on *N. glauca* stems using varying HP constructs for co-transformation. As a control *A. tumefaciens* LBA1010 wild-type (containing no HP constructs) was used. Each construct was tested independently twice on different plants (n = 2). Sections of the sliced tumor depicted are representative for the entire tumor. The tumors were harvested 21 d after inoculation of stems with *A. tumefaciens*. GUS expression was visualized using X-Gluc staining. Plasmids used were p14P (p14GUS), p14P3'HP (3'HP), p14P3´HPΔ7b (3´HPΔ7b), p14PDH (DH) and p14PDHΔ7b (DHΔ7b). For 3'HP as well as DH, another sub-clone (number 10) without a mutation was used as a comparison.

HPs still require function of single-strand binding protein VirE2$_{Ti}$

The effector protein VirE2$_{Ti}$ is either essential (Dumas *et al.*, 2000), or at least important (Yusibov *et al.*, 1994) for translocation of the T-DNA/protein complex from donor to recipient plant cells. VirE2$_{Ti}$ is thought to bind to the entirety of the T-strand once inside the recipient cell, protecting the T-strand from exonucleolytic attacks, and is suspected to assist in nuclear import (Abu-Arish *et al.*, 2004). If the HP folding or double-strand formation process would occur prior to the transfer or directly after transfer before encountering endogenous nucleases, a VirE2$_{Ti}$ deletion would have no effect on the HP T-strand constructs. All HP T-DNA constructs could affect VirE2$_{Ti}$ protein binding, but also could counter the potential effect from endogenous plant nucleases similar to the function of VirE2$_{Ti}$. Another hypothetical

scenario would be that the HP insertion triggered a rapid double-strand conversion of the construct, immediately occurring after the T-strand displacement. In this case, after transfer of the double-stranded construct into the recipient cell, $VirE2_{Ti}$ would not be able to bind and endogenous nucleases would not be able to function. If no transformation of the recipient is detected with the $VirE2_{Ti}$ deletion mutant, this would infer that $VirE2_{Ti}$ is required for the unfolding of the single-stranded DNA, and the subsequent transport to the nucleus. The effect of HP constructs in a $VirE2_{Ti}$ deletion mutant was tested using *A. tumefaciens* LBA2573, lacking $VirE2_{Ti}$ on its pAL1100 helper plasmid. An agroinfiltration assay with *N. benthamiana* for 48 h, with the DHΔ7b construct as well as the p14 control, showed minute background relative GUS expression (**Figure 8**). *A. tumefaciens* LBA1100 with the p14 construct was used as a functional control, to validate the possibility of transformation of the used recipient cells. The control transformation exhibited high levels of GUS expression for this time point, demonstrating that a successful transformation was possible with the used experimental conditions (**Figure 8**).

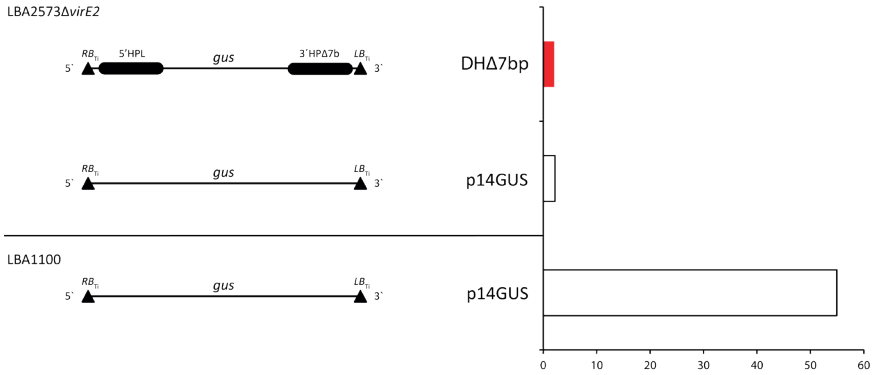


Figure 8. Influence of $VirE2_{Ti}$ on DHΔ7b HP mutant T-DNA transfer. GUS activity was determined 48 h after agroinfiltration of *N. benthamiana,* using *A. tumefaciens* LBA2573 lacking $VirE2_{Ti}$ and harboring the construct p14GUS (n = 6) without HP modification and the double-HP DHΔ7b (n = 6). Error bars are SEM and lie within the line of the graphs. GUS expression between the DHΔ7b construct and the p14GUS expression constructs was not significant using a t-test comparison ($t(10) = 1.417$, $p = 0.187$). As a functional control to validate the experimental protocol *A. tumefaciens* LBA1100 with construct p14GUS was used (n = 1). GUS expression was determined via a MUG assay. Plasmids used were p14P (p14GUS) and p14DHΔ7b (DHΔ7b).

The 3'DHΔ7b mutation has a negative effect on integration by homologous recombination

Analogous to the transformation in plants, the HP constructs were tested in *S. cerevisiae* to study the effect on T-DNA integration via HR. For this purpose, the p14 vectors were equipped with a *KanMX* yeast selection marker and left and right homology sequences of the *PDC6* locus of *S. cerevisiae* for integration via HR. When compared to the yeast p14Y vector, the single 5´-end 5'HP, and the single 3'-end 3´HPΔ7b, as well as the double-HP DHΔ7b, all triggered a measurable and statistically significant reduction in transformation efficiency after 48 h (**Figure 9**), indicating a negative effect on HR of the terminally modified T-DNA into the yeast genome. However, no difference between constructs equipped with different HP structures could be detected, when neither present individually nor combined. The negative effect on transformation efficiency is even visible in a short time window. Within 24 h or 48 h of co-cultivation, statistically significant reduced transformation efficiency was measured using the DHΔ7b construct (**Figure 10**).

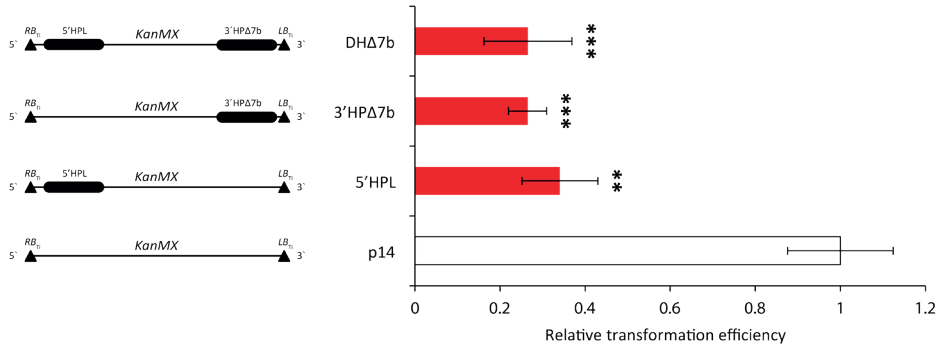


Figure 9. Effect of terminal HPs on transformation efficiency in *S. cerevisiae*. The tested constructs DHΔ7b (n = 10), 3`HPΔ4b (n = 10) and 5'HPL (n = 10) were normalized to the control vector p14 (n = 30) which was set as 1, showing the relative transformation efficiency after 48 h co-cultivation. The recipient strain used was *S. cerevisiae* YPH250 and transformation was interpreted as integration of double-stranded T-DNA via HR at the *PDC6* locus (**Figure 1**). The donor strain used was *A. tumefaciens* LBA1100. Error bars shown are SEM. The data was not normally distributed as assessed by Shapiro-Wilk test ($p < 0.05$) and there was heterogeneity of variance, as assessed by Levene's test ($p < 0.005$). A one-way Welch ANOVA showed a statistically significant difference within homologous recombination activity of the used HP constructs ($F(3, 24.069) = 9.910$, $p < 0.005$ with $\omega^2 = 0.29$). A *post hoc* Games-Howell test showed a statistical significant lower transformation efficiency of the constructs DHΔ7b ($p < 0.0005$), 3`HPΔ4b ($p < 0.0005$), 5'HPL ($p = 0.001$) compared to the p14 construct. There was no difference in relative transformation efficiency between the constructs DHΔ7b, 3`HPΔ4b and 5'HPL ($p > 0.05$). Plasmids used were p14Y (p14), p14Y5'HPL (5'HPL), p14Y3´HPΔ7b (3´HPΔ7b) and p14YDHΔ7b (DHΔ7b). ** for $p < 0.005$ and *** for $p < 0.0005$ compared to the p14 construct. The transformation efficiencies of this experiments can be found in **table S1A.**

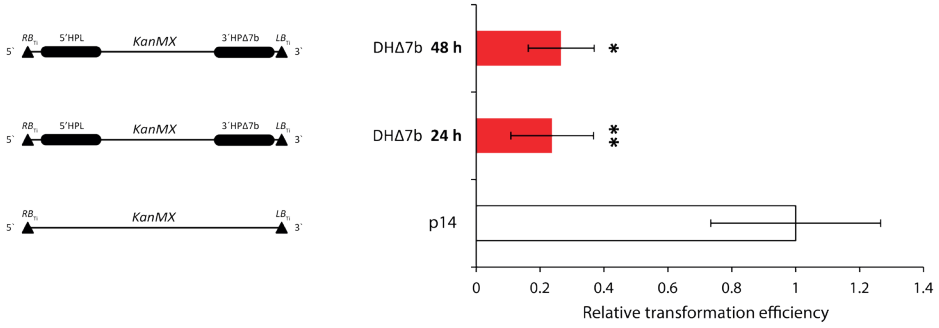


Figure 10. Immediate negative effect of HP construct DHΔ7b on transformation efficiency in *S. cerevisiae*. The negative effect on transformation efficiency is also detected within a timeframe of 24-48 h of co-cultivation with *S. cerevisiae* YPH250 and transformation was interpreted as integration of double-stranded T-DNA via HR at the *PDC6* locus (**Figure 1**). The donor strain used was *A. tumefaciens* LBA1100. The DHΔ7b constructs (both n = 10) were normalized to the transformation efficiency of p14 (n = 20) that was set as 1. Error bars depicted are SEM. The data was not normally distributed as assessed by Shapiro-Wilk test ($p > 0.05$). A Kruskal-Wallis H test showed that there was a statistically significant difference in relative transformation efficiency between the different constructs used, ($\chi 2(2) = 10.213$, $p = 0.006$), with a mean rank relative transformation score of 26.18 for the p14 construct, 12.85 for the DHΔ7b construct after 24 h and 16.80 for the DHΔ7b construct after 48 h. Pairwise comparisons were performed using Dunn's (1964) procedure with a Bonferroni correction for multiple comparisons. Adjusted p-values are presented. Values are mean ranks unless otherwise stated. This *post hoc* analysis showed statistically significant differences in relative transformation efficiency between the construct DHΔ7b after 24 h compared to p14 ($p = 0.003$) as well as after 48 h ($p = 0.036$). There was no difference between 24 h and 48 h compared to p14 ($p = 0.445$). Plasmids used were p14Y (p14), and p14YDHΔ7b (DHΔ7b). * for $p < 0.05$ and ** for $p < 0.005$ compared to the p14 construct. The transformation efficiencies of this experiments can be found in **table S1B.**

HPs still require the NHEJ pathway for genomic integration

To determine the influence on the integration of the HP constructs, the yeast strain *Kluyveromyces lactis* was used as recipient. The strain *K. lactis* exhibits a very low tendency for integration via HR of transformed DNA (Kooistra, Hooykaas & Steensma, 2004). Instead, genomic integration, in *K. lactis,* occurs increasingly via illegitimate integration via the NHEJ pathway, where no homology of the integrated DNA with genomic locus would be required (Kegel *et al.*, 2006). The *PDC6* homology region present on the T-DNAs has a high similarity (72%) to the same locus in *K. lactis;* as a result, some HR events can potentially occur. No transconjugants were detected when a *K. lactis* NHEJ mutant (Δ*ku80*) was transformed with p14 derived T-DNA containing HPs (**Table 1**).This showed the importance of the classical NHEJ pathway for T-DNA integration in yeast. When the wild-type *K. lactis* was transformed with T-DNA with HPs with or without the Δ7b mutation, no statistically significant difference in transformation efficiency compared to the p14Y baseline was observed (**Figure 11**). In addition, when the HR pathway was knocked out (*Δrad52*) in *K. lactis*, no statistically difference in transformation efficiency using the HP variants could be detected (**Figure 12**). The transformation efficiencies seen in the experiments with *K. lactis* wild-type and *Δrad52* can be found in **tables S1D** and **S1C**.

Table 1. Effect of HPs on transformation efficiency on *K. lactis* Δ*ku80* without non-homologous end joining. Duration of the co-cultivation time was 48 h with the recipient strain. The donor strain was *A. tumefaciens* LBA1100. Plasmids used were p14Y (p14), p14Y5'HPL (5'HPL), p14Y3'HP (3'HP), p14Y3´HPΔ7b (3´HPΔ7b), p14YDH (DH) and p14DHΔ7b (DHΔ7b).

***K. lactis* Δ*ku80* output:** 1.3×10^8	
Constructs used (n = 3)	**Transconjugants**
p14	0
5'HPL	0
3´HPΔ7b	0
DHΔ7b	0
3'HP	0
DH	0

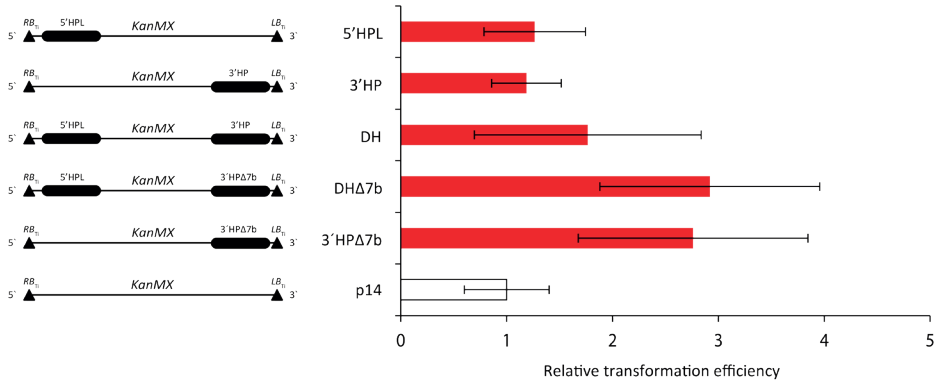


Figure 11. Effect of HP constructs on *K. lactis* wild-type (WT) transformation**.** Co-cultivation was performed for 48 h with the recipient strain. The donor strain was *A. tumefaciens* LBA1100. In this dataset, the HP constructs (n = 6) were normalized to the p14 base transformation efficiency (n = 6), which was set as 1. Error bars and margin are SEM. The data was not normally distributed as assessed by Shapiro-Wilk test ($p > 0.05$). A Kruskal-Wallis H test showed that there was no statistically significant difference in relative transformation efficiency between the different constructs used, ($\chi2(2) = 5.480$, $p = 0.360$). Plasmids used were p14Y (p14), p14Y5'HPL (5'HPL), p14Y3'HP (3'HP), p14Y3´HPΔ7b (3´HPΔ7b), p14YDH (DH) and p14DHΔ7b (DHΔ7b). The transformation efficiencies of this experiment can be found in **table S1C.**

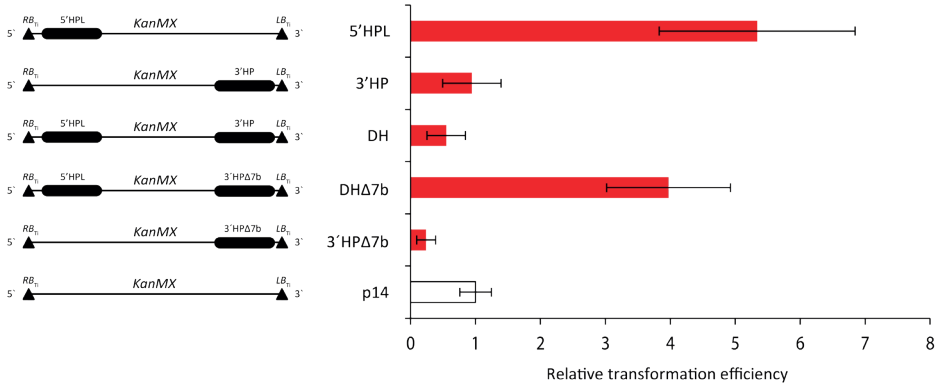


Figure 12. Effect of HPs on transformation efficiency on *K. lactis* Δ*rad52* without a HR pathway. Duration of co-cultivation with the recipient strain was 48 h. The donor strain was *A. tumefaciens* LBA1100. All HP structures (n = 3) have been normalized for the p14 base transformation efficiency (n = 3) that was set as 1. Error bars and margin provided are SEM. The data was not normally distributed as assessed by Shapiro-Wilk test ($p > 0.05$). A Kruskal-Wallis H test showed that there was no statistically significant difference in relative transformation efficiency between the different constructs used, ($\chi2(2) = 10.572$, $p = 0.061$). Plasmids used were p14Y (p14), p14Y5'HPL (5'HPL), p14Y3'HP (3'HP), p14Y3´HPΔ7b (3´HPΔ7b), p14YDH (DH) and p14DHΔ7b (DHΔ7b). The transformation efficiencies of this experiment can be found in **table S1D.**

Using an artificial $tRNA_i^{met}$-priming site does not induce increased transformation efficiency in yeast

An endogenous tRNA-binding site was introduced on the terminus of the T-strand to test the possible priming of a tRNA molecule to induce double-strand formation via the host endogenous DNA replication machinery. The $tRNA_i^{met}$-binding site was introduced on the 3´-end of the p14 T-DNA, having homology to the endogenous yeast $tRNA_i^{met}$. This modification was introduced to induce double-strand formation in the recipient cell. When the T-DNA was transferred via AMT to *S. cerevisiae* YPH250, the construct exhibited no change in transformation efficiency compared to the baseline transformation efficiency of the control p14Y construct without terminal modifications (**Figure 13**). The assay was not possible in plants, as the sensitive GUS expression assay using agroinfiltration was only possible with *N. benthamiana*, and the tRNA sequence for this species was not available at the time of this study. A tRNA binding site for *Arabidopsis thaliana* was created, but could not be used, as agroinfiltration into its leaves was not possible.

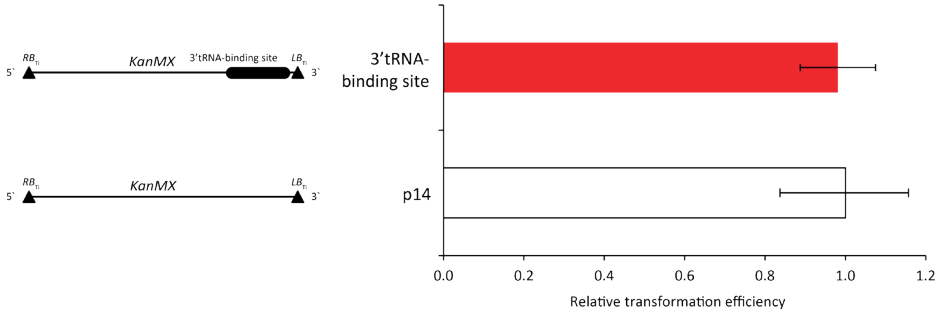


Figure 13. Effect of 3'tRNA-binding site at the 3´-end on transformation efficiency of *S. cerevisiae*. The tested construct containing the 3'tRNA-binding site (n = 10) was normalized to the control vector p14 (n = 10), set as 1, showing the relative transformation efficiency after 48 h of co-cultivation. The recipient strain used was *S. cerevisiae* YPH250 and transformation occurred via HR at the *PDC6* locus. The donor strain was *A. tumefaciens* LBA1100. Error bar depicted is SEM. There was no statistically significant difference between the 3'tRNA construct and the control construct p14 using a t-test comparison ($t(18) = 0.018$, $p = 0.916$). Plasmids used were p14Y (p14) and p14YSctRNA (3'tRNA-binding site). The transformation efficiencies of this experiment can be found in **table S1E.**

Discussion

Enhancing the transient *Agrobacterium*-mediated transformation of plants and other organisms could be a stepping-stone for future transformation technologies. The improvement of transformation was explored by the physical modification of T-DNA to facilitate priming of double-strand T-DNA formation in recipient cells, either for more rapid T-DNA expression in the recipient cell, and/or to provide for T-DNA molecules that are refractory to integration into the recipient genome still (temporarily) expressing the encoded genes. In this study, HP structures were tested to provide a 'built-in' priming site, to be generated when the single-stranded T-DNA end would be folded back on itself. In order to facilitate double-strand formation prior to loosing genetic information, the 3´-terminus of the T-DNA is the ideal initiating site to be employed by the host endogenous double-strand formation machinery (5´-3´ synthesis). A plant system was used to study the effect on transient expression, and a yeast system to see the effect on integration.

When screening different sub-clones of the 3´-HP structure (3'HP), one clone (3´HPΔ7b) exhibited a strong increase in transient GUS expression (**Figure 3**). Sequencing revealed a deletion that could create a 4 base free end of the HP structure when folded and paired with its inverted repeats (**Figure 2**). As the sequencing of the 3´HPΔ7b construct in this study was incomplete, possibly due to the HP structure, the final sequence was derived as consensus sequence of multiple reads of the construct (**Figures 2B, S2A and S2B**). It could be attainable that the actual mutation could be slightly different and was not detected due to the problematic sequencing results. The increased transient expression effect of the 3´HPΔ7b continued even when combined with a 5´-HP as a double-HP construct (**Figure 4**). Both constructs, containing the HP mutations, showed a significant difference in comparison to the 5'-HP and the control construct (**Figure 5**). The effect peaked around 48-72 h, but showed a steady long-term level (**Figure 6**). Also after 21 d (in tumors), no reduction of expression activity could be detected. In tumors, blue staining at an early time point would indicate transient expression from non-integrated T-DNA, which would lead to no visible staining at a later time point, as the T-DNA would be lost. Whereas a positive staining at a much later time point is suggestive of integration of the T-DNA with subsequent expression (**Figure 7**). HP structures showed similar transient transformation as T-DNA without HPs. In general, T-DNA with HP structures did not influence the T-DNA processing, transport and expression. Also in yeast, integration, via the HR pathway, occurred at a similar frequency as with T-DNA structures with perfect HPs (**Figure 9**). This further leads to the conclusion that the double-HP construct, if rendered double-stranded, is not able to withdraw itself from the integration pathways into the recipient genome, by potentially masking its free ends. Theoretically, polymerase Pol-θ can process a potential overhang derived from the 3´HPΔ7b double-stranded T-DNA construct before integration into the plant genome (van Kregten *et al.*, 2016). Double-strand formation is measured indirectly by determining the activity of the expressed GUS protein. The *gusA* sequence on the displaced and transferred T-strand was in such an orientation that it is non-coding as single-strand. Only after double-strand formation, a coding strand would be available for GUS protein expression. It was also very difficult to discern between transient expression and expression after integration in plant transformation experiments. It is still not clear if the initially transiently expressed GUS proteins are still detected after several days (**Figures 6 and 7**), as the measured relative GUS expression levels stayed the same after 48 h. Very little is known about the long-term stability of the GUS protein. Even if the original vector would be no longer present in the cell, active stable proteins could be retained and their activity picked up by the sensitive MUG assay. The half-life of GUS protein in plant cells was estimated to be around 3 d (Jefferson *et al.*, 1989) or even up to 14 d (Oliver, 1994), but no clear consensus or more elaborate data exists. All HPs seemed to have a negative effect on genomic integration using HR in yeast, by showing a decreased number of stably transformed yeast cells (**Figure 9**). When the HP constructs were compared

in a strain with very little HR activity, no discernable effect was detected (**Figure 11**). This could be due to folding of the hairpin structures influencing the end processing during HR. Similar, as with the experiments in plants, an effect was detected directly after 24 h at a 48 h time-point (**Figure 10**). It further was shown that the classical NHEJ pathway is crucial for integration in yeast (**Table 1**), which is in line with previous findings (van Attikum, Bundock & Hooykaas, 2001). This suggests the possibility, that in plants TMEJ mediated by polymerase Pol-θ is involved in the processing of the 3´HPΔ7b open single-stranded end, possibly having an immediate effect on double-strand formation of the construct.

It is possible that the HP structures have an influence on the transfer via the $T4SS_{Ti}$ itself. It was assumed that if T-strands outfitted with aberrant terminal structures could transfer at a normal transformation efficiency, the 3'-HP structure could be instrumental in aiding the double-strand formation. Normal transformation efficiency would lead to the conclusion that the HP structures would not hamper the initial Dtr_{Ti} DNA processing or the CP_{Ti}/$T4SS_{Ti}$ transfer. The double-strand formation of the construct would occur in the nucleus of the recipient cell. Experiments showed that the perfect HP structures (**Figures 3 and 4**) did not have an effect on the measured protein expression. This indirect measurement could also indicate that the transfer rate could be reduced, hampered by the HP structure, but the constructs that were successfully transferred could have an increased double-strand formation effect, in which the HP structures advantage and disadvantage would balance itself out. On the other hand, the 3´HPΔ7b structure showed increased protein expression, leading to the conclusion that there is an option to overcome this theoretical balancing act between transfer and double-strand formation.

Based on the hypothesis that $VirE2_{Ti}$ binds to the T-strand within the recipient cell prior to double-strand formation, the following conclusions could be made based on experiments conducted (**Figure 8**). Apparently, the DH construct was not able to evade nucleolytic attacks within the recipient cell and nucleus when transferred, when tested with a deletion mutant for $VirE2_{Ti}$ (**Figure 8**). This finding could mean that open ends exist in this construct and would be exposed to degradation processes that theoretically could be prevented by the association of $VirE2_{Ti}$ to the T-strand. $VirE2_{Ti}$ seems to be essential for the transfer to plants, and this most likely would be not due the lack of $VirE2_{Ti}$ disrupting the $T4SS_{Ti}$-mediated transfer, but by potentially not protecting the T-strands from degradation of plant specific nucleases. Nuclear import of the T-strand could be still regulated by the NLS contained in the bound $VirD2_{Ti}$ molecule. Evidence from studies with $VirE2_{Ti}$ deletions looking at the transfer to *S. cerevisiae* showed that $VirE2_{Ti}$ was not essential for transformation (Bundock *et al.*, 1995), as it also plays no role in the intra-bacterial transfer (Beijersbergen *et al.*, 1992). In conclusion, this could indicate that the HP-folding does not occur during the transfer or in the plant cytosol, but the theoretical folding would only be possible within the nucleus after $VirD2_{Ti}$ detachment, as a HP folding earlier could prevent degradation of the T-strand.

Considering the data generated in this work, it could be hypothesized that the 3´HPΔ7b modification (singly or in combination with the 5'-end HP) could impact strand processing or end processing within the recipient cell. In comparison, no difference in expression was detected between the perfect 3'HP and the LB_{Ti} sequence without modification, and hence no effect on actual T-strand priming for faster double-strand processing. Also using a tRNA-binding site to induce T-strand priming for subsequent double-strand formation, did not show an increase in integration. HPs though could have a negative effect on HR, as the folded structure could mask the ends for single-stranded insertion into the double-strand break. This leads to the conclusion that the free 4 b at the 3´-end of the T-strand end induces a quicker end-processing, removing the construct from the HR pathway. This study describes for the first time that it is possible to positively influence the transformation

frequency by just modifying the substrate for AMT. A more general conclusion would be that HP structures are processed and transported apparently without inherent problems, but could have an effect on homologous integration. The finding of this research of the specific terminal modification mutant combined with the recent insight into the Pol-θ TMEJ pathway, could be an opportunity for further study and determine the actual *in vivo* ratio of transient expression and integration of the construct in the future. The ratio could be determined by testing expression of the mutant construct in *A. thaliana* plants that lack polymerase Pol-θ and comparing these to wild-type expression levels.

Material and methods

Organisms and constructs

All strains used in this experimental set-up are listed in **table 2**, and all constructs used for this study are provided in **table 3**. Insertions of constructs were performed as follows: *Escherichia coli* DH5α were transformed by heat shock (Inoue *et al.* 1990) and *A. tumefaciens* either by electroporation (den Dulk-Ras & Hooykaas, 1995) or by tri-parental mating using *E. coli* CEL40/HB101 (pRK2013) (Ditta *et al.*, 1980). All *E. coli* and *A. tumefaciens* strains used in this chapter were cultured on appropriate antibiotic selection (**Tables 2 and 3**) on LB agar or in LB liquid medium (10 g/L bacto tryptone, 5 g/L bacto yeast extract, 8 g/L NaCl, pH 7.0). *E. coli* was cultured at 37° C, *A. tumefaciens* at 29° C. The yeast strain *S. cerevisiae* was cultured with YPD (Zonneveld, 1986) agar or liquid medium and the strain *K. lactis* was cultured with MY (Zonneveld, 1986) agar or liquid medium with the appropriate selection at 29° C. As plant models, *Nicotiana benthamiana* and *Nicotiana glauca* were used and grown with 75% relative humidity at a 16 h photoperiod at 25° C.

Molecular cloning and sequencing

Standard molecular cloning procedures were performed according to standard laboratory methods (Sambrook & Russell, 2001). Plasmids were contained and amplified in *E. coli* DH5α and DH10β. Constructs were send for commercial automated high-throughput sequencing (Sanger sequencing) to Macrogen Europe Laboratories (The Netherlands).

Cloning of the T-DNA modification parental plasmid p14Y (p14PDCKXSS) precursor to all p14 HP plasmids

The plasmid p14Y (p14PDCKXSS) was used as the origin for all T-DNA double-strand formation inducing modifications. This is the precursor plasmid for all other p14 plasmid used in this study. The plasmid contains an octopine RB_{Ti} with and an octopine OD_{Ti}, and an octopine LB_{Ti}. Located between the border sequences are a *KanMX* cassette flanked by two *PDC6* locus homologous sequences to ensure integration into an *S. cerevisiae* recipient. The plasmid was based on the plasmid p14PDC6KX (Rolloos *et al.*, 2013), that was modified from the binary helper plasmid pSDM14 (Offringa, 1992) using a pBIN19 backbone (Bevan, 1984). The KanMX cassette was cloned from plasmid pSDM8000 containing a *TEF1* promoter (Rolloos *et al.*, 2013). To create a multipurpose platform to use in a different organism, an exchangeable marker cassette was inserted in between the *PDC6* homology regions. To allow convenient cloning, a second SalI site upstream of the *KanMX* site was removed, as the second SalI site (located at the LB_{Ti}) was used for later oligonucleotide insertion. First, the construct was digested with XhoI and AscI. The smaller 814 bp fragment containing the left flank of the *PDC6* homology sequence was isolated and saved for later re-insertion. The large fragment was ligated with the double-stranded oligonucleotides SXA FW and SXA RV, containing a new cloning site with the restriction sites XhoI-BamHI-NheI-AscI (**Table 4**), utilizing an oligo insertion method according to Neuteboom *et al.* (2006). After SXA oligo insertion, the backbone was digested with XhoI. The saved 814 bp fragment was digested with SalI and then ligated into the XhoI digested backbone, creating the plasmid p14Y (p14PDCKXSS).

Insertion of T-DNA modifications and creation of the p14 plasmid line

The parental plasmid p14Y (p14PDCKXSS) was used for all further insertions. First the *A. thaliana* specific tRNA-binding site (AttRNA) or the *S. cerevisiae*-specific tRNA-binding site (SctRNA), was inserted upstream of the LB_{Ti} at the 3´-end of the T-DNA as double-stranded oligonucleotides (Neuteboom *et al.*, 2005) into the SalI digested p14PDCKXSS. It is to note that that the AttRNA binding site was not experimentally tested for its function. Oligonucleotides Sc tRNA FW and Sc tRNA RV were used for the SctRNA-binding site and oligonucleotides At tRNA FW, and At tRNA RV were used for AttRNA-binding site (**Table 4**). The following plasmids were created, p14PDCKXSSAttRNA (with direct- and indirect-versions of the tRNA-binding site) and p14PDCKXSSSctRNA (with direct- and indirect-versions of the tRNA-binding site). With the direct-version, as seen in a relative view of the inserted oligonucleotides as the insertion could be in both directions due the use of only one restriction site, the actual tRNA-binding site was present on the bottom strand in direction RB_{Ti} to LB_{Ti} on 5´ to 3´. The bottom strand in the direct-version is not displaced via the relaxase. Only the top strand of the direct-version is transferred, but does not allow for the annealing of the tRNA. For the indirect-version, the actual binding site for the tRNA was on the top strand in direction RB_{Ti} to LB_{Ti} on 5´to 3´. The T-DNA that is displaced and transferred via the relaxase is, in this case, the top strand of the plasmid with the indirect-oligonucleotide insertion. This means, the transferred T-DNA single-strand has the correct complementary bases for the tRNA to anneal, creating a 5'-3' direction tRNA-primer for initiation of double-strand synthesis. For experiments and further cloning, the indirect-version allowing for tRNA-mediated priming was used. The direction was verified by sequencing and restriction analysis.

To insert a HP at the RB_{Ti}, at the 5´-end of the T-DNA, p14Y (p14PDCKXSS) was digested with XhoI. The double-stranded oligonucleotides 5'HPL and 5'HPS were then inserted individually (Neuteboom *et al.*, 2006) creating the plasmids p14PDCKXSS5'HPL and p14PDCKXSS5'HPS. Both plasmids were created in direct- and indirect-versions. With the direct-version, the 5´-end HP insertion was present on the displaced T-DNA strand. These versions were used for experiments (**Figure 1**). The direction was verified by sequencing and restriction analysis. The plasmids were renamed into p14Y5'HPL and p14Y5'HPS.

The 3´-end HP modifications 3'HP was inserted as double-strand oligo in SalI digested p14PDCKXSSAttRNA (indirect-version), creating the plasmid p14PDCKXSS3'HP. The 3'HP insertion was also present as a direct- and indirect-version. With the indirect-version, the 3'HP was located on the displaced T-DNA strand in the correct alignment for the HP to fold back once the LB_{Ti} was nicked. For experiments, the 3'HP indirect-version was used (**Figure 1**). This plasmid was renamed into p14Y3'HP.

To form the double-HP construct (p14DH), a NotI/PstI fragment containing the 5'-HPL of p14PDCKXSS5'HPL direct construct was ligated with the PstI fragment of the p14PDCKXSS3'HP indirect-construct. The resulting plasmid was renamed into p14YDH. All HP variations constructs were verified by restriction analysis.

Insertion of *gusA* expression cassette into HP plasmids

A *gus* (*gusA*) expression cassette under control of the *CaMV 35S* promoter was inserted into the HP containing p14Y plasmid family used for plant transformation experiments in exchange of the *KanMX* cassette The *KanMX* cassette was removed using NheI and SpeI. The *gusA* cassette was amplified from the plasmid pCAMBIA2301 (GenBank: AF234316.1) with primer FTW75 and BZ1 and ligated into plasmid pJet1.2. The *gusA* cassette was removed with

SpeI and BsaI. The plasmid backbone and *gusA* were then ligated to create HP plasmids with an exchange selection cassette (**Table 3**). For the double-HP T-DNA displacement strand, the direction of the *gusA* insert was of matter. After insertion, the *gusA* direction was chosen in such way that the displaced T-DNA was non-coding and could only be used for transcription after double-strand formation. To indicate the presence of the *gus* cassette in the created plasmid family, a P (for plant selection) was added to the plasmid. This generated the plasmids series p14Pxxx (xxx = different designation according to insertion, see plasmid list in **table 3**).

Agroinfiltration and tumor formation assay

To physically insert *A. tumefaciens* into plant leaves, the following agroinfiltration protocol was devised, based on Sakalis (2013). The *A. tumefaciens* strains were prepared, pre-cultured and induced according to the standard method (Bundock *et al.*, 1995). Pre-inducing was initiated 55 h prior to the agroinfiltration within 10 mL induction medium (IM) containing 200 µM acetosyringone (AS) per independent culture at 20°C. After the pre-induction, *Agrobacterium* was prepared into a volume of IM with 200 µM AS with a final concentration of OD_{600} of 1. This solution was then injected using a 1 mL syringe into the underside of 3 weeks old *N. benthamiana* leaves. The injected *N. benthamiana* plants were cultured until the time point of analysis with 75% relative humidity at a 16 h photoperiod at 25° C covered with plastic foil.

The tumor formation assay on *N. glauca* was performed according to the protocol of Schrammeijer, Hemelaar & Hooykaas (1998).

Co-cultivation

The *A. tumefaciens* and yeast strains were prepared, pre-cultured and induced according to the standard method of Bundock *et al.* (1995). The co-cultivation method was then modified from Bundock *et al.* (1995). *A. tumefaciens* was inoculated in 10 mL LC and cultured 14-16 h at 29° C, and then a volume containing a total OD_{600} of 0.25 was transferred into 5 ml IM with 200 µM acetosyringone (AS) and grown at 28° C for 5.5 h. *S. cerevisiae* was inoculated in 50 mL yeast extract-peptone-dextrose (YPD) liquid medium and grown for 14-16 h at 29° C, and then 5 ml was transferred into 45 ml fresh YPD and grown 30° C for 5.5 h. After the 5.5 h growth, the concentration of colony forming units (CFU) per mL of *A. tumefaciens* and *S. cerevisiae* was determined. For optimal co-cultivation a final amount of 1×10^7 CFU of *S. cerevisiae* was mixed with 2×10^8 CFU of *A. tumefaciens*. Both amounts were mixed together, pelleted and re-suspended in 100 µL IM containing 200 µM AS. The 100 µL co-cultivation cell suspension was then equally distributed on one well, of a 34.8 mm 6-well plate, containing 3 mL solid (prior dried) IM with 200 µM AS. After addition of the cell suspension to the surface of the IM medium, the well was sufficiently dried. After the incubation until the time point of analysis at 21° C, the biofilm was removed with 1 mL of 0.9% NaCl solution from the well with rigorous pipetting until no biofilm was left on the medium surface. The suspension was then pelleted, re-suspended in 500 µL liquid YPD and then plated on the YPD selection plates for analysis containing geneticin (G418; 150 µg/ml) and cefotaxamine (200 µg/mL). To determine the output ratio of *S. cerevisiae*, 10 µL of the resuspension was used for a serial dilution that was plated on YPD solid medium containing cefotaxamine (200 µg/mL). The final transformation efficiency was determined by dividing the output transconjugants CFU by the output ratio of yeast CFU.

X-Gluc staining and MUG analysis

The GUS protein detection methods were performed according to Jefferson, Kavanagh & Becan (1987). A modified protein extraction buffer was used for the MUG assay: 0.5 M Na_2HPO_4/50 mM EDTA (pH = 7.0), 0.1% SDS, 0.1% Triton X-100 and 10 mM β-mercaptoethanol. *N. benthamiana* leaf discs (6 mm diameter) were stenciled out along the middle vein moving from leaf tip towards the bottom, three on each side. Per leaf, 6 leaf discs were collected. For protein isolation, leaf discs were shock frozen in liquid nitrogen and homogenized.

Table 2. Organisms used in this study.

Strain	Description	Resistance	Reference
A. tumefaciens			
LBA202	C58 cured of pTiC58	none	Hooykaas *et al.*, 1980
LBA1010	LBA202; Rif^R; Nal^R; pTiB6	Rif, Nal	Koekman, Hooykaas & Schilperoort, 1982
LBA1100	LBA202; Rif^R; Nal^R; pAL1100 or octopine pTiB6: ΔT_L, ΔT_R, Δ*tra*, Δ*occ*	Rif, Nal, Spec	Beijersbergen *et al.*, 1992
LBA2210	LBA202; Rif^R; Nal^R; RP4	Rif, Nal, Cb, Km, Tc	PJJ Hooykaas (unpublished)
LBA2573	LBA1100; pAL1100Δ*virE2*	Rif, Nal, Spec	Hodges *et al.,* 2006
W			
DH5α	*F-*; *Φ80lacZΔM15*; *Δ(lacZYA-argF); U169*; *recA1*; *endA1*; *hsdR17 (rK–, mK+)*; *phoA*; *supE44*; *λ– thi-1*; *gyrA96*; *relA1*		Grant *et al.*, 1990
DH10β	*F-*; *araDJ39; A(ara, leu)7697*; *AlacX74*; *galU*; *galK*; *rpsL*; *deoR*; *Φ80dlacZAM15*; *endAI*; *nupG*; *recA1*; *mcrA*; *A(mrr hsdRMS*, *mcrBC)*		Grant *et al.*, 1990
CEL40	HB101 (pRK2013)		Ditta *et al.*, 1980
N. benthamiana			
WT	Seeds provided by HJM Linthorst	n.a.	Bombarely *et al.*, 2012
S. cerevisiae			
YPH250	*MATa*; *ade2-101^{ochre}*; *his3-Δ200*; *leu2-Δ1*; *lys2-801^{amber}*; *trp1-Δ1*; *ura3-52*	n.a.	Sikorski and Hieter, 1989
K. lactis			
WT	WT CBS2359	n.a.	Kooistra, Hooykaas & Steensma, 2004
Δ*rad52*	GG1634; *KIRAD52*	n.a.	Kooistra, Hooykaas & Steensma, 2004
Δ*ku80*	GG1632; *KlKu80*	n.a.	Kooistra, Hooykaas & Steensma, 2004

n.a.: not applicable; Cb: Carbenicillin; Km: Kanamycin; Nal: Nalidixic acid; Rif: Rifampicin; Spec: Spectinomycin; Tc: Tetracycline; WT: Wild-type

Table 3. Genetic constructs used in this study.

Plasmid	*oriV*	Description	Resistance	Reference
p14-PDC6KX	*IncP*	Parental plasmid; *KanMX* (*TEF1* promoter)	Km	Rolloos *et al.*, 2013
p14Y	*IncP*	Lab name: p14PDCKXSS. Removal of second XhoI site with only one site left for cloning; $OD_{Ti,Oct}$; $RB_{Ti,oct}$; $LB_{Ti,oct}$; *PDC6, KanMX*	Km	This study
p14YAttRNA	*IncP*	*A. thaliana 3*'tRNA-binding site; *KanMX*	Km	This study
p14YSctRNA	*IncP*	*S. cerevisiae* 3'tRNA-binding site; *KanMX*	Km	This study
p14Y5'HPL	*IncP*	5′ long HP; *KanMX*	Km	This study
p14Y3'HP	*IncP*	3′ HP insertion; *KanMX*	Km	This study
p14Y3′HPΔ7b	*IncP*	Mutation 3′ HP insertion; *KanMX*	Km	This study
p14YDH	*IncP*	5′ and 3′ double-HP; *KanMX*	Km	This study
p14YDHΔ7b	*IncP*	Double-HP; 5′ HP and mutation 3′ HP; *KanMX*	Km	This study
p14P	*IncP*	$RB_{Ti,oct}$; $LB_{Ti,oct}$; *PDC6*; *gusA* (*CaMV 35S* promoter)	Km	This study
p14P5'HPL	*IncP*	5′ long HP; *gusA*	Km	This study
p14P5'HPS	*IncP*	5′ short HP; *gusA*	Km	This study
p14P3'HP	*IncP*	3′ HP insertion; *gusA*	Km	This study
p14P3′HPΔ7b	*IncP*	Mutation 3′ HP insertion; *gusA*	Km	This study
p14PYDH	*IncP*	5′ and 3′ double-HP; *gusA*	Km	This study
p14PDHΔ7b	*IncP*	Double-HP, 5′ HP and mutation 3′ HP; *gusA*	Km	This study
pCAMBIA2301	*pBR322, pVS1*	*gusA*; *nptII*	Km	GenBank AF234316.1

oriV: Origin of vegetative replication (synonymously with origin of replication *oriRep* or *ori*); Km: Kanamycin; $RB_{Ti,Oct}$: Right Border pTi_{Oct}; $LB_{Ti,Oct}$: Left Border pTi_{Oct}; $OD_{Ti,Oct}$: Overdrive pTi_{Oct} (Rolloos *et al.*, 2013)

Table 4. Oligonucleotide sequences used in this study.

SXA FW	TCGAGGATCCGCTAGCAAGG
SXA RV	CGCGCCTTGCTAGCGGATCC
Sc tRNA FW	TCGACGCATGCTTAATTAACGGATCGAAACCGAGCGGCGCTACCAATTTA-AATACCGGTC
Sc tRNA RV	TCGAGACCGGTATTTAAATTGGTAGCGCCGCTCGGTTTCGATCCGTTAAT-TAAGCATGCG
At tRNA FW	TCGACGCATGCTTAATTAAAGGATCGAAACCTGGCTCTGATACCAATTTA-AATACCGGTC
At tRNA RV	TCGAGACCGGTATTTAAATTGGTATCAGAGCCAGGTTTCGATCCTTTAATTA-AGCATGCG
5'HPL FW	TCGACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCCTCATA-AGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTCGAGGCC
5'HPL RV	TCGAGGCCTCGACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATG-GCTTCTTATGAGGAAGCCATACCAAACGACGAGCGTGACACCACGATG-CCTG
5'HPS FW	TCGACAGGCATCTCGTCGTTTGGTATGGCTTCCTCATAAGAAGCCATAC-CAAACGACGAGATGCCTGTCGAGGCC
5'HPS RV	TCGAGGCCTCGACAGGCATCTCGTCGTTTGGTATGGCTTCTTATGAGGAAG-CCATACCAAACGACGAGATGCCTG
3'HP FW	CCGGTGAATTCTTAATTAAAGGATATATTCAATTGTAAATCTCGAAGTCGAC-GCATGCCTCATAAGCATG
3'HP RV	CTTATGAGGCATGCGTCGACTTCGAGATTTACAATTGAATATATCCTTTAAT-TAAGAATTCA
FTW75	TCACTAGTGCAGCTTGAGCTTGGAT
BZ1	CCACTAGTATCCTGTCAAACACTGATAGT
HpinExOut	TTGCAAAAACCGAATGGC

All oligonucleotides are depicted in 5´-3´direction.

Acknowledgements W

M Rolloos (IBL) helped with the *K. lactis* co-cultivation experiments. HJM Linthorst (IBL) provided the *N. benthamiana* seeds. GM de Winter (IBL) and MCF Lavrijsen (IBL) prepared media and provided glass-ware sterilization. JM Vink assisted with growing of tobacco plants. This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Abu-Arish A, Frenkiel-Krispin D, Fricke T, Tzfira T, Citovsky V, Wolf SG & Elbaum M (2004) Three-dimensional reconstruction of *Agrobacterium* VirE2 protein with single-stranded DNA. *J. Biol. Chem.* **279:** 25359–25363

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

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Bevan M (1984) Binary *Agrobacterium* vectors for plant transformation. *Nucleic Acids Res.* **12:** 8711–8721

Bikard D, Loot C, Baharoglu Z & Mazel D (2010) Folded DNA in action: hairpin formation and biological functions in prokaryotes. *Microbiol. Mol. Biol. Rev.* **74:** 570–588

Bombarely A, Rosli HG, Vrebalov J, Moffett P, Mueller LA & Martin GB (2012) A draft genome sequence of *Nicotiana benthamiana* to enhance molecular plant-microbe biology research. *Mol. Plant-Microbe Interact.* **25:** 1523–1530

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

Burnett BP & McHenry CS (1997) Post-transcriptional modification of retroviral primers is required for late stages of DNA replication. *Proc. Natl. Acad. Sci. U.S.A.* **94:** 7210–7215

Chilton M-DM & 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 Physiol.* **133:** 956–965

de Groot MJA, Bundock P, Hooykaas PJJ & Beijersbergen AGM (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* **16:** 839–842

den Dulk-Ras AD & Hooykaas PJJ (1995) Electroporation of *Agrobacterium tumefaciens*. *Methods Mol. Biol.* **55:** 63–72

Ditta G, Stanfield S, Corbin D & Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **77:** 7347–7351

Dumas F, Duckely M, Pelczar P, van Gelder P & Hohn B (2001) An *Agrobacterium* VirE2 channel for transferred-DNA transport into plant cells. *Proc. Natl. Acad. Sci. U.S.A.* **98:** 485–490

Dunn OJ (1964) Multiple comparisons using rank sums. *Technometrics* **6:** 241–252

Dürrenberger F, Crameri A, Hohn B & Koukolíková-Nicola Z (1989) Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 9154–9158

Eckardt NA (2004) Host proteins guide *Agrobacterium*-mediated plant transformation. *Plant Cell* **16:** 2837–2839

Friant S, Heyman T, Byström AS, Wilhelm M & Wilhelm FX (1998) Interactions between Ty1 retrotransposon RNA and the T and D regions of the tRNA(iMet) primer are required for initiation of reverse transcription in vivo. *Mol. Cell. Biol.* **18:** 799–806

Friant S, Heyman T, Wilhelm ML & Wilhelm FX (1996) Extended interactions between the primer tRNAi(Met) and genomic RNA of the yeast Ty1 retrotransposon. *Nucleic Acids Res.* **24:** 441–449

Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the 'gene-jockeying' tool. *Microbiol. Mol. Biol. Rev.* **67:** 16–37

Grant SG, Jessee J, Bloom FR & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 4645–4649

Herrera-Estrella A, Chen ZM, van Montagu M & 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. *EMBO J.* **7:** 4055–4062

Hodges LD, Vergunst AC, Neal-McKinney J, den Dulk-Ras AD, Moyer DM, Hooykaas PJJ & Ream W (2006) *Agrobacterium rhizogenes* GALLS protein contains domains for ATP binding, nuclear localization, and type IV secretion. *J. Bacteriol.* **188:** 8222–8230

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Howard E & Citovsky V (1990) The emerging structure of the *Agrobacterium* T-DNA transfer complex. *BioEssays* **12:** 103–108

Inoue H, Nojima H & Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:** 23–28

Janssen BJ & Gardner RC (1989) Localized transient expression of GUS in leaf discs following cocultivation with *Agrobacterium*. *Plant Mol. Biol.* **14:** 61–72

Jefferson RA, Kavanagh TA & Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6:** 3901–3907

Kapila J, De Rycke R, van Montagu M & Angenon G (1997) An *Agrobacterium*-mediated transient gene expression system for intact leaves. *Plant Sci.* **122:** 101–108

Kegel A, Martinez P, Carter SD & Astrom SU (2006) Genome wide distribution of illegitimate recombination events in *Kluyveromyces lactis*. *Nucleic Acids Res.* **34:** 1633–1645

Kleinboelting N, Huep G, Appelhagen I, Viehoever P, Li Y & Weisshaar B (2015) The structural features of thousands of T-DNA insertion sites are consistent with a double-strand break repair-based insertion mechanism. *Mol. Plant* **8:** 1651–1664

Koekman BP, Hooykaas PJJ & Schilperoort RA (1982) A functional map of the replicator region of the octopine Ti plasmid. *Plasmid* **7:** 119–132

Kooistra R, Hooykaas PJJ & Steensma HY (2004) Efficient gene targeting in *Kluyveromyces lactis*. *Yeast* **21:** 781–792

Liang Z & Tzfira T (2013) *In vivo* formation of double-stranded T-DNA molecules by T-strand priming. *Nat. Commun.* **4:** 1–8

Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD & Lopez R (2019) The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47:** W636–W641

Menassa R, Ahmad A & Joensuu JJ (2012) Transient expression using agroinfiltration and its applications in molecular farming. In *Molecular Farming in Plants: Recent Advances and Future Prospects*, Wang A & Ma S (eds) pp 183–198. Dordrecht: Springer Netherlands

Narasimhulu SB, Deng XB, Sarria R & Gelvin SB (1996) Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *Plant Cell* **8:** 873–886

Neuteboom LW, Lindhout BI, Saman IL, Hooykaas PJJ & van der Zaal BJ (2006) Effects of different zinc finger transcription factors on genomic targets. *Biochem. Biophys. Res. Commun.* **339:** 263–270

Niu X, Zhou M, Henkel CV, van Heusden GPH & Hooykaas PJJ (2015) The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene *VBF*. *Plant J.* **84:** 914–924

Offringa R (1992) Gene targeting in plants using the *Agrobacterium* vector systems. Doctoral dissertation, Leiden University

Offringa R, de Groot MJ, Haagsman HJ, Does MP, van den Elzen PJ & Hooykaas PJJ (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *EMBO J.* **9:** 3077–3084

Oliver D (1994) Gus stability - the replies. Available at: http://www.bio.net/mm/arab-gen/1994-January/001801.html [Accessed April 9, 2018]

Park SY, Vaghchhipawala Z, Vasudevan B, Lee LY, Shen Y, Singer K, Waterworth WM, Zhang ZJ, West CE, Mysore KS & Gelvin SB (2015) *Agrobacterium* T-DNA integration into the plant genome can occur without the activity of key non-homologous end-joining proteins. *Plant J.* **81:** 934–946

Pearson CE, Zorbas H, Price GB & Zannis-Hadjopoulos M (1996) Inverted repeats, stem-loops, and cruciforms: significance for initiation of DNA replication. *J. Cell. Biochem.* **63:** 1–22

Rolloos M (2013) The effect of the DNA metabolism of the recipient cell on the fate of T-DNA translocated by *Agrobacterium*. Doctoral dissertation, Chapter **4**:87-109, Leiden University

Rolloos M, Dohmen MHC, Hooykaas PJJ & van der Zaal BJ (2014) Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **91:** 1240–1251

Rossi L, Hohn B & Tinland B (1996) Integration of complete transferred DNA units is dependent on the activity of virulence E2 protein of *Agrobacterium tumefaciens*. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 126–130

Sakalis PA (2013) Visualizing virulence proteins and their translocation into the host during *Agrobacterium*-mediated transformation. Doctoral dissertation, Chapter **4**:125-150, Leiden University

Sambrook J & Russell DW (2001) Molecular cloning: a laboratory manual 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Schrammeijer B, Hemelaar J & Hooykaas PJJ (1998) The presence and characterization of a *virF* gene on *Agrobacterium vitis* Ti plasmids. *Mol. Plant-Microbe Interact.* **11:** 429–433

Shah KH, Almaghrabi B & Bohlmann H (2013) Comparison of expression vectors for transient expression of recombinant proteins in plants. *Plant Mol. Biol. Report.* **31:** 1529–1538

Sikorski RS & Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122:** 19–27

Singer K, Shiboleth YM, Li J & Tzfira T (2012) Formation of complex extrachromosomal T-DNA structures in *Agrobacterium tumefaciens*-infected Plants. *Plant Physiol.* **160:** 511–522

van Attikum H, Bundock P & Hooykaas PJJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *EMBO J.* **20:** 6550–6558

van Kregten M, De Pater S, Romeijn R, van Schendel R, Hooykaas PJJ & Tijsterman M (2016) T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nat. Plants* **2:** 1–6

van Kregten M, Lindhout BI, Hooykaas PJJ & van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Mol. Plant-Microbe Interact.* **22:** 1356–1365

Vergunst AC & Hooykaas PJJ (1998) Cre/lox-mediated site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* by transient expression of *cre*. *Plant Mol. Biol.* **38:** 393–406

Yadav NS, Vanderleyden J, Bennett DR, Barnes WM & Chilton M-D (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 6322–6326

Yusibov VM, Steck TR, Gupta V & Gelvin SB (1994) Association of single-stranded transferred DNA from *Agrobacterium tumefaciens* with tobacco cells. *Proc. Natl. Acad. Sci. U.S.A.* **91:** 2994–2998

Zaltsman A, Lacroix B, Gafni Y & Citovsky V (2013) Disassembly of synthetic *Agrobacterium* T-DNA-protein complexes via the host SCFVBF ubiquitin-ligase complex pathway. *Proc. Natl. Acad. Sci. U.S.A.* **110:** 169–174

Zonneveld BJM (1986) Cheap and simple yeast media. *J. Microbiol. Methods* **4:** 287–291

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res.* **31:** 3406–3415

Supplemental information

Supplemental Table S1. Tables showing the transformation efficiency of co-cultivations within figures.

A: Transformation efficiency in *S. cerevisiae* YPH250 of co-cultivation experiments from **figure 9** before normalization. The donor strain used was *A. tumefaciens* LBA1100. Each table shows one experimental set-up. Co-cultivation was performed for 48 h with the recipient strain. Errors shown are SEM.

Constructs used	Average transformation efficiency (n = 10)
p14	$(1.4 \pm 1.1) \times 10^{-6}$
3´HPΔ7b	$(3.7 \pm 1.9) \times 10^{-7}$

Constructs used	Average transformation efficiency (n = 10)
p14	$(4.4 \pm 3.6) \times 10^{-7}$
DHΔ7b	$(1.2 \pm 0.5) \times 10^{-7}$

Constructs used	Average transformation efficiency (n = 10)
p14	$(3.7 \pm 0.5) \times 10^{-7}$
5´HPL	$(1.4 \pm 0.3) \times 10^{-7}$

B: Transformation efficiency in *S. cerevisiae* YPH250 of co-cultivation experiments from **figure 10** before normalization. The donor strain used was *A. tumefaciens* LBA1100. Each table shows one experimental set-up. Co-cultivation was performed for 24 h and 48 h with the recipient strain. Errors shown are SEM.

Constructs used	Average transformation efficiency (n = 10)
p14	$(3.1 \pm 0.8) \times 10^{-8}$
DHΔ7b (24 h)	$(7.5 \pm 0.4) \times 10^{-8}$

Constructs used	Average transformation efficiency (n = 10)
p14	$(4.4 \pm 3.6) \times 10^{-7}$
DHΔ7b (48 h)	$(1.2 \pm 0.5) \times 10^{-7}$

C: Transformation efficiency in *K. lactis* wild-type of co-cultivation experiment from **figure 11** before normalization. The donor strain used was *A. tumefaciens* LBA1100. Co-cultivation was performed for 48 h. Errors shown are SEM.

Constructs used	Average transformation efficiency (n = 6)
5'HPL	$(5.0 \pm 1.9) \times 10^{-7}$
3'HP	$(4.7 \pm 1.3) \times 10^{-7}$
DH	$(7.0 \pm 4.2) \times 10^{-7}$
DHΔ7b	$(1.2 \pm 0.4) \times 10^{-6}$
3´HPΔ7b	$(1.1 \pm 0.4) \times 10^{-6}$
p14	$(3.4 \pm 1.6) \times 10^{-7}$

Table continued on the next page.

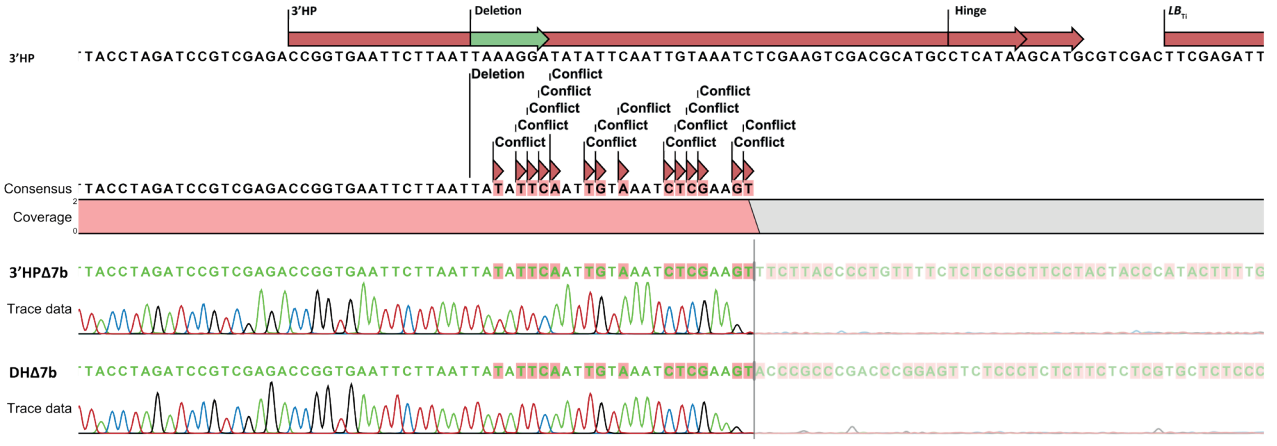
D: Transformation efficiency in *K. lactis* **Δ***rad52* of co-cultivation experiment from **figure 12** before normalization. The donor strain used was *A. tumefaciens* LBA1100. Co-cultivation was performed for 48 h. Errors shown are SEM.

Constructs used	**Average transformation efficiency (n = 3)**
5'HPL	$(5.0 \pm 1.4) \times 10^{-7}$
3'HP	$(8.8 \pm 4.2) \times 10^{-8}$
DH	$(5.1 \pm 2.8) \times 10^{-8}$
DHΔ7b	$(3.7 \pm 0.9) \times 10^{-7}$
3´HPΔ7b	$(2.2 \pm 1.4) \times 10^{-8}$
p14	$(9.3 \pm 2.3) \times 10^{-8}$

E: Transformation efficiency in *S. cerevisiae* YPH250 of co-cultivation experiment from **figure 13** before normalization. The donor strain used was *A. tumefaciens* LBA1100. Co-cultivation was performed for 48 h with the recipient strain. Errors shown are SEM.

Constructs used	**Average transformation efficiency (n = 10)**
3'tRNA	$(2.2 \pm 2.1) \times 10^{-6}$
p14	$(2.2 \pm 3.6) \times 10^{-6}$

A



B

```
CLUSTAL O(1.2.4) multiple sequence alignment

Construct       Sequence                                                Length (b)
3'HP            GAGACCGGTGAATTCTTAATTAAAGGATATATTCAATTGTAAATCTCGAAGT    52
3'HPΔ7b         GAGACCGGTGAATTCTTAAT-------TATATTCAATTGTAAATCTCGAAGT    45
DHΔ7b           GAGACCGGTGAATTCTTAAT-------TATATTCAATTGTAAATCTCGAAGT    45
                ********************       *************************
```

Supplemental Figure S2. Sequencing of the mutated hairpin sequence. **A:** Alignment of sequencing trace data of 3´HPΔ7b and DHΔ7b constructs with the 3'HP sequence. This depiction shows a simple sequence alignment with additional consensus sequence and coverage of the CLC Genomics Workbench 6.0 (https://www.qiagenbioinformatics.com). The trace data was taken from the chromatogram received from Macrogen Europe Laboratories. The sequencing primer used was HpinExOut. **B:** Multiple sequence alignment of the 3'HP, 3´HPΔ7b and DHΔ7b sequences using CLUSTAL Omega using seeded guide trees and HMM profile-profile to generate the alignment (Madeira *et al.*, 2019). The 7 b deletion is visible for the sequences of 3´HPΔ7b and DHΔ7b compared with the sequence of 3'HP.

Chapter 3

Horizontal gene transfer from *Agrobacterium* to *Streptomyces*

Felix T. Wittleben, Martijn Rolloos, Dennis Claessen,
Paul J.J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Abstract

Agrobacterium-mediated transformation (AMT), is a transfer process wherein the bacterium moves a defined part of a plasmid via an inducible type-IV secretion system (T4SS) into a eukaryotic recipient cell. AMT has been developed into an important technique for the generation of transgenic plants and fungi. To expand the use of AMT, more insight into the requirements for successful transfer of DNA is essential. This study was aimed to investigate AMT to the Gram-positive bacterium *Streptomyces lividans*. In theory such AMT to a genetically-tractable bacterium might provide a high-throughput assay to address remaining questions regarding the fate of the transferred DNA, and the role of the virulence proteins involved. AMT to *Streptomyces* was compared to the transfer via the bacterial conjugative plasmid RP4, containing a similar T4SS. Unfortunately, despite the use of a range of different media and incubation conditions, no AMT of *S. lividans* was detected mediated by the Ti plasmid. However, RP4-mediated DNA transfer from *A. tumefaciens* to *S. lividans* was observed. These findings can provide a basis for a dedicated approach to compare elements of both DNA transfer systems helping to unravel the host bias of the AMT system.

Introduction

Agrobacterium tumefaciens, a Gram-negative bacterium of the rhizosphere, is naturally able to transfer DNA and proteins into plant cells. In laboratory settings transfer is also possible to other eukaryotic organisms, for example yeasts and filamentous fungi (de Groot *et al.*, 1998; Lacroix *et al.*, 2006, Soltani *et al.*, 2008). The highly specialized process of *Agrobacterium*-mediated transformation (AMT), able to cross the *trans*-kingdom barrier, has evolved from bacterial conjugation (Alvarez-Martinez & Christie, 2009; Guglielmini, de la Cruz & Rocha, 2013). The specific transfer and transformation process is mediated by a specialized inducible type-IV secretion system (T4SS) encoded on the tumor-inducing (Ti) plasmid. This T4SS, named here $T4SS_{Ti}$, is functioning during AMT and shares homology with other bacterial T4SSs. However, conjugative plasmid transfer mediated by the $T4SS_{Ti}$ to other bacteria has been shown in only a few cases (Beijersbergen *et al.*, 1992; Kelly & Kado, 2002; Kado & Kelly, 2006).

In general, the process of conjugative DNA transfer can be divided into distinctive parts (Alvarez-Martinez & Christie, 2009). The T4SS or also called the mating-pair formation (Mpf) system, creates the physical channel between the donor and the recipient cells. The T4SS consists of several protein subunits creating a transmembrane spanning structure. A coupling protein (CP), situated at the cytoplasmic side of the T4SS system, is responsible for specifying which substrate will be transferred via the T4SS. And the DNA transfer and replication (Dtr) system, which processes DNA within the donor cell and further aids with the transfer of the DNA into recipient. The to be transferred DNA is specified by a distinct *cis*-acting region that is recognized by the Dtr system, the origin of transfer (*oriT*). The Dtr system consists amongst others of a relaxase protein that makes a nick at the *oriT*, after which it remains covalently bound to the DNA. The protein-DNA complex is then recognized by the CP and transferred via the T4SS. Several Dtr auxiliary proteins aid the relaxase in its function. To study AMT to bacteria, the comparison to a related transfer system of the conjugative plasmid RP4 was made. As recipient for DNA, the Gram-positive bacterium *Streptomyces lividans* was used. Moreover, *S. lividans* is present in the same biome as *A. tumefaciens*, which could make it a natural recipient for AMT.

A. tumefaciens normally infects wounded plant tissue, causing crown gall tumors (Escobar & Dandekar, 2003). Tumors are formed after transfer and integration of genes located on the T-DNA part of the Ti plasmid, while another part of this plasmid contains the virulence (vir_{Ti}) genes encoding the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$. For transfer to occur, induction by extracellular environmental triggers (low pH) and compounds (e.g. acetosyringone) released by wounded cells is required (Sheikholeslam & Weeks, 1987). Acetosyringone activates the $VirA_{Ti}/VirG_{Ti}$ signaling pathway that regulates vir_{Ti} gene expression (Krishnamohan, Balaji & Veluthambi, 2002). Contrary to other bacterial conjugative systems, the transferred DNA sequence located on the Ti plasmid is marked by two *cis*-acting elements for recognition by the Dtr_{Ti} system (Yadav *et al.*, 1982; Waters *et al.*, 1991; Alvarez-Martinez & Christie, 2009). These two imperfect direct repeats are called the right border (RB_{Ti}) and left border (LB_{Ti}). The region defined by these borders is called transfer-DNA (T-DNA). With the aid of auxiliary Vir_{Ti} proteins $VirC1_{Ti}$, $VirC2_{Ti}$, and $VirD1_{Ti}$ the relaxase $VirD2_{Ti}$ can nick the RB_{Ti} and LB_{Ti} sequence, whereby it becomes covalently bound to the 5′-ends. The RB_{Ti} is accompanied by an overdrive sequence (OD_{Ti}) to enhance border processing (Toro *et al.*, 1988). The area between the borders is liberated as a T-strand, bound at the 5'-end by $VirD2_{Ti}$, and is transported to the recipient cell via the $CP_{Ti}/T4SS_{Ti}$ system, migrating into the nucleus of the recipient cell for subsequent genomic integration (Gelvin, 2003; Christie & Gordon, 2014).

The IncPα plasmid RP4 is a conjugative plasmid able to transfer itself and other mobilizable plasmids from one bacterial cell to another (Pansegrau *et al.*, 1994). The

conjugative transfer is mediated via the Dtr_{RP4}, CP_{RP4} and $T4SS_{RP4}$ encoded on the plasmid. The relaxase of RP4 $TraI_{RP4}$, part of the Dtr_{RP4}, can recognize the *cis*-acting $oriT_{RP4}$ sequence. $TraI_{RP4}$, with the aid of auxiliary proteins, is able to nick at the $oriT_{RP4}$, whereafter a single-strand can be transferred via the $CP_{RP4}/T4SS_{RP4}$ into the recipient cell (Pansegrau *et al.*, 1990). As normal for bacterial conjugation, only one $oriT_{RP4}$ is present on the plasmid and therefore the entire plasmid sequence is transferred via the $CP_{RP4}/T4SS_{RP4}$. The transferred single-strand is converted into a double-strand and re-circularized with the aid of endogenous DNA processing systems within the recipient cell (Lanka & Wilkins, 1995; Pansegrau & Lanka, 1996; Parker & Meyer, 2005; Fernández-López *et al.*, 2014). One big difference to the AMT system is the presence of a plasmid-encoded primase in the Dtr_{RP4}, which is also transferred by $CP_{RP4}/T4SS_{RP4}$. The primase seems to play an important role for RP4-mediated conjugation (Merryweather, Barth & Wilkins, 1986). Within the recipient cell, the primase is thought to process the single-stranded plasmid to initiate double strand formation via the recipient cell endogenous replication machinery (Lanka & Barth, 1981). The host range of AMT are primarily eukaryotic cells, but it was also demonstrated that the $CP_{Ti}/T4SS_{Ti}$ Mpf system could mobilize plasmid RSF1010 into other *A. tumefaciens* cells (Beijersbergen *et al.*, 1992) and plant cells (Buchanan-Wollaston *et al.*, 1987). RSF1010 is lacking a CP/T4SS system and RB_{Ti} or LB_{Ti} sequences but contains instead its own $oriT_{RSF1010}$ and its own $Dtr_{RSF1010}$ (Scholz *et al.*, 1989). In this manner, the RSF1010 was able to complement its own $Dtr_{RSF1010}$ functions with the $CP_{Ti}/T4SS_{Ti}$ for plasmid transfer.

RP4 has a broad host-range, enabling it to be maintained in many Gram-negative bacteria (Grahn *et al.*, 2000). It is also able to mobilize plasmids from Gram-negative to Gram-positive bacteria (Mazodier, Petter & Thompson, 1989). Interestingly, in laboratory settings RP4 can transfer DNA from Gram-negative *E. coli* bacteria to yeast cells, thus crossing the *trans*-kingdom barrier too (Sawasaki *et al.*, 1996).

To compare the Ti plasmid and RP4 plasmid transfer systems for this study, two different experimental levels needed to be considered. One should be the environmental level, where the composition of the medium for the co-cultivation could play a role. A good setting for the medium should at the on hand to sustain growth of the donor and recipient cells, while on the other hand provide compatible conditions supporting transfer to occur. In general, for AMT low nutrient concentrations and low pH will be optimal, whereas for RP4-mediated transfer no transfer-specific medium seems to be required. The other level to be considered should deal with facilitation and comparison of the transfer process. To be able to compare AMT and RP4-mediated transfer on a simplified functional level, several adjustments had to be made to the AMT system. In contrast to AMT, RP4 typically transfers itself or a complete separate mobilizable plasmid by means of only one $oriT_{RP4}$ sequence. Hence, an analogous configuration needed to be established for AMT. It was discovered that AMT would be possible if the T-DNA and vir_{Ti} were separated from the Ti plasmid and presented on a different plasmid. DNA transfer via the $CP_{Ti}/T4SS_{Ti}$ is possible when T-DNA (defined by RB_{Ti} and LB_{Ti} sequences) is supplied *in trans* on a separate shuttle vector (Hoekema *et al.*, 1983), in combination with all essential vir_{Ti} genes making up the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ present on a separate plasmid, called the Ti helper plasmid (Hoekema *et al.*, 1983). This system was called the binary system. Additionally, it was demonstrated that a vector containing only a single border (RB_{Ti}) sequence could be processed by the Dtr_{Ti} and transferred via the $CP_{Ti}/T4SS_{Ti}$ to recipient cells (Kelly & Kado, 2002; Rolloos *et al.*, 2014). Thus, AMT can be modified to resemble the RP4-mediated transfer, by taking the RB_{Ti} sequence as a functional analog of an $oriT_{RP4}$. As a bacterial model organism to be used as recipient cells for studying any RB_{Ti}-only shuttle vector transfer via the $CP_{Ti}/T4SS_{Ti}$, *Streptomyces* could be a suitable substitute for plant cells as it was already described that AMT to *S. lividans* was possible (Kelly & Kado, 2002; Kado & Kelly, 2006). A high-throughput bacterial screening system could speed up the study of certain factors influencing AMT. As a standard method for

genetic manipulation of *Streptomyces*, a DNA transfer system based on RP4 is already widely used (Mazodier, Petter & Thompson, 1989; Paget *et al.*, 1999). With this method, a transfer-deficient RP4 plasmid lacking $oriT_{RP4}$ present in *E. coli* is used to deliver integrative vectors for genomic integration in *Streptomyces.* (Mazodier, Petter & Thompson, 1989; Paget *et al.*, 1999). If efficient, AMT might also serve as a future tool for dedicated genetic manipulation of *Streptomyces*, as they are very important organisms, being a major source of natural antibiotics (Chater, 2006; Newman & Cragg, 2007; Zhu, Sandiford & van Wezel, 2014).

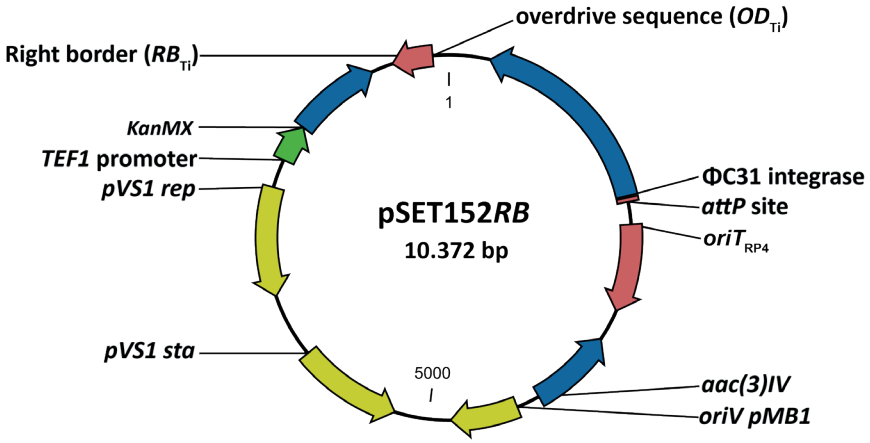


Figure 1. Shuttle vector pSET152*RB*, used for transfer between *Agrobacterium* and *Streptomyces*. The vector encodes a ΦC31 phage integrase, able to integrate the entire vector into the host genome, via site-specific recombination. The integrase recognition sequence on the shuttle vector is *attP* and in the *Streptomyces* genome *attB* (Kuhstoss & Rao, 1991). It further contains a *pVS1* minimal replicon (*sta* and *rep*) for *A. tumefaciens*, a replicon for *E. coli*, a *KanMX* resistance cassette for use in *S. cerevisiae*, an apramycin resistance cassette (*aac(3)IV*), a 25 bp octopine-type right border (RB_{Ti}) recognition sequence for $VirD2_{Ti}$ combined with an upstream adjacent 38 bp overdrive (OD_{Ti}) sequence for Dtr_{Ti} auxiliary protein recognition (Rolloos *et al.*, 2014), and a RP4 based origin of transfer ($oriT_{RP4}$) for mobilization by RP4 (Mazodier, Petter & Thompson, 1989). The shuttle vector pSET152Δ*RB* just lacks the RB_{Ti} for $VirD2_{Ti}$ recognition (vector map not shown). The shuttle vector pSET152*RB*Δ*oriT* is identical to pSET152*RB*, except for the deletion of the $oriT_{RP4}$ region (vector map not shown).

Results

Construction of a transfer system for *Agrobacterium*-mediated transfer to *Streptomyces*

Kelly and Kado (2002) previously reported AMT to *S. lividans*. The study described the transfer of two shuttle vectors, both containing a RB_{Ti} sequence, from the recombination deficient Ach5-based *A. tumefaciens* LBA4301 (Klapwijk, van Beelen & Schilperoort, 1979) to *S. lividans* 66 TK64 (Kelly & Kado, 2002), facilitated via the Dtr_{Ti}, CP_{Ti} and $T4SS_{Ti}$ encoded on the Ti helper plasmid pUCD2614. The Ti helper plasmid was created by subcloning only the vir_{Ti} regulon of the Ti plasmid pTiC58 (Rogowsky *et al.*, 1990). The first mobilizable vector contained the RB_{Ti} sequence, a hygromycin and kanamycin/gentamicin resistance cassette. The second mobilizable vector was a fusion of the plasmid pSET152 (Bierman, Logan & O'Brian, 1992) and the first vector (Zyprian & Kado, 1990). Plasmid pSET152 contained an apramycin resistance cassette as well as the gene encoding a ΦC31 phage integrase, for genomic integration of plasmid pSET152 by recombination between the *att* sites on the plasmid and the chromosome of *S. lividans* (Kuhstoss & Rao, 1991; Bierman, Logan & O'Brian, 1992). The second vector did not contain the hygromycin resistance cassette of the parental vector nor the $oriT_{RP4}$ sequence present on pSET152. At the onset of the study, it was attempted to repeat the experiments performed by the Kado lab. Unfortunately, the originally used strains and plasmids could no longer be obtained (CI Kado, personal communication). Therefore, this study used a set-up resembling the original set-up of Kelly & Kado (2002) but with different strains and plasmids. Kelly & Kado (2002) used strain *S. lividans* TK64 containing a loss of function mutation of proline synthesis and streptomycin resistance and lacking two linear plasmids. For this study the progenitor *S. lividans* wild-type 66 (1362) without genomic modifications was used. Here, the C58 based *A. tumefaciens* strain LBA1100 (Beijerbergen *et al.*, 1992) was used as donor. This strain harbors the Ti helper plasmid pAL1100 (Beijerbergen *et al.*, 1992) which is derived from the octopine Ti plasmid pTiB6, by a large deletion including the T-regions. The constructed shuttle vector pSET152*RB* (**Figure 1**), based on plasmid pSET152, contained in addition to the $oriT_{RP4}$ sequence a RB_{Ti} with an adjacent OD_{Ti} (**Figure 1**).

To verify the correct function of the pSET152*RB* shuttle vector (**Figure 1**), the plasmid was transformed via protoplast transformation directly into *S. lividans* according to Gust, Kieser & Chater (2002). The transformation resulted in apramycin resistant colonies (n = 3; data not shown). The shuttle vector was unable to replicate in *S. lividans*, and the apramycin resistance of *S. lividans* was thus only acquired after the integrase-mediated integration event of the plasmid into the recipient cell's genome (Paranthaman & Dharmalingam, 2003). The stable genomic integration of the plasmid was verified by PCR as well as by subsequent cultivation of positive colonies on medium selecting further on apramycin (n = 3; data not shown).

Next, it was tested if the shuttle vector pSET152*RB* could be processed and transferred via the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$. For this the plasmid was additionally equipped with an *ARS6/Cen4* sequence for stable low copy maintenance in yeast, resulting in plasmid pSET152Y*RB*. This was qualitatively tested by using AMT transfer to yeast following the modified protocol of Bundock *et al.* (1995). As plasmid donor *A. tumefaciens* strain LBA1100 containing pAL1100 and pSET15Y*RB* was used, and as recipient *Saccharomyces cerevisiae* strain YPH250. This co-cultivation resulted in observable kanamycin resistant yeast colonies with an estimated conjugation frequency of around 10^{-5} (n = 3; data not shown).

To exclude that free plasmid pSET152*RB* during AMT, present in the co-cultivation mix from lysed bacteria, could be responsible for transformation of *S. lividans*, the conjugation

protocol of Gust, Kieser & Chater (2002) was followed, with the change of adding only free plasmid instead of a donor strain. Around 1 × 10^7 pre-germinated spores were used and incubated with purified pSET152*RB* plasmid DNA (1 µg/1 × 10^7 pre-germinated spores). After incubation and selection, no apramycin positive colonies were observed, showing that free plasmids were not taken up under the conditions used (n = 3; data not shown).

AMT to *S. lividans* based on methodology by Kelly and Kado (2002)

To investigate the plasmid transfer between *A. tumefacien*s and *S. lividans* mediated by the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$, a co-cultivation experiment was performed. The original co-cultivation protocol of Kelly & Kado (2002) was followed. First, pre-germinated *S. lividans* spores were combined with induced *A. tumefaciens* LBA1100 cells containing pSET152*RB* and then cultivated together for 4 d. During this time *S. lividans* was able to form mycelia, which produced new spores at the end of the co-cultivation. The *S. lividans* spores found after the co-cultivation were recovered and a dilution series was plated on antibiotic selection plates. None of the transferred *S. lividans* spores exhibited growth on antibiotic selection (n = 9). As a control, part of the dilution series was plated on medium without antibiotic selection to verify growth of *S. lividans* cells, which led to bacterial growth in each experiment (qualitative measurement; no data shown).

During conjugation, cell-to-cell contact is crucial for the T4SS to create a transmembrane bridge for DNA transfer. To study the interaction between the organisms, Kelly & Kado (2002) used an aggregate formation assay. In their study, it was observed that acetosyringone induced *A. tumefaciens* LBA4301 cells to aggregate with *S. lividans* hyphae into structures visible in test tubes. To check for this aggregation, a comparable experimental set-up was used. As donor cells, two different *A. tumefaciens* strains were used: a derivative of *A. tumefaciens* LBA1100 containing the Ti helper plasmid pAL1100 as well as the shuttle vector pSET152*RB*, and a derivative of *A. tumefaciens* LBA2210 containing the plasmid RP4 as well as the shuttle vector pSET152*RB*. As recipients *S. lividans* 66 as well as the unrelated *Streptomyces venezuelae* were used. In the experiments, *A. tumefaciens* LBA1100 was tested induced as well as non-induced. Remarkably, in none of the experimental combinations and conditions, using induced or non-induced LBA1100 as well as in non-induced conditions LBA2210, any visible aggregation was detected such as reported by Kelly & Kado (2002) when using strain LBA4301 (**Table 1**).

Table 1. Aggregate formation assay between *Streptomyces* hyphae and *A. tumefaciens*. Induced (with acetosyringone; AS) or non-induced *A. tumefaciens* was cultured together with 48 h old *S. lividans* hyphae at room temperature to observe possible formation of aggregates between both organisms, visible to the eye. Observations were taken for following time points: 5, 15 and 30 min and 12 h (each n = 1).

Streptomyces	***A. tumefaciens***	***A. tumefaciens* induction**	**Observation**
S. lividans	LBA1100 (pSET152*RB*)	5.5 h induced in IM+AS	No aggregation
S. lividans	LBA1100 (pSET152*RB*)	5.5 h in LC	No aggregation
S. venezuelae	LBA1100 (pSET152*RB*)	5.5 h induced in IM+AS	No aggregation
S. lividans	LBA2210 (pSET152*RB*)	5.5 h in LC	No aggregation

IM: Induction medium; AS: Acetosyringone (200 µM)

Modifications of co-cultivation conditions for AMT to *S. lividans*

The co-cultivation, essentially replicating the published protocol for AMT transfer of *S. lividans* by Kelly & Kado (2002), did not result in transformants with the strains and plasmids used in this study. This raised the question, if the applied co-cultivation conditions allowed DNA transfer between the strains. To find compatible conditions for Dtr_{Ti}-, CP_{Ti}-, $T4SS_{Ti}$-mediated DNA transfer to occur, several variations alongside the conditions used by Kelly & Kado (2002) were tested: use of different media, a different co-cultivation method required for the use with a specific medium, as well as using a different strain of *A. tumefaciens*. The compared media were: (1) Induction medium (IM) used normally for AMT to yeast, containing minimal nutrients and specifically a low pH supporting the expression and function of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$; (2) R2 medium, a specific growth medium for *Streptomyces* used by Kelly & Kado (2002) with a pH of around 7; and (3) R5 medium, a variation of the R2 medium, containing yeast extract for better growth of *Streptomyces* with a pH of around 7. Further, different methods of selection for recipient cells were tested. One selected method was the overlay procedure commonly used for conjugation to *Streptomyces* (Gust *et al.*, 1999), also used in the Kelly & Kado (2002) experiment. In this case only the *Streptomyces* specific media R2 and R5 could be used for co-cultivation, selection and growth of *Streptomyces*. In the overlay method after co-cultivation, an aqueous overlay containing an antibiotic for selection of the transformants and an antibiotic for selection against the donor strain was applied directly on the co-cultivation plate. Subsequently, surviving *Streptomyces* cells were grown until new spores were available for isolation and further selection on the medium that was also used for the co-cultivation, for example R2 or R5 medium as they support *Streptomyces* growth. The overlay method is optimized for *Streptomyces* growth and does not account for any requirements for *Agrobacterium* cells and induction of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$. The other method was based on a standard AMT protocol that specifically allowed *Agrobacterium* cells to induce their $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ (Bundock *et al.*, 1995). In this method a filter was placed on the co-cultivation medium to keep the mixture of donor and recipient cells separate from the medium. The filter still allowed transfer of nutrients and compounds. The advantage of this method was that the co-cultivation could be performed on one specific co-cultivation medium. After co-cultivation the filter could be removed and the biofilm (containing the donor and recipient cells) could be transferred to a different *Streptomyces* medium for selection and growth (R2 or R5). This method was in particular necessary, when IM medium was used for the co-cultivation, as continued presence of this medium did not support growth of *S. lividans* cells after co-cultivation. But *S. lividans* was found to stay viable when cultured on cellophane filters on IM for 4 d and then being transferred to R5 growth medium (data not shown). As a final variable, an *A. tumefaciens* strain LBA1126 was used that did not require induction of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$, as mutations in the Ti plasmid $virA_{Ti}$ and $virG_{Ti}$ regulatory genes ensure expression of the vir_{Ti} genes.

First, the protocol for yeast transformation using IM, providing optimal condition for *A. tumefaciens* virulence, was tested (Bundock *et al.*, 1995). The donor strain was *A. tumefaciens* LBA1100 containing the shuttle vector pSET152*RB* and the recipient strain was *S. lividans*. The used cell-to-cell ratio for the co-cultivation was 1×10^8 donor cells to 1×10^8 pre-germinated recipient spores. The co-cultivation was performed for different time spans (48 h, 96 h, and 5 d) on a filter on IM at 21° C. No apramycin resistant S. *lividans* colonies were subsequently observed after plating all different co-cultivations on R5 selection medium for up to 3 weeks (n = 3, for each time point; data not shown).

To directly compare the co-cultivation on IM medium to the co-cultivation on *Streptomyces* specific media, a co-cultivation experiment was performed on IM, R2, and R5 medium, using either the overlay or the filter method. For this comparison, the co-cultivation length was set at 4 d, just as for the initial experiments based on the Kelly & Kado (2002). *A. tumefaciens* LBA1100 containing the Ti plasmid and the shuttle vector pSET152*RB* was

used again as plasmid donor strain, with *S. lividans* as recipient strain. For further comparison another recipient, *S. venezuelae* was used as well. However, none of these approaches and variations yielded any apramycin resistant *Streptomyces* colonies (**Table 2).**

Table 2. Overview of different methods for AMT co-cultivation with *Streptomyces*. The *A. tumefaciens* strain LBA1100 with Ti helper plasmid pAL1100 and shuttle vector pSET152*RB* was tested as donor in different co-cultivation conditions with two different *Streptomyces strains* as recipient. During co-cultivation, the phenolic-inducer acetosyringone (AS) was present in each medium at a concentration of 200 μM. The co-cultivation period was 4 d at 21° C. After co-cultivation on filter, the resulting biofilm was removed and plated on strain specific medium with selection for apramycin resistance. With the overlay method, both strains were placed together directly on the agar plate, and after co-cultivation the biofilm was treated directly with antibiotics to select for apramycin resistance and against *A. tumefaciens* (for *S. lividans* n = 3; and for *S. venezuelae* n = 2).

Streptomyces	**Method**	**Medium**	**Transformants**
S. lividans	Cellophane filter	IM + AS	No
S. lividans	Cellophane filter	R2 + AS	No
S. lividans	Overlay	R2 + AS	No
S. lividans	Cellophane filter	R5 + AS	No
S. lividans	Overlay	R5 + AS	No
S. venezuelae	Cellophane filter	IM + AS	No
S. venezuelae	Cellophane filter	R5 + AS	No

IM: Ti plasmid specific induction medium with minimal nutrients and specific pH; R2: *Streptomyces* specific growth medium; R5: a variation of R2 medium containing additional yeast extract; AS: Acetosyringone

To investigate whether the use of a filter (with the filter method) during AMT would have negatively affected plasmid transfer, a co-cultivation was set-up directly on the medium As *S. lividans* cells were not able to continue their life cycle on IM medium, the biofilm containing the donor and recipient strain now needed to be removed directly from the IM medium surface, in order to be transferred to a selection medium. For making direct comparisons between IM and R5 medium, the same direct plating method was applied for R5 medium as well. The co-cultivation was performed on 6-well plates as the smaller surface area of the medium allowed for a more thorough removal of the biofilm. It was found that the maximal length of the co-cultivation period was limited to a window between 2 and 3.5 h. After that time window *S. lividans* spores would start creating hyphae, which would grow into the medium making it from then onwards impossible to remove the biofilm from the surface. After removal from the surface, the co-cultivation mixture was plated on antibiotic selection plates selecting for transformants and selecting against the donor cells. In none of the experiments any apramycin positive *S. lividans* colonies could be detected (n = 3; data not shown).

As already indicated above, another strain of *A. tumefaciens* was tested in order to have a comparison to the previously used strain LBA1100, in case this would have been an unfortunate choice. *A. tumefaciens* LBA1126 has mutations in the $virA_{Ti}$ and $virG_{Ti}$ regulatory genes on the Ti plasmid (Bundock & Hooykaas, 1996), which leads to constitutive activation of the vir_{Ti} genes. Therefore, no induction of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ should be required for this strain. This would in principle allow for full $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ activity even on *Streptomyces* specific media, which do not provide for the conditions triggering virulence gene expression. Hence, *A. tumefaciens* LBA1126 harboring pSET152*RB* was used for a co-cultivation with *S. lividans* on R5 medium (filter and overlay method) as well as on IM medium (filter method) for 4 d. Also, after these co-cultivation experiments no apramycin resistant colonies of *S. lividans* were observed (n = 3; data not shown). These results indicated that if any conjugative transfer via the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ would have occurred, it would have been below the frequency of 10^{-9}, the lower detection limit of this experiment.

RP4-mediated DNA transfer to *Streptomyces*

Although the plasmid substrate for AMT could be equipped with a single RB_{Ti} sequence to functionally resemble an *oriT* as used in the conjugative process of RP4, no experimental evidence for AMT transfer to *Streptomyces* could be obtained in this study. To have a direct comparison though, RP4-mediated transfer to *Streptomyces* had to be tested using *A. tumefaciens* as donor strain. So far, RP4 had not been tested or reported for transfer of DNA from *Agrobacterium* to *Streptomyces*.

A routine method for DNA transfer to *Streptomyces* has been established using *E. coli* ET12567. The ET12567 strain harbors the non-transmissible conjugative plasmid pUZ8002 (Paget *et al.*, 1999), which is a derivative of the conjugative plasmid RP4 (Pansegrau *et al.*, 1994) lacking the $oriT_{RP4}$ region. Therefore, pUZ8002 is not able to transfer itself, but is still able to mobilize other plasmids containing an $oriT_{RP4}$. To compare the routine pUZ8002 mediated *E. coli* to *Streptomyces* conjugative plasmid transfer with transfer starting from *A. tumefaciens*, the plasmid pUZ8002 was isolated from *E. coli* ET12567 and introduced into *A. tumefaciens* LBA288 lacking a Ti plasmid. Next, a co-cultivation experiment was performed using *A. tumefaciens* LBA288 containing pUZ8002 as the donor strain for the shuttle vector pSET152*RB* and *S. lividans* as recipient, using the overlay method on R5 medium for 4 d. No apramycin resistance *S. lividans* colonies were observed on the selection plates (n= 3; data not shown).

To verify if the isolated pUZ8002 plasmid would also be functional outside of the *E. coli* ET12567 strain, pUZ8002 was also introduced into another strain *E. coli* DH10β. Following the same protocol (Paget *et al.*, 1999), *E. coli* DH10β containing pUZ8002 and pSET152*RB* was co-cultivated with *S. lividans*. The co-cultivation was performed on specific MS medium, providing optimal conditions for a fast sporulation of *Streptomyces*. In this experiment after selection, no apramycin resistant *S. lividans* colonies could be observed (n = 3; data not shown). As the threshold of detection in this assay would be at frequency of 10^{-9} for this experiment, if transformation had been successful the frequency would be below the threshold frequency. In a parallel experiment, *E. coli* ET12567 containing pUZ8002 and pSET152*RB* was used as donor strain for a co-cultivation with *S. lividans* as recipient strain on MS medium following the protocol of Paget *et al.* (1999). After this co-cultivation apramycin resistant colonies were observed, with an estimated average transformation frequency of 10^{-4} (n = 3; no data shown).

In view of these variable results with pUZ8002, it was studied whether using RP4 itself instead of pUZ8002 would allow conjugative DNA transfer from *Agrobacterium* to *Streptomyces*. To this end the *A. tumefaciens* strain LBA2210 lacking a Ti plasmid but containing the RP4 plasmid was used. Co-cultivating *A. tumefaciens* LBA2210 harboring RP4 and pSET152*RB* with *S. lividans* according to Gust, Kieser & Chater (2002) resulted in transformation of *S. lividans*. Apramycin-resistant *S. lividans* colonies were observed with an estimated frequency of 10^{-5} (**Figure 2**). The co-cultivation was performed on R5 medium. This is the first report of *A. tumefaciens* to transfer an integrative shuttle vector to *S. lividans* within 24 h via the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$. As control transfer to *S. lividans* was tested with *A. tumefaciens* LBA2210 containing RP4 with either pSET152*RB*Δ*oriT* lacking the $oriT_{RP4}$ sequence, or pSET152Δ*RB*Δ*oriT* lacking the border sequence as well as $oriT_{RP4}$. After use of these two shuttle vectors in the co-cultivation, no apramycin positive colonies were visible (**Figure 2**).

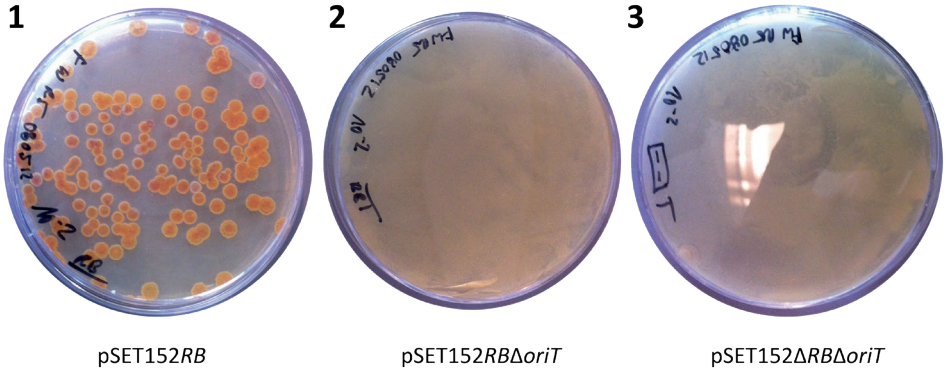


Figure 2. *A. tumefaciens* RP4-mediated transfer of shuttle vectors to *S. lividans* with or without $oriT_{RP4}$. (**1**). Apramycin resistant *S. lividans* colonies visible when plasmid pSET152*RB* was used. (**2**). No *S. lividans* colonies visible when conjugation was attempted with plasmid pSET152*RB*Δ*oriT*. (**3**). No *S. lividans* colonies visible when transfer was attempted with plasmid pSET152Δ*RB*Δ*oriT*. Each picture shown was a good representation of the other observed plates (n = 3). The screening was performed to determine only a visible qualitative result, but the transformation frequency was estimated to be around 10^{-5}. The co-cultivation (24 h) and selection was performed on R5 medium at 30° C.

Influence of co-cultivation media on DNA transfer from to *S. lividans*

As the RP4-mediated DNA transfer was possible from *Agrobacterium* to *S. lividans*, the co-cultivation conditions used for the AMT experiments were compared to assess if DNA transfer could occur during the specific AMT conditions (**Table 3**). This was a qualitative approach, and no transformation frequencies were determined. Only the occurrences of apramycin resistant colonies were noted. The media tested were IM, R2 as *Streptomyces* specific growth medium, and R5 as a more complex *Streptomyces* specific growth medium, being a derivative of R2 containing yeast extract, also already mentioned above. To emulate conditions required for AMT, the vir_{Ti} inducer acetosyringone was added to the R2 and R5 *Streptomyces* media or was left out for comparison. As RP4-mediated DNA transfer to *S. lividans* was possible, acetosyringone was added to indicate if there would be an influence on the DNA transfer from *A. tumefaciens.* In addition, either the filter method or the overlay method was used. Also the temperature was varied to be either 21° C, as optimal for AMT, or 30° C, as optimum for *S. lividans* transformation. The length of the co-cultivation was 24 h as this is an optimal time window for RP4/pUZ8002-mediated DNA transfer. Apramycin positive colonies were observed after co-cultivation on R2 and R5 media, with or without acetosyringone, at 21° C or 30° C, and with both co-cultivation methods used (n = 3; for each data point). No apramycin positive colonies were observed on any of the co-cultivations using IM (n = 3; for each data point). These data are again indicative for problems with the cultivation of *Streptomyces* on IM during co-cultivation, as also encountered during earlier performed co-cultivations.

Table 3. RP4-mediated transfer from *A. tumefaciens* LBA2210 containing pSET152*RB* to *S. lividans* in different co-cultivation conditions. Co-cultivation occurred on the medium and by the method indicated in the table. After co-cultivation only the qualitative occurrence of apramycin resistant colonies was determined (n = 3; each data point). Acetosyringone (AS) was added to the medium where indicated, at a concentration of 200 μM.

Medium	Method	Temp. in [° C]	Time in [h]	Transformants
R2	Cellophane filter	21/30	24	Yes
R2 + AS	Cellophane filter	21/30	24	Yes
R2	Overlay	21/30	24	Yes
R2 + AS	Overlay	21/30	24	Yes
R5 +AS	Cellophane filter	21/30	24	Yes
R5	Cellophane filter	21/30	24	Yes
R5 +AS	Overlay	21/30	24	Yes
R5	Overlay	21/30	24	Yes
IM + AS	Cellophane filter	21/30	24	No
IM	Cellophane filter	21/30	24	No
IM + AS	Overlay	21/30	24	No
IM	Overlay	21/30	24	No

IM: Ti plasmid specific induction medium with minimal nutrients and specific pH; R2: *Streptomyces* specific rich growth medium; R5: a variation of R2 medium containing additional yeast extract; AS: Acetosyringone

Comparison of RP4-mediated transfer from *A. tumefaciens* and *E. coli* to *S. lividans*

To gain more clarity about the RP4-mediated DNA transfer to *Streptomyces*, the transformation efficiencies of *E. coli* ET12567 harboring the pUZ8002 and the shuttle vector pSET152*RB,* and of *A. tumefaciens* LBA2210 containing the RP4 plasmid as well as pSET152*RB* were determined. The standard *Streptomyces* transformation protocol using *E. coli* uses MS medium, a rich medium specific for fast sporulation of transformed *Streptomyces* cells (Paget *et al.*, 1999). The transformation frequency of *A. tumefaciens* with *S. lividans* on MS medium could not be determined, as *A. tumefaciens* then showed an increased production of extracellular matrix; a very thick slime layer was formed on top of the bacterial lawn and no colonies of *S. lividans* could be identified. Furthermore, it is unknown if the antibiotic overlay can permeate such a barrier. Slime formation was avoided by using R5 medium, used for *S. lividans* transformants selection. No apramycin resistant *S. lividans* were observed after co-cultivation of *E. coli* ET12567 using pUZ8002 and pSET152*RB* on R5 medium. Apramycin resistant *S. lividans* colonies were only detected after a RP4-mediated transfer both with *A. tumefaciens* on R5 medium at 21° C or *E. coli* on MS medium at 30° C (**Table 4**). Also, apramycin-positive cells were only found after a co-cultivation time of 24 h. Co-cultivation for 30, 60, and 120 min did not result in any positive transformation of *S. lividans*, neither from *A. tumefaciens* at 21° C nor from *E. coli* at 30° C (**Table 4**).

Table 4. Comparison of RP4-mediated conjugation of *E. coli* and *A. tumefaciens* to *S. lividans* on different media. As plasmid donors *E. coli* ET12567 containing the plasmid pUZ8002 and *A. tumefaciens* LBA2210 containing the plasmid RP4 were used. The recipient strain was *S. lividans*. Co-cultivation was performed for 30 min, 60 min, 120 min and 24 h (n = 3; for each time point). Co-cultivation with *E. coli* was performed at 30° C and with *A. tumefaciens* at 21° C. Error shown is SEM. The conjugation efficiency was determined by dividing the output transconjugants CFU by the input recipient cells.

Donor strain (plasmids used)	Medium	Conjugation efficiency			
		30 min	60 min	120 min	24 h
E. coli ET12567 (pUZ8002, pSET152*RB*)	MS	0	0	0	$(3.3 \pm 1.0) \times 10^{-4}$
	R5	0	0	0	0
A. tumefaciens LBA2210 (RP4, pSET152*RB*)	MS	n.d.	n.d.	n.d.	n.d.
	R5	0	0	0	$(6.5 \pm 4.1) \times 10^{-5}$

MS: *Streptomyces* sporulation medium; R5: *Streptomyces* growth medium; n.d.: not determined

Discussion

This study investigated DNA transfer from *A. tumefaciens* to *S. lividans* via two functionally different but structurally related transfer systems, those of the RP4 and Ti plasmid. To compare both transfer systems, a compatible functionally similar DNA substrate was required, carrying a single RB_{Ti} site as an equivalent of an $oriT_{RP4}$, thus making the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ transfer system a functional mimic of the conjugative plasmid transfer by the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$. In this study, previously published results showing AMT to *S. lividans* (Kelly & Kado, 2002) could not be corroborated. Variations in co-cultivation conditions also did not result in any detectable AMT via the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$. However, it was shown that *A. tumefaciens* containing the RP4 plasmid was able to mobilize an $oriT_{RP4}$ containing integrative vector to *S. lividans*. Comparing all data, it had to be concluded that none of the conditions regularly used for AMT transfer to yeast and plant cells was compatible with DNA transfer from *A. tumefaciens* to *S. lividans*.

In 2002, Kelly & Kado reported a method for AMT to *S. lividans*, followed by a methodological publication (Kado & Kelly, 2006). The strains and constructs originally used for this study (Kelly & Kado, 2002) were not available anymore. In the set-up of Kelly & Kado (2002), rich *Streptomyces* specific R2 medium was used for the co-cultivation of organisms (Kieser *et al.*, 2000) at a temperature of 25° C. These conditions seemed to be not optimal for $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$-mediated DNA transfer conditions. Key factors for successful AMT are specific medium, pH and temperature. According to literature, rich medium is sub-optimal for AMT (Turk *et al.*, 1991; Liu *et al.*, 1993). A minimal medium like IM is optimal for AMT, mimicking the natural environment for transformation (Beijersbergen *et al.*, 1992; Bundock *et al.*, 1995). The optimum pH for AMT is an acidic pH between 5 and 6 (Turk *et al.*, 1991; Liu *et al.*, 1993; Gelvin, 2006). The pH of R2 medium is around 7.2 (Okanishi, Suzuki & Umezawa, 1973) providing less optimal conditions for the transformation. According to literature, the window between 21°-22° C provides an optimal window for a more efficient assembly of the $T4SS_{Ti}$ pilus (Dillen *et al.*, 1997; Fullner and Nester 1996; Baron *et al.*, 2001). The 25° C co-cultivation temperature chosen by Kelly & Kado (2002) is within the upper limit of possible AMT. To investigate why no DNA transfer was observed with the published conditions, a different medium and a different method was tried. In particular IM, a specialized medium for optimized expression and function of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$, was used for co-cultivation and combined with a filter method to be able to retrieve *S. lividans* for selection after the co-cultivation. This approach unfortunately also did not result in AMT of *S. lividans* (**Table 2**).

Though *S. lividans* was shown to be viable on IM, RP4-mediated DNA transfer to *S. lividans* on IM did not occur (**Table 3**). Therefore, different media, co-cultivation times and temperatures were tested subsequently. R5 rich medium (specific for *Streptomyces*) was used instead of IM medium, cultured on cellophane filter or directly on the agar surface, but this also did not generate positive results (**Table 2**). For activation of the virulence system the medium should contain preferably a lower pH (between 5 and 6), and fewer nutrient sources than present in the tested R5 and R2 media at a temperature of around 21° C. The virulence proteins $VirA_{Ti}$ and $VirG_{Ti}$ form a two-component signaling pathway, which mediates the expression of the other Ti plasmid encoded virulence proteins at low pH and room temperature in the presence of acetosyringone (Stachel, Nester & Zambryski, 1986; Shaw *et al.*, 1988; Turk *et al.*, 1991). However, previously an *A. tumefaciens* helper strain called LBA1126 was constructed that showed hypersensitivity for induction of vir_{Ti} expression even at a pH of 7 (Bundock & Hooykaas, 1996). LBA1126 contains modifications in both $virA_{Ti}$ and $virG_{Ti}$. Its $VirA_{Ti}$ protein has its periplasmic domain replaced by that of the *E. coli* TAR chemoreceptor, so that the protein is locked in a highly responsive state making it also less thermo sensitive and less pH sensitive (Melchers *et al.*, 1989; Turk *et al.*, 1993). The $VirG_{Ti}$ protein with substitution (I77V) already gives maximal vir_{Ti} gene expression in the presence of

low amounts of acetosyringone (Scheeren-Groot *et al.*, 1994). Unfortunately, also when using the strain LBA1126 no transformants of *S. lividans* could be observed after co-cultivation on R5 or even IM medium.

When tested in this study, *E. coli* ET12567 containing pUZ8002 was able to mobilize the shuttle vector pSET152*RB* containing the *oriT*$_{RP4}$ into *S. lividans* (**Table 4**). However, when plasmid pUZ8002 and pSET152*RB* were co-introduced into *A. tumefaciens* LBA288 or *E. coli* DH10β no conjugative transfer to *S. lividans* was observed. Both, *A. tumefaciens* and *E. coli* DH10β are known to methylate DNA (Gelvin, Karcher & DiRita, 1983; Kahng & Shapiro, 2001; Grant *et al.*, 1990). *Streptomyces* species in general contain a methylation restriction system for foreign DNA (González-Cerón, Miranda-Olivares & Servín-González, 2009), and for that reason the methylation deficient *E.coli* strain ET12567 is used as a donor for RP4-mediated *Streptomyces* transformation (Flett, Mersinias & Smith, 1997). But for *S. lividans* the restriction system is functioning in a low state (McNeil, 1988; González-Cerón, Miranda-Olivares & Servín-González, 2009). Therefore, methylated DNA that would come from the donor strains used during co-cultivation should only show a roughly 10% reduction in transformation efficiency as previously described (González-Cerón, Miranda-Olivares & Servín-González, 2009). In this study it was not attractive to replicate the experiment of González-Cerón, Miranda-Olivares & Servín-González (2009) to check if isolated DNA from *A. tumefaciens* would have a reduced transformation efficiency due to methylation, as high-quality plasmid DNA is difficult to obtain from *Agrobacterium*. Indeed RP4-mediated DNA transfer from *A. tumefaciens* LBA2210 to *S. lividans* turned out to be possible, although at a lower frequency than from *E. coli* ET12567 (**Table 4**). This raised questions about the original modification of pUZ8002. If pUZ8002 would only lack the *oriT*$_{RP4}$ sequence, the lack of mobilization activity was unexpected. The plasmid pUZ8002 was initially created for use in the *E. coli* strain ET12567 (McNeil *et al.*, 1992). The modification of the plasmid (except the mutation in the *oriT*$_{RP4}$ region) was not fully known and citations lead to a final reference marked as personal communication of the author (Paget *et al.*, 1999). It could be for example that the *E. coli* strain ET12567 could supply a function for pUZ8002 to mediate DNA transfer. It is also to note that ET12567 pUZ8002-transfer *to S. lividans* did not occur when co-cultivation was performed on R5 medium, where *A. tumefaciens* RP4-mediated DNA transfer was possible.

It could be assumed that the environmental impact of the used co-cultivation medium, pH and temperature would be the main parameters influencing the transformation as discussed earlier. During conjugative plasmid transfer between bacteria, the DNA-bound relaxase is transferred into the recipient as well, enabling re-circularization of the single-stranded DNA molecule (Dostal *et al.*, 2011). However, in the recipient cell endogenous DNA processing proteins are required for double-strand DNA formation. Remarkably, the RP4 plasmid encodes a primase that is transferred in the course of co-cultivation. The primase acts on the *oriT*$_{RP4}$ present on the shuttle vector, initiating priming of the single-stranded DNA for DNA polymerization by the host DNA replication machinery (Yakobson *et al.*, 1990; Miele *et al.*, 1991; Pansegrau *et al.*, 1994). The presence of a transferred primase could explain the successful RP4-mediated transfer of the shuttle vector pSET152*RB* from *A. tumefaciens* to *S. lividans*, while AMT remained undetectable. Also it is not certain if the re-circularization of the plasmid at its *RB*$_{Ti}$ could occur in *S. lividans* after AMT, although evidence has been obtained that this can happen in yeast (Bundock *et al.*, 1995; Rolloos *et al.*, 2014). If the RP4 primase is necessary for RP4-mediated transfer to *S. lividans*, adding this primase to the Ti system may lead to a more versatile T-DNA transformation with which transfer to *Streptomyces* may be possible. Another observation was made related to the speed of RP4-mediated transfer to *Streptomyces*. The successful transformation of *S. lividans* using *A. tumefaciens* with RP4 and pSET152*RB* was observed after 24 h only, but no transfer was detected in the first 2 h **(Table 4)**. This is in accordance with Mazodier, Petter & Thompson

(1989), showing that transfer occurs only with *Streptomyces* germ tubes that emerge from the spore around 2-6 h after germination, and not with vegetative mycelium (Jyothikumar *et al.*, 2008; Chater, 2011). The RP4-mediated transfer from *A. tumefaciens* occurs at a similar speed as pUZ8002-mediated transfer from *E. coli* (Gust, Kieser & Chater, 2002). This indicates that the length of co-cultivation beyond 48 h, as used in many AMT protocols (optimized for yeast transformation) would not be not necessary when transfer to *Streptomyces* would be attempted again. The sensitive period would be between 2-6 h of *Streptomyces* spore germination, reducing the effective co-cultivation time to within the 24 h window, thus making the method comparable to that with RP4 transformation methods.

In conclusion, the initial findings of Kelly & Kado (2002) could not be replicated. In different co-cultivation environments, no AMT to *S. lividans* was observed. With the collected data *S. lividans* could no longer be considered as a research objective to create a high-throughput screening platform to study AMT. But DNA transfer mediated by the RP4 plasmid was shown to be successful between *A. tumefaciens* and *S. lividans*, indicating that the continuation to study the comparison between both transfer systems would be highly interesting.

Materials and methods

Organisms and constructs

All used organisms are listed in **table 5**. All used plasmids are listed in **table 6**. *E. coli* DH10β was made competent and transformed using the protocol of Inoue *et al.* (1990). Transformation of *E. coli* ET12567 was done according to Chung *et al.* (1989). Introduction of plasmids into *A. tumefaciens* was done by electroporation (den Dulk-Ras & Hooykaas, 1995). Transformation of *S. lividans* was done according to Gust, Kieser & Chater (2002). All strains of *E. coli* and *A. tumefaciens* containing plasmids were cultured on appropriate antibiotic selection (**Tables 5 and 6**) on LB agar or in liquid LB medium (10 g/L bacto tryptone, 5 g/L bacto yeast extract, 8 g/L NaCl, pH 7.0). *E. coli* was cultured at 37° C, *A. tumefaciens* at 29° C and *Streptomyces* at 30° C.

Cloning

Molecular cloning was performed according to standard laboratory methods (Sambrook & Russell, 2001). Plasmids were contained and amplified in *E. coli* DH10β. The plasmid pUZ8002 was isolated from *E. coli* ET12567 (pUZ8002) using the plasmid isolation kit GeneJet Plasmid Midi-prep Kit (Fermentas).

Construction of pSET152 based shuttle vectors

The shuttle vectors used were constructed starting from the pSET152 vector, which was optimized for *Streptomyces* genomic integration using the integrase of bacteriophage ΦC31 (Bierman, Logan & O'Brian, 1992). The border sequence, *KanMX* resistance cassette and the *pVS1* minimal replicon (*sta* and *rep*) were isolated from pOphis vectors (Rolloos *et al.*, 2014). The plasmid pOphisΔ*RB* (original name: pOphis Borderless/BL) was digested with PvuII and the fragment containing a *KanMX* cassette and the *pVS1* replicon was inserted into the PvuII digested pSET152 backbone containing the $oriT_{RP4}$, *pMB1* replicon, and ΦC31 integrase gene. This created the vector pSET152Δ*RB* containing. Plasmid pOphis*RB* was digested with PvuII and the fragment containing the 25 bp RB_{Ti} sequence in combination with a upstream adjacent 38 bp OD_{Ti} sequence, *KanMX* cassette and *pVS1* replicon was as well ligated into the PvuII digested pSET152 backbone containing the $oriT_{RP4}$, *pMB1* replicon, and ΦC31 integrase gene. This created the following vector pSET152*RB*. The $oriT_{RP4}$ region (synonymously with $oriT_{RK2}$; Bierman, Logan & O'Brian, 1992) was removed cutting pSET152*RB* and pSET152Δ*RB* with ApaI and SphI. The 840 bp fragment containing only the $oriT_{RP4}$ was separated from the backbone fragment. The $oriT_{RP4}$ sequence on pSET152 was itself based on the 760 bp HaeII fragment of the RP4 plasmid (Guiney & Yakobson, 1983). The ApaI and SphI overhangs of the linear plasmids were blunted using T4 DNA polymerase (Thermo Scientific) and ligated with T4 DNA ligase (Thermo Scientific) creating the plasmids pSET152*RB*Δ*oriT* and pSET152Δ*RB*Δ*oriT*. The plasmids pSET152*RB* and pSET152Δ*RB*, pSET152*RB*Δ*oriT* and pSET152Δ*RB*Δ*oriT* have the blunt ended PvuII fragment inserted in the same orientation as was verified via restriction digest. To test transfer to yeast pSET152*RB* was additionally equipped with the autonomous replication and centromeric region sequence *ARS6/Cen4* to allow for stable replication in yeast cells after transfer. The *ARS6/Cen4* sequence was amplified from vector pUC34CPF with primer pair 5'-AGG TAC CCC ACC TGG GTC CTT TTC ATC ACG TG and 5'-AGG TAC CTA ATG GTT TCT TAG GAC GGA TCG CTT GC, both containing KpnI sites. The KpnI digested PCR fragment was introduced in the KpnI sites of the plasmids resulting in the plasmid pSET152Y*RB* (Y: indicative of yeast compatible shuttle or mobilizable vector).

Aggregate formation assay

The aggregate formation assay between *A. tumefaciens*, *S. lividans* or *S. venezuelae* was performed as described in Kelly & Kado (2002). *A. tumefaciens* LBA1100 (pSET152*RB*) was induced prior for 5.5 h in induction medium (IM) at 29° C, containing 200 μM acetosyringone (as described by Bundock *et al.* (2005). For the coagulation experiment, *A. tumefaciens* LBA2210 (pSET152*RB*) was prior cultured for 5.5 h with a starting OD_{600} of 0.25 to reach growth in the logarithmic phase of the culture at 29° C. The combined cultures in the test tubes were observed at room temperature for coagulation at 5, 15, and 30 min, as well as after 12 h.

Co-cultivation experiments

AMT co-cultivation of *A. tumefaciens* to *S. lividans* was essentially as performed after Kelly & Kado (2002), using R2 medium (Kieser *et al.*, 2000) as well as R5 medium (Kieser *et al.*, 2000) both containing acetosyringone as inducing agent for AMT.

Other AMT co-cultivations were performed using the protocol of Bundock *et al.* (1995). Modifications in this protocol were as follows: *A. tumefaciens* was pre-cultured over night before the co-cultivation in LB broth containing appropriate antibiotics, the culture was then diluted into a final concentration of OD_{600} of 0.25 in IM containing 200 μM acetosyringone, and further cultured whilst shaking at 29° C, 5 h prior to the co-cultivation. Sterile cellophane filters were used during *Streptomyces* co-cultivation (Kieser *et al.*, 2000). Commercially available cellophane wrapping film was used, cut into discs fitting a 90 mm petri dish, and autoclaved while submersed in H_20. Per co-cultivation, an average of 1×10^6 spores were used. Spores were heat shocked for pre-germination for 10 min at 50° C and kept for 1 h at 30° C, until mixed with *A. tumefaciens*. The length of co-cultivation varied according to the experiment, as described in the results. Co-cultivation was performed at either 21° C or 30° C, as described in the results. The biofilm was removed from the cellophane filter with a metal scraper and placed in 1 mL H_2O. The mixture was vigorously shaken until the biofilm was homogenously suspended in the solution. Selection of transformed *Streptomyces* cells was performed either on R2 medium (Kieser *et al.*, 2000), or R5 medium (Kieser *et al.*, 2000) at 30° C containing 50 μg/mL apramycin and 200 μg/mL cefotaxamine. For growth on 6-well plates, the modified protocol for optimized AMT was used **(Chapter 2)**. After 2 and 3.5 h the recovered biofilm was removed and placed directly on selective plates.

The transformation of *Streptomyces* using *E. coli* ET12567 (pUZ8002, pSET152*RB*), *E. coli* DH10β (pUZ8002, pSET152*RB*), *A. tumefaciens* LBA288 (pUZ8002, pSET152*RB*) and *A. tumefaciens* LBA2210 (RP4, pSET152*RB*) were performed according to Gust, Kieser & Chater (2002). Co-cultivation was performed at 29° C. When *S. lividans* (1×10^7 pre-germinated spores) was cultivated without an *E. coli* donor strain but just with 1 μg of isolated pSET152*RB* plasmid DNA, the co-cultivation was also performed according to Gust, Kieser & Chater (2002). Transformed *S. lividans* cells, after co-cultivation, were cultured and selected on appropriate on MS medium or R5 medium plates (Kieser *et al.*, 2000), containing 50 μg/mL apramycin and either 25 μg/mL nalidixic acid to select against *E. coli,* or 200 μg/mL cefotaxamine to select against *A. tumefaciens*. *S. venezuelae* selection was done on MYM medium (4 g/L maltose, 4 g/L yeast extract, 10 g/L malt extract and 20 g/L bacto agar) that was supplemented after autoclaving with 2.5 mL/L R2 trace element solution (Kieser *et al.*, 2000). The transformation efficiency was determined by dividing the output transconjugants CFU by the input of *Streptomyces* cells.

Verification pSET152*RB* genomic integration

To verify the genomic integration of the shuttle vector pSET152*RB*, primer Apra-INT-FW 5'-ACG CCG AGG AGA AGT ACC TG and primer Apra-INT-RV 5'-ACG CCG AGG AGA AGT ACC TG were used to detect the apramycin cassette in the *S. lividans* genome with PCR, using DreamTaq DNA polymerase (Thermo Scientific). Genome isolation was performed with the GenElute bacterial genomic DNA kit (Sigma-Aldrich), after positive cultures were isolated from co-cultivation with *E. coli* ET12567 (pUZ8002, pSET152*RB*).

Table 5. Organisms used in this study.

Strain	Description	Resistance	Reference
A. tumefaciens			
LBA288	based on WT C58, Cured of pTiC58, based on LBA201 but with spontaneous Rif and Nal; cryptic plasmid pAtC58	Rif	Hooykaas *et al.*, 1980
LBA1100	pAL1100 or octopine pTiB6 ΔT_L, ΔT_R , Δ*tra*, Δ*occ*; C58 chromosomal background; cryptic plasmid pAtC58	Rif, Spec	Beijersbergen *et al.*, 1992
LBA1126	VirG I77V, VirA-TAR, C58 chromosomal background; cryptic plasmid pAtC58	Rif, Spec	Bundock & Hooykaas, 1996
LBA2210	RP4, cryptic plasmid pAtC58, C58 chromosomal background	Rif, Km, Tc, Cb, Nal	PJJ Hooykaas (unpublished)
E. coli			
DH10β	*F- araDJ39*, *A(ara, leu)7697*, *AlacX74*, *galU*, *galK*, *rpsL*, *deoR*, *Φ80dlacZAM15*, *endAI*, *nupG*, *recA1*, *mcrA*, *A(mrr hsdRMS mcrBC)*		Grant *et al.*, 1990
ET12567	pUZ8002; *dam*$^-$, *dcm*$^-$, *hsdM*$^-$	Cm, Km	Paget *et al.*, 1999
Streptomyces			
S. lividans	66 (1326); WT		Cruz-Morales *et al.*, 2013
S. venezuelae	ATCC 10712		Pullan *et al.*, 2013

Cb: Carbenicillin; Cm: Chloramphenicol; Km: Kanamycin; Nal: Nalidixic acid; Rif: Rifampicin; Spec: Spectinomycin; Tc: Tetracycline; WT: Wild-type

Table 6. Genetic constructs used in this study.

Plasmid	*oriV*	Description	Resistance	Reference
pAL1100	*In-cRh-1*	pAL1100:ΔT_L, ΔT_R, Δ*tra*, Δ*occ*	Spec	Beijerbergen *et al.*, 1992
pSET152	*pMB1*	lacZαMCS; *ori*T_{RP4}; Int$^{\Phi C31}$	Apra	Bierman, Logan & O'Brian, 1992
pSET152*RB*	*pMB1, pVS1*	*OD*$_{Ti,Oct}$; *RB*$_{Ti,Oct}$; *ori*T_{RP4}; *KanMX*	Apra	This study
pSET152Δ*RB*	*pMB1, pVS1*	Δ*OD*$_{Ti,Oct}$; Δ*RB*$_{Ti,Oct}$; *ori*T_{RP4}; *KanMX*	Apra	This study
pSET152*RB*Δ*oriT*	*pMB1, pVS1*	*OD*$_{Ti,Oct}$; *RB*$_{Ti,Oct}$; Δ*ori*T_{RP4}; *KanMX*	Apra	This study
pSET152Δ*RB*Δ*oriT*	*pMB1, pVS1*	Δ*OD*$_{Ti,Oct}$; Δ*RB*$_{Ti,Oct}$; Δ*ori*T_{RP4}; *KanMX*	Apra	This study
pSET152Y*RB*	*pMB1, pVS1*	*OD*$_{Ti,Oct}$; *RB*$_{Ti,Oct}$; *ori*T_{RP4};*KanMX*; *ARS6/Cen4*,	Apra	This study
pUG34CFP	*pMB1*	pRUL1001; contains *ARS6/Cen4*	Cb	Sakalis, 2013
pOphis*RB*	*pVS1*	*OD*$_{Ti,Oct}$; *RB*$_{Ti,Oct}$, Δ*LB*$_{Ti,Oct}$, *KanMX*	Km	Rolloos *et al. 2014*
pOphisΔ*RB*Δ*LB*	*pVS1*	pOphis Borderless/BL; Δ*OD*$_{Ti,Oct}$; Δ*RB*$_{Ti,Oct}$; Δ*LB*$_{Ti,Oct}$; *KanMX*	Km	Rolloos *et al.*, 2014
pUZ8002	*Inc*Pα	RP4/RK2 derivative Δ*ori*T_{RP4}	Km	Paget *et al.*, 1999
RP4	*Inc*Pα	*mob*; *tra*; *ori*T_{RP4}	Km, Tc, Cb	Pansegrau et *al.*, 1994

oriV: Origin of vegetative replication (synonymously with origin of replication *oriRep* or *ori*); Apra: Apramycin; Cb: Carbenicillin; Km: Kanamycin; Spec: Spectinomycin; Tc: Tetracycline; *ori*T_{RP4} is synonymously with *ori*T_{RK2}; Oct: octopine from pTi15955; Y: *ARS6/Cen4*; *RB*$_{Ti,Oct}$: Right Border pTi$_{Oct}$; *LB*$_{Ti,Oct}$: Left Border pTi$_{Oct}$; *OD*$_{Ti,Oct}$: Overdrive pTi$_{Oct}$ (Rolloos *et al.*, 2014)

Acknowledgements

This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* **73:** 775–808

Baron C, Domke N, Beinhofer M & Hapfelmeier S (2001) Elevated temperature differentially affects virulence, VirB protein accumulation, and T-pilus formation in different *Agrobacterium tumefaciens* and *Agrobacterium vitis* strains. *J. Bacteriol.* **183:** 6852–6861

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Beijersbergen AGM (1993) Trans-kingdom promiscuity. Similarities between T-DNA transfer by *Agrobacterium tumefaciens* and bacterial conjugation. Doctoral dissertation**,** Chapter **3**: 69–80, Leiden University

Bierman M, Logan R & O'Brien K (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. *Gene* **116:** 43–49

Buchanan-Wollaston V, Passiatore JE & 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 AD, Beijersbergen AGM & Hooykaas PJJ (1995) *Trans*-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J.* **14:** 3206–3214

Bundock P & Hooykaas PJJ (1996) Integration of *Agrobacterium tumefaciens* T-DNA in the *Saccharomyces cerevisiae* genome by illegitimate recombination. *Proc. Natl. Acad. Sci. U.S.A.* **93:** 15272–15275

Chater K (2006) *Streptomyces* inside-out: a new perspective on the bacteria that provide us with antibiotics. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361:** 761–768

Chater K (2011) Differentiation in *Streptomyces*: the properties and programming of diverse cell types. In *Streptomyces: Molecular biology and biotechnology*, Dyson P (ed) pp 43–86. Norfolk, UK: Caister Academic Press

Chung CT, Niemela SL & Miller RH (1989) One-step preparation of competent *Escherichia coli*: transformation and storage of bacterial cells in the same solution. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 2172–2175

Cruz-Morales P, Vijgenboom E, Iruegas-Bocardo F, Girard G, Yáñez-Guerra LA, Ramos-Aboites HE, Pernodet J-L, Anné J, van Wezel GP & Barona-Gómez F (2013) The genome sequence of *Streptomyces lividans* 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. *Genome Biol. Evol.* **5:** 1165–1175

de Groot MJA, Bundock P, Hooykaas PJJ & Beijersbergen AGM (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* **16:** 839–842

den Dulk-Ras AD & Hooykaas PJJ (1995) Electroporation of *Agrobacterium tumefaciens*. *Methods Mol. Biol.* **55:** 63–72

Dillen W, Clercq J, Kapila J, Zambre M, van Montagu M & Angenon G (1997) The effect of temperature on *Agrobacterium tumefaciens*-mediated gene transfer to plants. *Plant J.* **12:** 1459–1463

Dostal L, Shao S & Schildbach JF (2011) Tracking F plasmid TraI relaxase processing reactions provides insight into F plasmid transfer. *Nucleic Acids Res.* **39:** 2658–2670

Escobar MA & Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* **8:** 380–386

Fernández-López C, Bravo A, Ruiz-Cruz S, Solano-Collado V, Garsin DA, Lorenzo-Díaz F & Espinosa M (2014) Mobilizable rolling-circle replicating plasmids from Gram-positive bacteria: a low-cost conjugative transfer. *Microbiol. Spectr.* **2:** 257–276

Flett F, Mersinias V & Smith CP (2006) High efficiency intergeneric conjugal transfer of plasmid DNA from *Escherichia coli* to methyl DNA-restricting *Streptomycetes*. *FEMS Microbiol. Lett.* **155:** 223–229

Fullner KJ & Nester EW (1996) Temperature affects the T-DNA transfer machinery of *Agrobacterium tumefaciens*. *J. Bacteriol.* **178:** 1498–1504

Gelvin SB (2003) *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiol. Mol. Biol. Rev.* **67:** 16–37

Gelvin SB (2006) *Agrobacterium* virulence gene induction. *Methods Mol. Biol.* **343:** 77–84

Gelvin SB, Karcher SJ & DiRita VJ (1983) Methylation of the T-DNA in *Agrobacterium tumefaciens* and in several crown gall tumors. *Nucleic Acids Res.* **11:** 159–174

González-Cerón G, Miranda-Olivares OJ & Servín-González L (2009) Characterization of the methyl-specific restriction system of *Streptomyces coelicolor* A3(2) and of the role played by laterally acquired nucleases. *FEMS Microbiol. Lett.* **301:** 35–43

Gordon JE & Christie PJ (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Grahn AM, Haase J, Bamford DH & Lanka E (2000) Components of the RP4 conjugative transfer apparatus form an envelope structure bridging inner and outer membranes of donor cells: implications for related macromolecule transport systems. *J. Bacteriol.* **182:** 1564–1574

Grant SG, Jessee J, Bloom FR & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 4645–4649

Guglielmini J, de la Cruz F & Rocha EPC (2013) Evolution of conjugation and type IV secretion systems. *Mol. Biol. Evol.* **30:** 315–331

Guiney DG & Yakobson E (1983) Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. *Proc. Natl. Acad. Sci. U.S.A.* **80:** 3595–3598

Gust B, Kieser T & Chater K (2002) PCR targeting system in *Streptomyces coelicolor* A3(2). John Innes Centre, Norwich Research Park, Colney, Norwich NR4 7UH, UK: 13–14

Hoekema A, Hirsch PR, Hooykaas PJJ & Schilperoort RA (1983) A binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature* **303:** 179–180

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Hooykaas PJJ, Roobol C & Schilperoort RA (1979) Regulation of the transfer of Ti plasmids of *Agrobacterium tumefaciens*. *J. Gen. Microbiol.* **110:** 99–109

Jyothikumar V, Tilley EJ, Wali R & Herron PR (2008) Time-lapse microscopy of *Streptomyces coelicolor* growth and sporulation. *Appl. Environ. Microbiol.* **74:** 6774–6781

Kado CI & Kelly B (2006) *Actinomycetes* (*Streptomyces lividans*). In *Agrobacterium Protocols Volume 2*, Wang K (ed) pp 395–401. Totowa, NJ: Humana Press

Kahng LS & Shapiro L (2001) The CcrM DNA methyltransferase of *Agrobacterium tumefaciens* is essential, and its activity is cell cycle regulated. *J. Bacteriol.* **183:** 3065–3075

Kelly B & Kado CI (2002) *Agrobacterium*-mediated T-DNA transfer and integration into the chromosome of *Streptomyces lividans*. *Mol. Plant Pathol.* **3:** 125–134

Kieser T, Bibb M, Buttner M, Chater K & Hopwood D (2000) Practical *Streptomyces* genetics 2nd ed. Norwich, UK: John Innes Foundation

Klapwijk PM, van Beelen P & Schilperoort RA (1979) Isolation of a recombination deficient *Agrobacterium tumefaciens* mutant. *Mol. Gen. Genet. MGG* **173:** 171–175

Krishnamohan A, Balaji V & Veluthambi K (2001) Efficient *vir* Gene Induction in *Agrobacterium tumefaciens* requires *virA*, *virG*, and *vir* Box from the same Ti plasmid. *J. Bacteriol.* **183:** 4079–4089

Kuhstoss S & Rao RN (1991) Analysis of the integration function of the *Streptomycete* bacteriophage phi C31. *J. Mol. Biol.* **222:** 897–908

Lacroix B, Tzfira T, Vainstein A & Citovsky V (2006) A case of promiscuity: *Agrobacterium*'s endless hunt for new partners. *Trends Genet.* **22:** 29–37

Lanka E & Barth PT (1981) Plasmid RP4 specifies a deoxyribonucleic acid primase involved in its conjugal transfer and maintenance. *J. Bacteriol.* **148:** 769–781

Lanka E & Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64:** 141–169

Liu C-N, Steck TR, Habeck LL, Meyer JA & Gelvin SB (1993) Multiple Copies of *virG* allow induction of *Agrobacterium tumefaciens vir* genes and T-DNA processing at alkaline pH. *Mol. Plant-Microbe Interact.* **6:** 144–156

MacNeil DJ (1988) Characterization of a unique methyl-specific restriction system in *Streptomyces avermitilis*. *J. Bacteriol.* **170:** 5607–5612

MacNeil DJ, Gewain KM, Ruby CL, Dezeny G, Gibbons PH & MacNeil T (1992) Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111:** 61–68

Mazodier P, Petter R & Thompson C (1989) Intergeneric conjugation between *Escherichia coli* and *Streptomyces* species. *J. Bacteriol.* **171:** 3583–3585

Melchers LS, Regensburg-Tuïnk TJG, Bourret RB, Sedee NJ, Schilperoort RA & Hooykaas PJJ (1989) Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *EMBO J.* **8:** 1919–1925

Merryweather a, Barth PT & Wilkins BM (1986) Role and specificity of plasmid RP4-encoded DNA primase in bacterial conjugation. *J. Bacteriol.* **167:** 12–17

Miele L, Strack B, Kruft V & Lanka E (1991) Gene organization and nucleotide sequence of the primase region of IncP plasmids RP4 and R751. *DNA Seq.* **2:** 145–162

Newman DJ & Cragg GM (2007) Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.* **70:** 461–477

Okanishi M, Suzuki K & Umezawa H (1974) Formation and reversion of *Streptomycete* protoplasts: cultural condition and morphological study. *J. Gen. Microbiol.* **80:** 389–400

Paget MSB, Chamberlin L, Atrih A, Foster SJ & Buttner MJ (1999) Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **181:** 204–211

Pansegrau W, Balzer D, Kruft V, Lurz R & Lanka E (1990) In vitro assembly of relaxosomes at the transfer origin of plasmid RP4. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 6555–6559

Pansegrau W & Lanka E (1996) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Paranthaman S & Dharmalingam K (2003) Intergeneric conjugation in *Streptomyces peucetius* and *Streptomyces* sp. strain C5: chromosomal integration and expression of recombinant plasmids carrying the chiC gene. *Appl. Environ. Microbiol.* **69:** 84–91

Parker C & Meyer R (2005) Mechanisms of strand replacement synthesis for plasmid DNA transferred by conjugation. *J. Bacteriol.* **187:** 3400–3406

Pullan ST, Chandra G, Bibb MJ & Merrick M (2011) Genome-wide analysis of the role of GlnR in *Streptomyces venezuelae* provides new insights into global nitrogen regulation in *Actinomycetes*. *BMC Genomics* **12:** 175

Rogowsky PM, Close TJ, Chimera JA, Shaw JJ & Kado CI (1987) Regulation of the *vir* genes of *Agrobacterium tumefaciens* plasmid pTiC58. *J. Bacteriol.* **169:** 5101–5112

Rolloos M, Dohmen MHC, Hooykaas PJJ & van der Zaal BJ (2014) Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **91:** 1240–1251

Sakalis PA (2013) Visualizing virulence proteins and their translocation into the host during *Agrobacterium*-mediated transformation. Doctoral dissertation, Chapter **4**:125-150, Leiden University

Sambrook J & Russell DW (2001) Molecular Cloning: A Laboratory Manual 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Sawasaki Y, Inomata K & Yoshida K (1996) *Trans*-kingdom conjugation between *Agrobacterium tumefaciens* and *Saccharomyces cerevisiae*, a bacterium and a yeast. *Plant Cell Physiol.* **37:** 103–106

Scheeren-Groot EP, Rodenburg KW, den Dulk-Ras AD, Turk SCHJ & Hooykaas PJJ (1994) Mutational analysis of the transcriptional activator VirG of *Agrobacterium tumefaciens*. *J. Bacteriol.* **176:** 6418–6426

Scholz P, Haring V, Wittmann-Liebold B, Ashman K, Bagdasarian M & Scherzinger E (1989) Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene* **75:** 271–288

Shaw CH, Ashby AM, Brown A, Royal C, Loake GJ & Shaw CM (1988) *virA* and *virG* are the Ti-plasmid functions required for chemotaxis of *Agrobacterium tumefaciens* towards acetosyringone. *Mol. Microbiol.* **2:** 413–417

Sheikholeslam SN & Weeks DP (1987) Acetosyringone promotes high efficiency transformation of *Arabidopsis thaliana* explants *by Agrobacterium tumefaciens*. *Plant Mol. Biol.* **8:** 291–298

Soltani J, Heusden GPH & Hooykaas PJJ (2008) *Agrobacterium*-mediated transformation of non-plant organisms. In *Agrobacterium from biology to biotechnology*, Tzfira T & Citovsky V (eds) pp 650–677. New York: Springer

Stachel SE, Nester EW & Zambryski PC (1986) A plant cell factor induces *Agrobacterium tumefaciens vir* gene expression. *Proc. Natl. Acad. Sci. U.S.A.* **83:** 379–383

Turk SCHJ, Melchers LS, den Dulk-Ras H, Regensburg-Tuïnk TJG & Hooykaas PJJ (1991) Environmental conditions differentially affect *vir* gene induction in different *Agrobacterium* strains. Role of the VirA sensor protein. *Plant Mol. Biol.* **16:** 1051–1059

Turk SCHJ, van Lange RP, Sonneveld E & Hooykaas PJJ (1993) The chimeric VirA-Tar receptor protein is locked into a highly responsive state. *J. Bacteriol.* **175:** 5706–5709

Vergunst AC, van Lier MCM, den Dulk-Ras AD, Grosse Stüve TA, Ouwehand A & Hooykaas PJJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 832–837

Waters VL, Hirata KH, Pansegrau W, Lanka E & Guiney DG (1991) Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of *Agrobacterium* Ti plasmids. *Proc. Natl. Acad. Sci. U.S.A.* **88:** 1456–1460

Yadav NS, Vanderleyden J, Bennett DR, Barnes WM & Chilton M-D (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 6322–6326

Yakobson E, Deiss C, Hirata K & Guiney DG (1990) Initiation of DNA synthesis in the transfer origin region of RK2 by the plasmid-encoded primase: detection using defective M13 phage. *Plasmid* **23:** 80–84

Zhu H, Sandiford SK & van Wezel GP (2014) Triggers and cues that activate antibiotic production by *Actinomycetes*. *J. Ind. Microbiol. Biotechnol.* **41:** 371–386

Zyprian E & Kado CI (1990) *Agrobacterium*-mediated plant transformation by novel mini-T vectors in conjunction with a high-copy *vir* region helper plasmid. *Plant Mol. Biol.* **15:** 245–256

Chapter 4

Identification of a C-terminal translocation signal in the TraI relaxase protein of the RP4 plasmid

Felix T. Wittleben, Silvia Lang, Paul J.J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Abstract

The conjugative plasmid RP4 with its type-IV secretion system ($T4SS_{RP4}$) is widely used in the laboratory to transfer DNA constructs between organisms. These DNA molecules are transferred to recipient cells in a single-strand DNA form with the RP4 relaxase $TraI_{RP4}$ covalently attached to the 5'-end. Although the main molecular elements of this conjugative transfer system have been described, no adequate descriptions of potential translocation signals in the relaxase have been published. In this study, the translocation signal (TS) of the DNA present in $TraI_{RP4}$ was identified using a modified Cre Reporter Assay for Translocation (CRAfT). The TS was determined to be located in the very C-terminus of the $TraI_{RP4}$ relaxase. The TS shares similarity with several previously published TS used by related T4SS; it has characteristic arginine residues and a similar hydropathic profile. The addition of the isolated $TraI_{RP4}$ translocation signal to a heterologous protein was sufficient to facilitate their delivery via the $T4SS_{RP4}$, from *Escherichia coli* to *E. coli*, and from *Agrobacterium tumefaciens* to *E. coli* even in the absence of conjugal DNA transfer.

Introduction

Conjugation allows the transfer of genetic material between bacteria. One type of conjugational transfer system employs the type-IV secretion system (T4SS) as a conduit. The T4SS can transfer DNA and (effector) proteins to other bacteria and, in some cases to eukaryotes. DNA molecules are transferred because they are covalently bound by a relaxase protein at their 5'-end (van Kregten *et al.*, 2009). In general, a conjugational system consists of the following components: a *trans*-membrane spanning protein complex for mate pair formation (Mpf) and facilitating nucleoprotein transfer, most commonly the T4SS, ending in an extracellular pilus structure; a DNA processing machinery (Dtr) encompassing a relaxase/nickase facilitating plasmid processing leading to formation of a transferred nucleoprotein complex; and the 'gatekeeper' between both components, the coupling protein (CP) binding both to the substrates as well as the Mpf/T4SS and thus selecting the substrates that are transferred. The CP recognizes a translocation signal (TS) located within the relaxase and other effector proteins that are transferred. Conjugation systems may have evolved from protein translocation systems which use a T4SS and CP for protein translocation, but lack the Dtr machinery (Alvarez-Martinez & Christie 2009; Grohman *et al.*, 2018).

One of the more thoroughly studied transfer systems is the inducible $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ system found in the natural plant genetic engineer *Agrobacterium tumefaciens*, encoded on its tumor-inducing (Ti) plasmid (Hooykaas & Beijersbergen 1994; Alvarez-Martinez & Christie, 2009). Part of the Ti plasmid, called transfer DNA (T-DNA), is transferred via this $CP_{Ti}/T4SS_{Ti}$ system into recipient plant cells. The proteins making up the $T4SS_{Ti}$ are encoded by 11 virulence B genes ($virB_{Ti}$). Induction and function of the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ system are triggered by specific environmental conditions. In particular, vir_{Ti} gene expression is triggered by phenolic compounds like acetosyringone and mediated by a $VirA_{Ti}/VirG_{Ti}$ 2-component system (Brencic & Winans, 2005). Besides low nutrient availability, temperatures around 20° C and a slightly acidic pH level are important (Melchers *et al.*, 1989; Turk *et al.*, 1991). The T-DNA is defined by specific sequences called the right border (RB_{Ti}) and left border (LB_{Ti}). The relaxase $VirD2_{Ti}$, part of the inducible Dtr_{Ti} system, can recognize the border sequences and create nicks there, which is thought then to lead to displacement of one single-strand of the T-DNA called the T-strand. During the nicking reaction, $VirD2_{Ti}$ remains bound to the 5'-end at the RB_{Ti}. The nucleoprotein complex (T-strand with bound relaxase), as well as other so-called effector Vir_{Ti} proteins, are then recognized by the CP_{Ti} $VirD4_{Ti}$ and transferred via the $T4SS_{Ti}$ into the recipient cell. The T-strand bound to $VirD2_{Ti}$ with the aid of the effector Vir_{Ti} protein $VirE2_{Ti}$ is then imported via the host endogenous importing system into the nucleus (Ziemienowicz *et al.*, 2001). In the nucleus the T-strand is randomly integrated into the host genome largely determined by host endogenous factors (Bundock *et al.*, 1995). The T-DNA encoded genes converting the plant cells into tumor cells, starting unregulated cellular growth and synthesis of specific compounds called opines for the survival of *A. tumefaciens*, in effect causing disease within the affected organism (Escobar & Dandekar, 2003). The Ti plasmid is also able to mediate T-DNA transfer to other eukaryotes, for example, filamentous fungi (de Groot *et al.*, 1998) and *Saccharomyces cerevisiae* (Bundock *et al.*, 1995).

A mechanistically similar type of transfer systems can be found encoded on the conjugative plasmid RP4. The main function of the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ is the conjugative transfer of the plasmid itself and other mobilizable plasmids to recipient cells. RP4, a 60,099 bp large IncPα plasmid with a broad host range, is able transfer and replicate in a variety of Gram-negative hosts including *Escherichia coli* and *A. tumefaciens* (Jain & Srivastava, 2013). Further, RP4 can facilitate plasmid transfer to Gram-positive bacteria as well as to the eukaryote *S. cerevisiae* (Trieu-Cuot *et al.*, 1987; Bates, Cashmore & Wilkins, 1998). Under laboratory conditions, transfer of RP4 itself or mobilization of other plasmids does not require a specific induction of the transfer system. One of the key elements of the RP4 mediated conjugative

transfer is the relaxase $TraI_{RP4}$ (**Figure 1A**), part of the Dtr_{RP4} system, encoded in the *tra1*$_{RP4}$ region (Pansegrau *et al.*, 1994). Guided by several system-specific auxiliary proteins $TraH_{RP4}$, $TraJ_{RP4}$ and $TraK_{RP4}$ (Pansegrau & Lanka; 1996A), $TraI_{RP4}$ recognizes a 6 bp sequence within the *oriT*$_{RP4}$ (origin of transfer) of the RP4 plasmid, where it introduces a single-strand nick and binds covalently to the 5'-end via a tyrosine residue present at position 22 by means of a transesterification reaction (**Figure 1A**; Pansegrau, Schröder & Lanka, 1993; Pansegrau, Schröder & Lanka, 1994; Balzer, Pansegrau & Lanka, 1994). After the nick at the *oriT*$_{RP4}$, a single-strand of the plasmid is presumably displaced via rolling-circle-type processing (Balzer, Pansegrau & Lanka 1994) and then recognized by the CP_{RP4} $TraG_{RP4}$ and transferred via the RP4 pilus structure $T4SS_{RP4}$ (Pansegrau *et al.*, 1994). Further, $TraI_{RP4}$ contains a possible $TraH_{RP4}$ interaction domain on the C-terminus (Pansegrau & Lanka, 1996B). $TraH_{RP4}$ is thought of stabilizing the protein-protein interactions of the relaxosome, thereby aiding in the nicking of the border sequences, without binding to DNA itself (Pansegrau & Lanka, 1996B).

Despite structural similarity and ability to mediate *trans*-kingdom transformation, there are several differences between the transfer as mediated by Dtr/CP/T4SSs of the Ti plasmid and the RP4 plasmid. In the Ti system, the substrate T-DNA contains two border sequences for relaxase nicking, defining a specific linear sequence. The linear T-strand, once it is transferred, has no intrinsic feature for double-strand formation. There are also no T-strand specific effector proteins present to initiate double-strand formation. The T-strand does not contain an origin of vegetative replication (*oriV*), and thus can only be stably maintained once it is integrated into the host genome (Gordon & Christie, 2014). In the RP4 system, the transferred DNA is a complete plasmid containing an *oriV* allowing it to replicate in a wide range of bacterial hosts (Pansegrau *et al.*, 1994). RP4 DNA contains only one motif for relaxase nicking, the *oriT*$_{RP4}$, as well as a DNA primase binding site for the primase $TraC_{RP4}$, initiating DNA double-strand formation in the recipient cell (Lanka & Fürste, 1984). The primase $TraC_{RP4}$ is also transferred via the $CP_{RP4}/T4SS_{RP4}$ into the recipient cell (Rees & Wilkins, 1990). As mentioned, both systems share a relatedness regarding the composition of their Dtr, CP and T4SS, in particular, their analogous relaxases (Scheiffele, Pansegrau & Lanka, 1995; Garcillán-Barcia, Francia & de La Cruz, 2009). The relaxase $VirD2_{Ti}$ contains an N-terminal relaxase domain (Vogel & Das, 1992) and has a TS located at the C-terminal region (van Kregten *et al.*, 2009). The relaxase $TraI_{RP4}$ also has the catalytic domain located at the N-terminus (**Figure 1A**; Pansegrau & Lanka, 1996A), as well as the TS in the C-terminus (**Figure 1A**; Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995). It is to note that both relaxases $VirD2_{Ti}$ and $TraI_{RP4}$ are essential for DNA transfer via their respective CP/T4SS (Balzer, Pansegrau & Lanka, 1994; van Kregten *et al.*, 2009). Translocation of various proteins independent of DNA transfer has been shown to occur with the $CP_{Ti}/T4SS_{Ti}$ (Vergunst *et al.*, 2000; Schrammeijer *et al.*, 2003; Vergunst *et al.*, 2005; van Kregten *et al.*, 2009). With the $CP_{RP4}/T4SS_{RP4}$, protein translocation into the recipient cell has thus far only been inferred for the primase $TraC_{RP4}$ (Rees & Wilkins, 1990).

One striking phenomenon that was discovered for the $CP_{Ti}/T4SS_{Ti}$ was its ability to translocate proteins without any concomitant DNA transfer. This occurs naturally in specialized transfer systems also containing a CP/T4SS that are known to only translocate effector proteins, for example, the CP/T4SSs of specialized human pathogens as *Heliobacter pylori* and *Legionella pneumophilia* (Nagai & Roy, 2003; Cascales & Christie, 2004). Naturally, as described above, the $CP_{Ti}/T4SS_{Ti}$ is transferring T-DNA combined with effector proteins as well as effector proteins mechanistically independent of DNA transfer. In 2000, Vergunst *et al.* were able to prove that the effector protein $VirF_{Ti}$ could be transferred via the $CP_{Ti}/T4SS_{Ti}$ independent of T-DNA transfer into plant cells. It was also shown that the TS of the effector proteins $VirD2_{Ti}$, $VirF_{Ti}$, and $VirE2_{Ti}$ could be used for translocation *in vivo* (Schrammeijer *et al.*, 2003; van Kregten *et al.*, 2009). Until now, the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ has only been studied under the aspect of being a conjugative DNA delivery system (Ditta *et al.*,

1980; Simon, 1984; Wiater, Marra & Shuman, 1994; Paget *et al.*, 1999; Babic, Guérout & Mazel 2008, Strand *et al.*, 2014).

This study explored potential similarities between the $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ and $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ and thereby investigated whether the relaxase $TraI_{RP4}$ did possess a discrete TS that directed its transport via the $CP_{RP4}/T4SS_{RP4}$. Evidence for a $TraI_{RP4}$ TS was found located at the relaxase C-terminus, and in addition it could be demonstrated for the first time that protein translocation between bacteria via the $CP_{RP4}/T4SS_{RP4}$ was possible independent of DNA transport. The implications and possible use of these observed phenomena are discussed within.

Results

$TraI_{RP4}$ TS is situated between the amino acids 625 and 732 of the protein

So far, no TS domain has been defined for the relaxase $TraI_{RP4}$; therefore, this study aimed to locate and characterize it. The position of the TS on the relaxase $TraI_{RP4}$ was determined by a modified Cre Reporter Assay for Translocation (CRAfT) assay. In this assay $TraI_{RP4}$ segments possibly containing a TS were expressed as a C-terminal fusion with a Cre recombinase protein encoded on an expression plasmid. When a CP_{RP4}/$T4SS_{RP4}$ compatible TS would be present, the fusion protein could then be transported from *E. coli* DH5α cells containing an RP4 plasmid into the recipient strain *E. coli* CSH26Cm:LTL, harbouring a Cre reporter target cassette integrated into the genome. The Cre reporter target cassette contains a tetracycline resistance cassette (*tet*) flanked by *loxP* sites, which is embedded within a chloramphenicol resistance cassette (*cat*). After Cre-mediated excision, the *loxP* sites are re-ligated, leaving only a 6 bp in-frame scar, and thereby creating a functional chloramphenicol resistance cassette. Successful fusion protein transfer can therefore be detected by a change from tetracycline to chloramphenicol antibiotic resistance. Several parts of the 732 amino acids (AA) $TraI_{RP4}$ protein sequence were fused to the C-terminus of the Cre recombinase. A first indicative CRAfT assay implied that a $TraI_{RP4}$ TS was located in the C-terminus, between 532-732 AA and a second one between 243-515 AA (**Figure S2**). Subsequent CRAfT assays with further truncations of the C-terminus were performed to narrow down the precise location of the TS (**Figure 1B**). In these experiments the fusion protein Cre::$TraI_{RP4}$ 625-732 AA demonstrated the most efficient translocation, indicating that a TS is located within the C-terminal 625-732 AA sequence. The translocation efficiency, as measured by the CRAfT assay, was higher than observed after fusion to the full-length $TraI_{RP4}$ 1-732 AA sequence, or the 243-732 AA sequence (lacking the relaxase domain) fused to Cre. The fusion constructs with the $TraI_{RP4}$ 1-531 AA fragment containing the relaxase domain as well as the middle part of the protein, but excluding the $TraI_{RP4}$ C-terminal part, exhibited a low but clearly detectable translocation frequency, indicative of the presence of a second TS between 243-415 AA (**Figure 1B**). The Cre::$TraI_{RP4}$ 599-679 AA and Cre::$TraI_{RP4}$ 532-632 AA fragments showed no evidence for transfer (**Figure 1B**), as the numbers of resistant colonies did not exceed those in negative controls (Cre protein only) of the comparable CRAfT assays in **figure S2**. When comparing the truncation 599-679 AA (no transfer) with the truncation 625-732 AA (transfer), it becomes apparent that the TS is likely to be located in the part of 680-732 AA. These would be the last C-terminal 53 AA. In conclusion, the CRAfT assay revealed the presence of a TS in the C-terminus of $TraI_{RP4}$ between 625-732 AA (last 107 AA of the protein), and a putative TS with lower activity located between 243-415 AA just beyond the N-terminal relaxase domain.

The $TraI_{RP4}$ TS shares a consensus arginine motif, charge and similar hydropathic profile with other T4SS TSs

The 680-732 AA C-terminal sequence (last 53 AA) of $TraI_{RP4}$ theoretically containing the translocation motif (**Figure 1B**) was compared to the C-termini of other relaxases and a primase encoded by RP4 (Miele *et al.*, 1991; Pansegrau *et al.*, 1994; Schröder & Lanka, 2005; Vergunst *et al.*, 2005; Garcillán-Barcia, Francia & de La Cruz, 2009). The following proteins were used for the C-terminal comparison: the relaxase $MobA_{RSF1010}$ of the mobilizable RSF1010 plasmid, which can be mobilized by RP4 (Haase *et al.*, 1995), with a C-terminal TS, which is recognized by the RP4 CP_{RP4} $TraG_{RP4}$. Importantly, RSF1010 can also be transferred from *Agrobacterium* by the plasmid Ti and its CP_{Ti} $VirD4_{Ti}$ (Beijersbergen *et al.*, 1992), which makes the comparison of the $MobA_{RSF1010}$ C-terminus containing the TS particularly interesting.

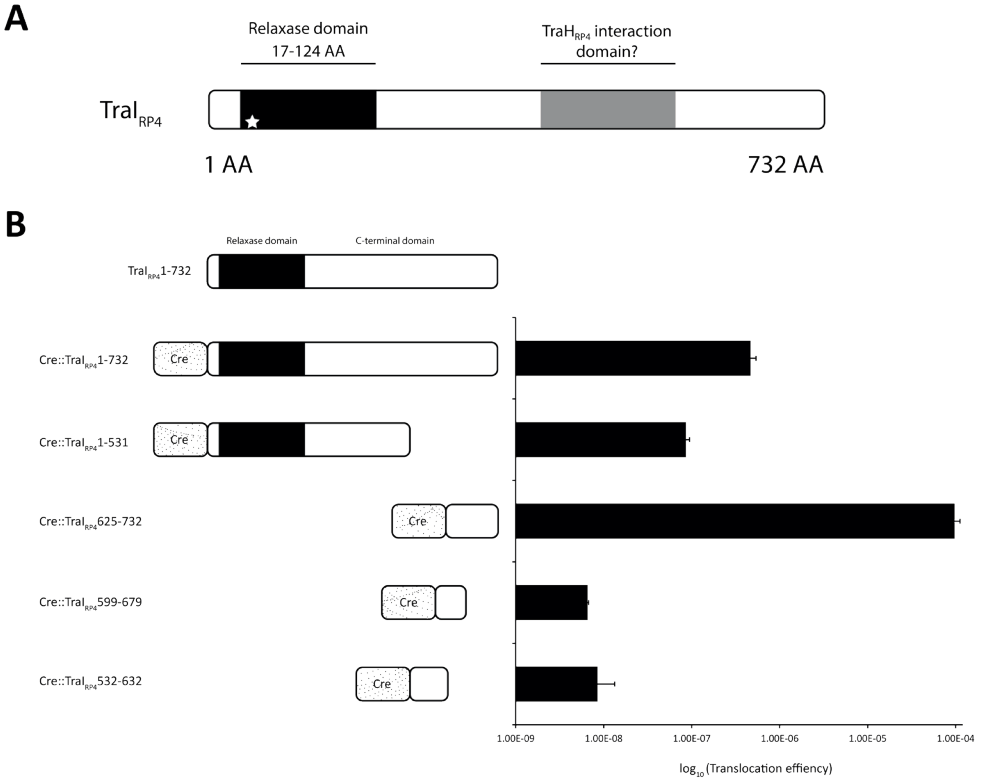


Figure 1. A: Schematic overview of the relaxase $TraI_{RP4}$. The full-length $TraI_{RP4}$ protein 1-732 AA, contains a relaxase domain situated between 17-124 AA at the N-terminus. The conserved essential tyrosine 22 forming the covalent bond with DNA at the *ori*T_{RP4}, is marked with a white star (Pansegrau & Lanka, 1996A). The $TraH_{RP4}$ interaction domain is potentially located at the C-terminus (Pansegrau & Lanka, 1996B). **B:** Characterization of the TS of the relaxase $TraI_{RP4}$ using a CRAfT assay. Cre recombinase was fused to $TraI_{RP4}$ truncations (indicated on the left) and was tested for transfer via RP4 CP_{RP4}/$T4SS_{RP4}$ from *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL. The x-axis (log_{10} scale) depicts the translocation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins, and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1**. Error bars depict SEM (n = 3).

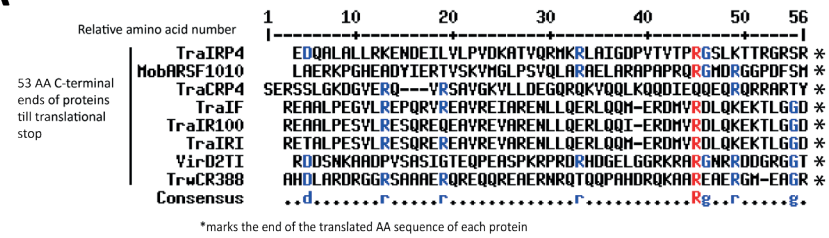
The TS of $MobA_{RSF1010}$ has been described in detail by Meyer (2015). Additionally, the relaxase $VirD2_{Ti}$ was compared, encoded on the Ti plasmid, described and characterized in more detail by Vergunst *et al.* (2005) and van Kregten *et al.* (2009). Furthermore, the relaxase of the conjugative plasmid R388 $TrwC_{R388}$ (Fernández-López *et al.*, 2006), the related relaxases $TraI_F$ of plasmid F, $TraI_{R100}$ of plasmid R100 and $TraI_{R1}$ of plasmid R1 (Cox & Schildbach, 2017) and the RP4 primase $TraC_{RP4}$ were included in the comparison to the C-terminus of $TraI_{RP4}$ (Miele *et al.*, 1991). An earlier study (Vergunst *et al.*, 2005), looking at the characteristics of the virulence proteins translocated via the CP_{Ti}/$T4SS_{Ti}$, described conservation of arginine residues in the TS, which were located in the last 30 AA of $VirF_{Ti}$, $VirE2_{Ti}$, $VirE3_{Ti}$, $VirD2_{Ti}$ and $VirD5_{Ti}$. When comparing the $TraI_{RP4}$ C-terminal end with the C-terminal ends (last 53 AA) of the relaxases and the primase mentioned above (**Figure 2A**), an arginine residue (R721) of $TraI_{RP4}$ appears to be conserved in the other relaxases but not in the primase. A second less conserved arginine (R709) in $TraI_{RP4}$ appears to be in consensus with $MobA_{RSF1010}$ and

$VirD2_{Ti}$ **(Figure 2A)**. Another important characteristic of CP/T4SS TSs was the net positive charge of the presumed TS domains, primarily due to the present arginine residues (Vergunst *et al.*, 2005; Alperi *et al.*, 2013; Meyer 2015). In the case of $TraI_{RP4}$, the C-terminal sequence contains 13.2% arginine residues (**Figure 2B**), which seems to be an enrichment over a normal occurrence of 5%, when each of the 20 AA residues would occur with equal chance. A rather similar arginine enrichment was found in other C-terminal sequences of related relaxases (**Figure 2B**). To further illustrate this, the frequency of occurrence of arginine residues in the C-terminal 30 AA sequence of $VirF_{Ti}$ TS sequence as well for $VirD2_{Ti}$ TS are around 20% to 30% compared to an average arginine frequency in *A. tumefaciens* of around 6.64% (Vergunst *et al.*, 2005). For 53 AA, statistically significant enrichment over an expected 5% to 6.64% using a chi-squared test is already reached when observing 7 to 8 arginine residues, which is the case with all C-terminal sequences shown (**Figure 2B**). Moreover, as another characteristic feature Vergunst *et al.* (2005) argued for the importance of a total positive charge of the TS for efficient translocation of the protein. This is also the case for the TS sequence of $TraI_{RP4}$ as it exhibited an all-over positive charge, just as its closest related relaxases $MobA_{RSF1010}$, $VirD2_{Ti}$, and $TrwC_{R388}$ as well as for the primase $TraC_{RP4}$ (**Figure 2B**). The more distantly related relaxases TraI of F, R1 and R100, which are in between themselves very related, exhibited a negative total charge in their C-terminal sequence (**Figure 2B**). Another factor assumed important for translocation is the hydrophilicity of the TS; reduction of hydrophilic properties of the $CP_{Ti}/T4SS_{Ti}$ TS strongly decreased translocation (Vergunst *et al.*, 2005). The 53 AA $TraI_{RP4}$ C-terminal sequence containing the TS is slightly hydrophilic, although mainly at the very terminal end. Comparative analysis of hydrophobicity plots (**Figure 2C**) showed the same trend for the $TraI_{RP4}$ C-terminal as for the other relaxase C-terminal ends as well as the primase C-terminus; the last 14 AAs of the terminal ends are all hydrophilic, with the conserved arginine R721 of $TraI_{RP4}$ being part of this sequence. Summarizing the results, the 53 AA C-terminal sequence of $TraI_{RP4}$ showed clear characteristics of a previously described relaxases $VirD2_{Ti}$ and $MobA_{RSF1010}$ TS motif that seems also to be present in other related relaxases as well as in the primase of RP4.

Protein translocation via the $CP_{RP4}/T4SS_{RP4}$ is independent of DNA transfer and is determined by the C-terminal TS motifs of $TraI_{RP4}$

When using the Ti plasmid for transfer, the $VirD2_{Ti}$ covalently bound to DNA is transferred into the recipient cells. Nevertheless, it was shown that $VirD2_{Ti}$ itself, in the absence of DNA transfer, could be translocated into recipient cells (Vergunst *et al.*, 2005). **Figure S2** provided evidence that the $TraI_{RP4}$ 243-732 AA fragment, without the relaxase domain and therefore not able to bind to DNA, could be translocated from cells containing RP4. It was subsequently tested using a CRAfT assay if the $TraI_{RP4}$ fragment could also be translocated from cell carrying RP4Δ*traI*, which is not transferred itself due to a deleted relaxase. In this experiment, the donor strain *E. coli* DH5α containing the RP4Δ*traI* plasmid and an expression plasmid expressing a Cre::$TraI_{RP4}$ 243-733 AA were used. The latter protein sequence has a deletion of the relaxase domain, which is located within the N-terminal 1-242 AA and therefore DNA transfer is prohibited from this donor. As a recipient strain *E. coli* CSH26Cm:LTL was used. While no DNA transfer occurred to the recipient as expected (**Figure S4B**), evidence for translocation of the Cre::$TraI_{RP4}$ 243-733 AA protein into recipient cells was found (**Figure 3A**). To further corroborate lack of RP4Δ*traI* plasmid transfer in the absence of any complementing full-length $TraI_{RP4}$, 5 to 10 randomly picked chloramphenicol-positive colonies were tested in each assay and always found to be sensitive to the RP4Δ*traI* based kanamycin marker, proving that RP4Δ*traI* plasmid transfer had not occurred. To demonstrate the functionality of RP4Δ*traI*, Cre fused to a full-length $TraI_{RP4}$ was used to substitute for the relaxase function. This fusion protein was translocated (**Figure 3A**), and it complemented the RP4Δ*traI* plasmid

A



B

TS (53 AA total)	Total charge	Arginine composition
$TraI_{RP4}$	+4	13.2% (7*)
$TraC_{RP4}$	+3	14.6% (8*)
$VirD2_{Ti}$	+3	18.9% (10*)
$MobA_{RSF1010}$	+2	13.2% (7*)
$TrwC_{R388}$	+3	22.6% (12*)
$TraI_{F}$	-1	17.0% (9*)
$TraI_{R100}$	-3	15.1% (8*)
$TraI_{R1}$	-2	17.0% (9*)

*total number of arginine residues in the 53 AA sequence

C

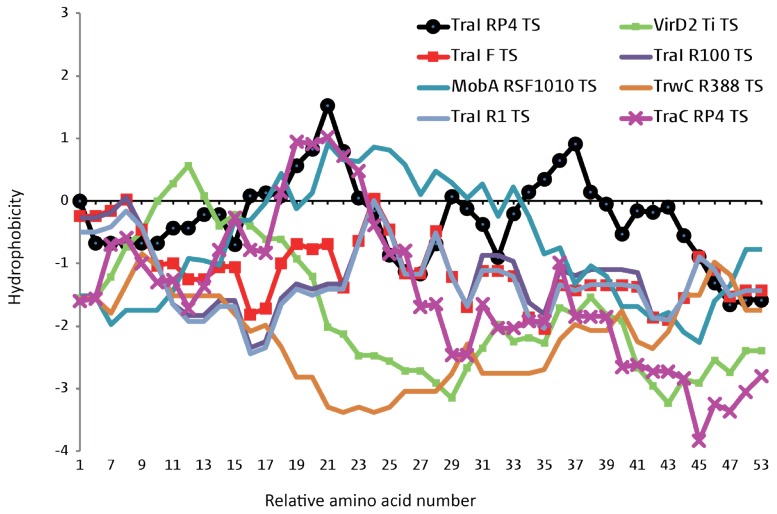


Figure 2. A: Alignment of the end C-terminal sequences (last 53 AA) of different CP/T4SS mobilizable proteins. The C-terminal 53 AA of $TraI_{RP4}$, potentially containing the TS, was compared to the C-terminal region of the same length of the relaxase $MobA_{RSF1010}$, $VirD2_{Ti}$, $TraI_{F}$, $TrwC_{R388}$, $TraI_{R100}$ and $TraI_{R1}$, and the primase $TraC_{RP4}$. High consensus AA are highlighted in a capital read letter, and lower consensus AA are highlighted in blue lowercase letters. The end of the translated AA sequence is indicated with an asterisk. The protein alignment was performed using the online Multalin software (http://multalin.toulouse.inra.fr/multalin). **B:** Comparison of the total charge of the 53 AA C-terminal residues of different relaxases and a primase translocated by a CP/T4SS and their arginine composition. The arginine composition in percent is based on the relative normal occurrence of all different essential 20 AAs in the sequence with equal chance. **C:** Kyte-Doolittle hydrophobicity plot. Comparison of the 53 AA C-termini of different relaxases and a primase translocated by a CP/T4SS. The x-axis shows the hydrophobicity index of the plot. Positive values indicate a potential hydrophobic part of the sequence (Kyte & Doolittle, 1985). The y-axis shows the 53 AAs of each C-terminus. Calculation of the Kyte-Doolittle index numbers for each C-terminal residue was performed using the software CLC Genomics Workbench 7 (https://www.qiagenbioinformatics.com).

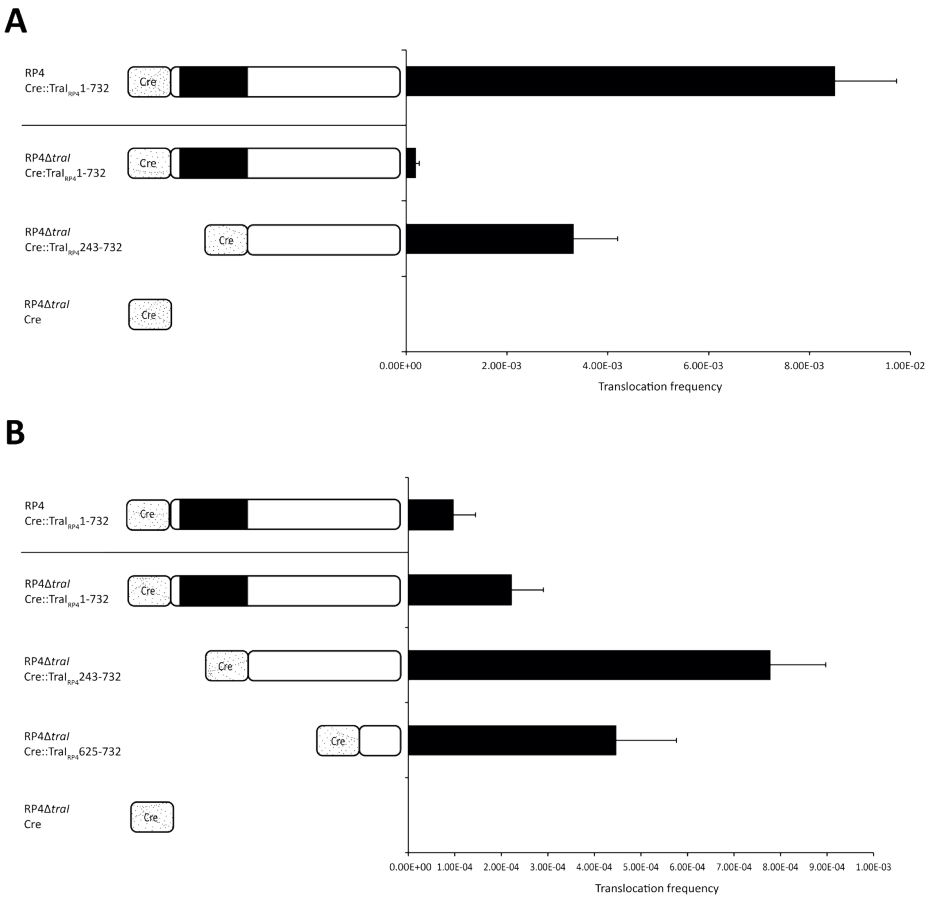


Figure 3. A: Protein translocation activity mediated by the TS of $TraI_{RP4}$ between E. *coli* via $CP_{RP4}/T4SS_{RP4}$ without conjugative DNA transfer using a CRAfT assay. The donor strain was *E. coli* DH5α and the recipient strain was *E. coli* CSH26Cm:LTL. The x-axis depicts the translocation frequency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1.** Error bars depict SEM (n = 3). **B:** Protein translocation activity mediated by the C-terminal 625-732 AA of $TraI_{RP4}$ between *E. coli* via $CP_{RP4}/T4SS_{RP4}$ without conjugative DNA transfer using a CRAfT assay. As donor strain *E. coli* DH5α and recipient strain *E. coli* CSH26Cm:LTL was used. The x-axis depicts the transformation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids. The Cre protein box is not in size relation to the $TraI_{RP4}$ depiction. The $TraI_{RP4}$ truncations are in relative size to each other. The translocation frequencies are provided in detail in **figure S1.** Error bars depict SEM (n = 3).

for full plasmid transfer. As negative control, the Cre protein by itself was expressed in the strain containing RP4Δ*traI*, but from that strain, neither protein translocation nor plasmid transfer was detected (**Figure 3A**). To narrow down the TS motif requirements for protein-only translocation via CP_{RP4}/$T4SS_{RP4}$, these experiments were repeated and included a fusion protein where Cre was fused with the C-terminal TS of $TraI_{RP4}$ described in **figure 1B**. This 625-732 AA C-terminal fragment of $TraI_{RP4}$ fused to Cre was indeed also transferred from strains containing RP4Δ*traI* (**Figure 3B**). The frequency of transfer of Cre::$TraI_{RP4}$ 625-732 AA was comparable to transfer of the larger fusion protein only lacking the relaxase domain (**Figure 3B**). In summary, protein translocation independent of DNA transfer is possible via the CP_{RP4}/$T4SS_{RP4}$, and only the TS present in the small C-terminal part consisting of 625-732 AA of $TraI_{RP4}$ is sufficient for effective protein translocation.

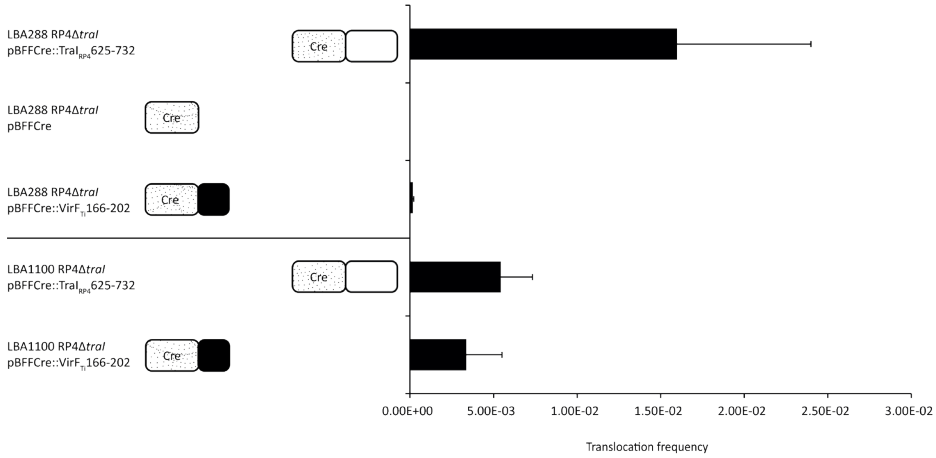


Figure 4. Protein translocation activity mediated by the C-terminal 625-732 AA of $TraI_{RP4}$ and the TS of $VirF_{Ti}$ via the CP_{RP4}/$T4SS_{RP4}$ from *A. tumefaciens* to *E. coli* without conjugative DNA transfer using a CRAfT assay. As donor strains, *A. tumefaciens* LBA288 (no Ti plasmid) and LBA1100 (containing a Ti plasmid) and recipient strain *E. coli* CSH26Cm:LTL were used. The x-axis depicts the translocation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids and donor strains. The grey dotted box is the Cre protein, the white box is the C-terminal part of $TraI_{RP4}$, and the black box is the TS for $VirF_{Ti}$. The size of the boxes is not in relation to the real size of the protein constructs and to each other. The translocation frequencies are provided in detail in **figure S1.** The error bars depicted are SEM (n = 3).

RP4 can deliver proteins, independent of DNA transfer, from *A. tumefaciens to E. coli*

The conjugative plasmid RP4 has a broad host range and is able to transfer into a large variety of bacteria. To explore the ability of RP4 to translocate proteins independent of DNA transfer into different bacterial species, translocation from *A. tumefaci*ens to *E. coli* was tested with CRAfT assays using Cre proteins fused with the short 625-732 AA $TraI_{RP4}$ sequence that was shown to contain a functional TS motif. Transfer requires not only the $T4SS_{RP4}$, but also the CP_{RP4} $TraG_{RP4}$, which selects the proteins to be transferred by their specific TS. In order to check for the selectivity of the CP_{RP4} $TraG_{RP4}$ not only to transfer protein constructs with the RP4 TS, but also protein constructs with the C-terminal 166-202 AA TS motif of the Ti

plasmid effector protein $VirF_{Ti}$ (Vergunst *et al.*, 2000; van Kregten *et al.*, 2009) was evaluated. Transfer of $VirF_{Ti}$ to recipient cells via the $T4SS_{Ti}$ normally requires the CP_{Ti} $VirD4_{Ti}$. The expression plasmid for Cre-fusion protein was introduced into *A. tumefaciens* donor strain LBA288 containing RP4Δ*traI* but lacking a Ti plasmid**.** In this set-up, no DNA transfer is possible and protein transfer can only occur via the CP_{RP4}/$T4SS_{RP4}$. As recipient strain *E. coli* CSH26Cm:LTL was used. Strains were co-cultivated at 30° C on rich LB medium. For expression of Cre fusion proteins in *A. tumefaciens*, the previously assembled expression plasmid pBFF (**Table 2**) was used. The expression cassette on pBFF is under the control of a $virF_{Ti}$ promoter, and protein sequences are fused with an N-terminal FLAG tag and SV40 NLS (van Kregten *et al.*, 2009). Although the natural $virF_{Ti}$ promoter is activated only under inducing conditions for the $VirA_{Ti}$/$VirG_{Ti}$ signalling pathway (Melchers *et al.*, 1990), van Kregten *et al.* (2009) already described that the $virF_{Ti}$ promoter construct in pBFF was active in rich LB medium at 30° C, which was verified for this promoter (**Figure S3**). CRAfT assays showed that Cre::$TraI_{RP4}$ 625-732 AA was transferred from *A. tumefaciens* LBA288 containing RP4Δ*traI* to *E. coli* (**Figure 4**). In contrast, from the same donor strain Cre::$VirF_{Ti}$ TS transfer was not seen, revealing the selectivity of the CP_{RP4}.

Discussion

This study focused on the characterization of the TS of the relaxase $TraI_{RP4}$ of the conjugative plasmid RP4. While the $Dtr_{RP4}/CP_{RP4}/T4SS_{RP4}$ is a highly efficient system to transfer large size plasmids between cells, the related $Dtr_{Ti}/CP_{Ti}/T4SS_{Ti}$ of *A. tumefaciens* Ti plasmid can be used for either DNA delivery or protein only transfer (Schrammeijer *et al.*, 2003, Vergunst *et al.*, 2005). Here the position of a discrete TS motif of the relaxase $TraI_{RP4}$ is described. In addition, the TS motif fused to a Cre protein was shown to be able to mediate Cre translocation via the $CP_{RP4}/T4SS_{RP4}$, independent of transfer of a DNA substrate.

The main TS motif of the relaxase $TraI_{RP4}$ was shown to be situated on its C-terminus between 680-732 AA **(Figure 1B)**. This is congruent with the TS position of other relaxases, including $VirD2_{Ti}$, (Vergunst *et al.*, 2005; van Kregten *et al.*, 2009) and $MobA_{RSF1010}$ (Parker & Meyer, 2007; Meyer, 2015). The 108 AA $TraI_{RP4}$ C-terminal sequence when fused to a Cre protein was able to mediate translocating of the Cre protein into a recipient cell via the $CP_{RP4}/T4SS_{RP4}$ (**Figures 1B, 3A, 3B and 4**). The present study also indicated that there might be a second and much weaker TS motif in the middle of the $TraI_{RP4}$ protein, between 243 and 531 AA (**Figures 1B and 3B**). Evidence for the presence of a second TS in the middle of the protein adjacent to the relaxase domain was seen in several relaxases studied so far, including $MobA_{RSF1010}$, $VirD2_{Ti}$, $TraI_R$, $TraI_F$ (Parker & Meyer, 2007; Lang *et al.*, 2010; Meyer, 2015). This is in contrast to the translocated effector proteins, which only have a C-terminal TS (Vergunst *et al.*, 2005). The TS of $TraI_{RP4}$ was further characterized and compared to other C-terminal ends of relaxases as well of the C-terminal end of the RP4 primase. Previous studies proposed that functional specificity of TSs for mediating transport via CPs and cognate T4SSs does not entirely rely on the precise primary AA sequence, but may be mainly determined by the charge and hydrophilicity of the signal (Miele *et al.*, 1991; Vergunst *et al.*, 2005 Thomas & Hecht, 2007); different substrates (effector proteins), but all with a similar profile, can therefore bind to the CP for further translocation. Yet, some AA residues seem to be shared. A conserved arginine motif was found to be present in the C-termini of several T4SS transfer system relaxases of the IncP, IncQ and IncF plasmid families, and a second arginine residue more downstream is conserved also in $MobA_{RSF1010}$ and $VirD2_{Ti}$ (**Figure 2A**). The different relaxase TSs, as well as the primase C-terminus, share a similar hydrophilic profile, particularly for the last 14 AA that are all hydrophilic **(Figure 2C)**. Furthermore, the total charge comparison emphasized the characteristics of the TSs as well as that of the primase C-terminus of the IncP and IncQ plasmid family **(Figure 2B)**. It thus seems that any specific charge, specific AA, or other characteristics cannot be a determining factor for making a C-terminal sequence suitable as CP/T4SS TS, but that it rather will be a combination of the specific local charge, hydrophilic pattern, and still a certain sequence specificity. This combination will create a TS with specificity to the cognate CP of the translocation machinery. However, even within related translocation machineries, the specificity varies and may not allow transfer of heterologous substrates, *e.g.* the $VirF_{Ti}$ TS is not transferred via $CP_{RP4}/T4SS_{RP4}$ (**Figure 4**).

As mentioned, the 108 AA sequence of $TraI_{RP4}$ clearly mediated reporter protein translocation via the $CP_{RP4}/T4SS_{RP4}$. Such translocation was observed from *E. coli* to *E. coli* (**Figures 1B, 3A and 3B**), as well as also from *A. tumefaciens* to *E. coli* (**Figure 4**). This is in line with earlier observations that truncations of $TraI_{RP4}$ without C-terminus, reduced transfer of RP4 drastically (Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995). In both studies (Cole, Lanka & Guiney, 1993; Cole & Guiney, 1995) the C-terminally truncated $TraI_{RP4}$ proteins consisted of the first 471 or 514 AAs that still mediated a reduced transfer of RP4, probably because they still contain the weaker secondary TS motif (**Figure 1B**).

Co-cultivation experiments determining conjugative transfer or protein translocation, all exhibit a large *inter*-experimental variation. For independent experiments using the same strains, constructs and parameters, a large variation in their measured

transformation efficiency was observed. Several repeats can manifest discrepancies of several magnitudes of the measured transformation efficiency, whereas measured phenomena consistently manifest, but are also bound to this variation. It is difficult to pinpoint the exact cause of this variation (*e.g.* temperature fluctuations, nutrient fluctuation, growth inhibition), but it is a common occurrence within this technique (author personal communication). The variation does not affect the main conclusion of each experiment. One example would be the discrepancy between the relative translocation efficiency of the full-length $TraI_{RP4}$ fused with Cre transferred with RP4 between **figure 1B** (translocation efficiency: $(4.6 \pm 1.4) \times 10^{-7}$) and **figures 3A** and **3B** (transformation efficiency: $(8.5 \pm 1.2) \times 10^{-3}$ and $(9.6 \pm 4.8) \times 10^{-5}$ respectively), that could be explained by the all-over fluctuations of a given conjugative CRAfT assay. The seemingly lower full-length $TraI_{RP4}$ translocation efficiency in **figure 1B**, should not be used in this case as an indication for its all-over transfer rate, but only should be viewed as indication for the presence of a TS within the $TraI_{RP4}$ sequence.

Some CP/T4SSs have the ability to translocate a relaxase and/or effector proteins independent of DNA transfer (Vergunst *et al.*, 2000). Experimentally, this was demonstrated by using CRAfT assays in which the candidate proteins were fused to the Cre protein and translocation could be monitored through a site-specific recombination event on a floxed substrate in the recipient cell. Thus it was shown that the TS of the relaxase $VirD2_{Ti}$ and effector proteins such as $VirF_{Ti}$ was enough to translocate the Cre protein independently of DNA transfer via the $CP_{Ti}/T4SS_{Ti}$ (Vergunst *et al.*, 2000; Vergunst *et al.*, 2005); the relaxase $TrwC_{388}$ could be translocated without DNA transfer via the $CPR_{388}/T4SS_{R388}$ of the IncW plasmid family (Draper *et al.*, 2005); also the primase Sog_{Colb} of the IncI plasmid ColIb could be transferred independently of DNA via the $CP_{Colb}/T4SS_{Colb}$ (Wilkins & Thomas, 2000). However, in other cases *e.g.* with the $CP_{R1}/T4SS_{R1}$ of plasmid R1, protein transfer independent of DNA transfer was not seen (Lang & Zechner, 2012). It was indirectly observed that a DNA substrate coupled to the relaxase plays a role in the energetical activation of the CP as well as specific parts of the T4SS for translocation of the DNA-protein complex (Cascales & Christie, 2004; Jakubowski *et al.*, 2009). When the DNA substrate is not present, relaxase protein translocation seems inhibited in some systems, as shown for $TraI_{R1}$, which is only translocated via its $CP_{R1}/T4SS_{R1}$when bound to DNA (Lang & Zechner, 2012). In the present study, it was shown for the first time that Cre fused with $TraI_{RP4}$ **(Figures 3A and 3B)** as well as Cre fused to the $TraI_{RP4}$ C-terminal sequence 625-732 AA **(Figure 3B)** was translocated via the $CP_{RP4}/T4SS_{RP4}$ independent of DNA transfer. The protein only translocation was also shown to be functional from *A. tumefaciens* to *E. coli* **(Figure 4)**. Further *trans*-kingdom delivery of proteins via the $CP_{RP4}/T4SS_{RP4}$, *e.g.* from *A. tumefaciens* to *S. cerevisiae*, or *E. coli* to *S. cerevisiae* still needs to be tested.

The recorded phenomenon of protein only delivery via $CP_{RP4}/T4SS_{RP4}$ has potential biotechnological applications using a modified RP4 plasmid as protein delivery tool. Currently, no comparable RP4 based protein delivery system exists, and so far, little literature is available on alterations and specifications using RP4 as a potential biotechnological tool beyond just DNA delivery (Babic, Guérout & Mazel, 2008). RP4 transfer is highly efficient in *E. coli,* and can be easily implemented and handled in laboratory conditions. RP4 can be easily isolated and modified, using standard molecular biological methods. In addition, the $T4SS_{RP4}$ pilus structure does not require specific induction conditions, and can polymerize within a wide temperature range. Further, loose contact of an *E. coli* donor cell to a recipient cell is sufficient for translocation of a protein (Samuels, Lanka & Davies, 2000). No induction and loose cell contact means, that simple and easy to use co-cultivation protocols can be established for protein translocation. In order to create an RP4-based delivery tool, the RP4 plasmid can modified. For example, when $oriT_{RP4}$ is be removed inhibiting the self-transfer of RP4, in combination with a second plasmid for fusion protein expression a versatile and powerful *inter*-bacterial T4SS protein delivery system is available.

Material and methods

Organisms and constructs

All strains used are listed in **table 1**, and plasmid constructs in **table 2**. *E. coli* DH5α, DH10β, CSH26CM::LTL and MS614 were made competent and transformed using the protocol of Inoue *et al.* (1990). *A. tumefaciens* strains were created using electroporation (den Dulk-Ras & Hooykaas, 1995) or tri-parental mating (Ditta *et al.*, 1980). All *E. coli* and *A. tumefaciens* strains used in this chapter were cultured on the appropriate antibiotic selection **(Tables 1 and 2)** on LB agar or in LB medium (10 g/L bacto tryptone, 5 g/L bacto yeast extract, 8 g/L NaCl, pH 7.0). *E. coli* was cultured at 37° C and *A. tumefaciens* at 29° C.

Molecular cloning

Standard molecular cloning procedures were performed according to established laboratory methods (Sambrook & Russell, 2001). Plasmids were contained and amplified in *E. coli* DH5α and DH10β. The plasmid RP4 and its derivatives were isolated from *E. coli* using the plasmid DNA purification kit Nucleobond PC100 (Macherey-Nagel), and from *A. tumefaciens* using the plasmid isolation kit GeneJet Plasmid Midi-prep Kit (Fermentas).

CRAfT-assay plasmids

The used plasmids and the primers used to generate inserts are described in **table 2** and the sequences of the primers are given in **table 3**.

For *E. coli* based protein expression, the following variations of the parental plasmid CFPBSm (Lang *et al.*, 2010) were created. Full-length and truncated versions of *traI*$_{RP4}$ DNA were amplified via PCR from an RP4 template, and digested with KpnI/SalI. The primers contained an opal TGA stop codon, the natural occurring stop codon of TraI$_{RP4}$, after each TraI$_{RP4}$ encoding fragment. Those fragments were then inserted in between the KpnI/SalI sites of the CFPBSm plasmid (Lang *et al.*, 2010). The constructs were verified by restriction digestion and sequencing.

For *A. tumefaciens* based expression, the parental plasmid pBFFCre(N) (van Kregten *et al.*, 2009) was used, containing a full-length Cre coding region fused at the N-terminus with a FLAG tag and SV40 NLS and fused at the C-terminus to a VirF$_{Ti}$ 166-202 AA TS. The following variations of the plasmid were created. The TraI$_{RP4}$ 625-732 AA fragment DNA sequence contained an opal TGA stop codon at its end and was amplified via PCR from an RP4 plasmid template and digested with SalI/XhoI. This fragment was then inserted into the XhoI digested plasmid pBFFCre(N) creating the plasmid pBFFCre::TraI625-732. The plasmid pBFFCre::VirF$_{Ti}$TS, was created by digesting the plasmid pBFFCre(N) with XhoI, removing a 90 bp sequence containing an N-terminal polylinker between Cre and the VirF$_{Ti}$ TS. The plasmid pBFFCre contains an oligo insertion (5′- to 3′-direction; GGC CGC TAT AGT GAT AAT TAG TAA CTA GGA ATT CAG CTG TT) in the NotI site of pBFFCre(N) supplying 3 in-frame stop codons shortly after the Cre coding sequence, preventing read through into the *virF*$_{Ti}$ TS. The constructs were verified by restriction digestions and sequencing.

CRAfT-assay

A Cre Reporter Assay for Translocation (CRAfT) assay (Vergunst *et al.*, 2000) as modified by Lang *et al.* (2010) was used to determine the translocation sequence of $TraI_{RP4}$, and to investigate protein translocation via $CP_{RP4}/T4SS_{RP4}$. The modifications to the original CRAfT assay by Lang *et al.*, (2010) were only minimal: a single recipient strain was used and different antibiotic selections due to the different selection markers present in the strains. The donor strain *E. coli* DH5α contained the expression vector CFPBSM coding for the fusion of the Cre protein and protein of interest (**Table 2**), as well as either the wild-type RP4 or the RP4Δ*traI* plasmid (see experiments). The donor strains of *A. tumefaciens* (LBA288 and LBA1100) contained the expression plasmid pBFF encoding Cre fusion proteins (**Table 2**) as well as the RP4 plasmid. See **table 1** for details of the used strains in this assay.

When *E. coli* DH5α was used as a donor with the recipient *E. coli* CSH26Cm::LTL or MS614, both strains were cultured in liquid LB medium with appropriate antibiotics, depending on constructs used, for selection for 24 h at 37° C. Then 100 µL of donor strain was transferred into 3 mL fresh LB medium without antibiotic selection and cultured for around 1 h at 37° C until the strains reached an OD_{600} of 0.2. For the recipient strain, after overnight growth, a volume with an OD_{600} of 0.2 was centrifuged to remove the medium containing antibiotics. The recipient cells were mixed with OD_{600} of 0.2 of the selected donor strain, providing a mixing ratio of 1:1. The combined cell suspension was centrifuged to remove the medium and then re-suspended in 30 µL fresh LB and transferred onto nitrocellulose filters (0.48 µm pore size) on solid LB. The co-cultivation was carried out at 37° C for 3 h.

When *A. tumefaciens* (LBA288 or LBA1100) was used as a donor for the recipient strain *E. coli* CSH26Cm::LTL, different cultivation temperature needed to be used. *A. tumefaciens* strains were cultured in liquid LB medium on rifampicin (10 µg/mL) selection for 24 h at 30° C. After incubation, 100 µL each of donor strain was transferred into 3 mL LB medium without selection until the strains reached an OD_{600} of 0.2. For the recipient strain, after overnight growth, a volume with an OD_{600} of 0.2 was centrifuged to remove the medium containing antibiotics. The recipient cells were mixed with OD_{600} of 0.2 of the selected donor strain, providing a mixing ratio of 1:1. The cell suspension was centrifuged to remove the medium, and then re-suspended in 30 µL fresh LB to be transferred onto nitrocellulose filters (0.48 µm pore size) on solid LB. The co-cultivation was carried out at 30° C for 24 h.

After co-cultivation, the filter was transferred from the agar surface into reaction tubes containing 1 mL LB, which were vortexed until the biofilm was re-suspended. A dilution series was plated onto LB agar plates with chloramphenicol (25 µg/mL) selection to determine the amount of recombinants cells. To only determine the recipient cell output, a dilution series was plated on LB agar plates with tetracycline (10 µg/mL) selection. All plates were cultivated at 37° C and scored after 24 h. The protein translocation frequencies were expressed as transconjugants per total number of recipient cells output after the co-cultivation.

RP4 and RP4Δ*traI* self-transfer was tested for each construct by transferring chloramphenicol resistant colonies onto a fresh LB agar plate containing kanamycin (100 µg/mL) and culturing it at 37° C for 24 h. The RP4 plasmid was considered as transferred when a re-streaked colony was able to grow on a kanamycin (100 µg/mL) selection plate. For RP4Δ*traI*, if a re-streaked colony was not able to grow, it was considered that the plasmid was not transferred.

C-terminal amino acid sequences of relaxases and analysis

The last 53 C-terminal amino acid sequences of the relaxases were taken from sequences from publicly available repositories: $TraI_{RP4}$ (GenBank L27758.1), $TraC1_{RP4}$ referred to as $TraC_{RP4}$ in this chapter, as the last 53 AA sequence of $TraC2_{RP4}$ is homologous to $TraC1_{RP4}$ (GenBank BN000925.1), $MobA_{RSF1010}$ (GenBank M28829.1), $VirD2_{Ti}$ (GenBank AF242881.1), $TrwC_{R388}$ (GenBank NC_028464.1), $TraI_{F}$ (GenBank NC_002483), $TraI_{R100}$ (GenBank NC_002134.1), and $TraI_{R1}$ (GenBank AY423546.1). To align and compare the protein sequences the Multalin software (http://multalin.toulouse.inra.fr/multalin) was utilized, and for the calculation of the Kyte-Doolittle index, the total protein charge calculation and arginine composition the software CLC Genomics Workbench 7 (https://www.qiagenbioinformatics.com) was used.

RP4 modification

To test protein translocation via the $CP_{RP4}/T4SS_{RP4}$ independent of DNA transfer, the relaxase $traI_{RP4}$ was removed from the RP4 plasmid using the λ-Red homologous recombination method based on Datsenko & Wanner (2000). The linear targeting DNA, used to knock-out $traI_{RP4}$ on RP4, contained an apramycin cassette flanked by *loxP* sites for later removal of the resistance cassette using Cre recombination, followed by homologous sequences of $traI_{RP4}$. To create the $traI_{RP4}$ target DNA**,** the apramycin resistance cassette (*aac(3)IV*) was amplified via PCR from plasmid pSET152 with primer ApraFW and ApraRV introducing SalI and SpeI restriction sites, respectively. The PCR product was ligated into the plasmid pJET1.2 (Fermentas), and transformants exhibiting apramycin resistance were selected. The SalI/SpeI apramycin fragment was removed from the positive selected pJET1.2 plasmids with the PCR insert, and then inserted into XhoI/SpeI digested p2G2L plasmid. The p2G2L plasmid contained a SalI/SpeI cloning site that is flanked by *LoxP*; this added the sites to be later removed by the Cre recombination. This created an apramycin resistance cassette flanked by *LoxP* sites creating the plasmidp2G2LApra. Primers that were used to amplify the *LoxP* flanked apramycin cassette, contained $traI_{RP4}$ homologous sequences; the amplified cassettes were then used as linear target DNA. Due to the length of the $traI_{RP4}$ homology, a two-step amplification protocol using two overlapping primer pairs was applied. The PCR protocol was as follows: 4 cycles of pre-amplification using the primer pair TraIFW1 and TraIRV1 at 72° C for the annealing step. Subsequently, 1/10 of the initial PCR product was transferred into a new PCR mix containing the second primer pair TraIFW2 and TraIRV2, initially with 4 cycles using a 60° C annealing temperature, which was increased to 72° C for another 31 cycles. The PCR product was purified by agarose gel electrophoresis using the Zymoclean gel DNA recovery kit. The purified PCR product was then used for disrupting the $traI_{RP4}$ locus on the RP4 plasmid. For this, the temperature-sensitive pKD46 plasmid containing the λ-Red recombinase was inserted into *E. coli* MS411 containing RP4. Of the PCR amplified targeting DNA (from plasmid p2GSLApra), 500 ng were transformed into this strain allowing for homologous recombination according to Datsenko & Wanner (2000). Insertion of the disruption sequence was checked with primers TraICheckFW, TraICheckRV, ApraCheckFW and ApraCheckRV. The apramycin cassette, now present at the $traI_{RP4}$ locus, was excised using Cre recombination. RP4 plasmids containing the disruption sequence were inserted into *E. coli* MS411 pKD-*cre*, as both plasmids contained chloramphenicol resistance. Chloramphenicol- and kanamycin-resistant transformants were selected at 30° C after 48 h, and re-streaked four times at 37° C with kanamycin selection only, to lose the pKD-*cre* vector. This created the plasmid RP4Δ*traI*. The knock-out of $traI_{RP4}$ was checked with primers TraICheckFW and TraICheckRV (**Figure S4A**). The plasmid RP4Δ*traI* was transformed into the donor *E. coli* DH5α and checked for lack of self-transfer to *E. coli* recipient MS614. No transconjugants were observed (**Figure S4B**). When RP4Δ*traI* was inserted into *E. coli* DH5α containing the

plasmid pGZ119EHTraI (containing a *traI*$_{RP4}$), plasmid transfer (pGZ119EHTraI) as well as self-transfer (RP4Δ*traI*) was observed (**Figure S4B**). For all experiments, the RP4Δ*traI* clone #1 was used (**Figure S4A**). The plasmid pGZ119EHTraI was created as follows. The full-length *traI*$_{RP4}$ was PCR amplified from plasmid RP4 containing SalI digestion sites on both ends. The SalI *traI*$_{RP4}$ fragment was inserted into SalI digested pGZ119EH plasmid, creating the plasmid pGZ119EHTraI. The plasmid was checked by restriction digest and sequencing.

Table 1. Organisms used in this study.

Strain	Description	Resistance	Reference
A. tumefaciens			
LBA288	based on WT C58 (LBA201); Rif; Nal; cured of pTiC58; cryptic plasmid pAtC58	Rif, Nal	Hooykaas *et al.*, 1980
LBA1100	LBA 288 with pAL1100 or octopine pTiB6: ΔT_L, ΔT_R, Δ*tra*, Δ*occ*; pAtC58	Rif, Nal, Spec	Beijersbergen *et al.*, 1992
E. coli			
DH5α	*F-*; *Φ80lacZΔM15*; *Δ(lacZYA-argF)*; *U169*; *recA1*; *endA1*; *hsdR17 (rK–, mK+)*; *phoA*; *supE44*; *λ– thi-1*; *gyrA96*; *relA1*		Grant *et al.*, 1990
DH10β	*F-*; *araDJ39*; *A(ara, leu)7697*; *AlacX74*; *galU*; *galK*; *rpsL*; *deoR*; *Φ80dlacZAM15*; *endAI*; *nupG*; *recA1*; *mcrA*; *A(mrr hsdRMS*, *mcrBC)*		Grant *et al.*, 1990
CSH26CM::LTL	CSH26 galK::cat::loxP-Tet-loxP	Tc	Lang *et al.*, 2010
MS411	*ilvGI*; *rfb-50*; *thi*		Lang *et al.*, 2010
MS614	*ilvG*; *rfb-50*; *thi*; *rpsL*	Sm	Lang *et al.*, 2010

Nal: Nalidixic acid; Rif: Rifampicin; Sm: Streptomycin; Spec: Spectinomycin; Tc: Tetracycline; WT: Wild-type

Table 2. Genetic constructs used in this study.

Plasmid	*oriV*	Description	Resistance	Reference
CFPBSm	*pMB1*	Cre fusion plasmid	Sm	Lang *et al.*, 2010
Cre::$TraI_{RP4}$1-732	*pMB1*	CFPBSm: $TraI_{RP4}$ 1-732 AA Primer: CreTraIfw01/CreTraIrev01	Sm	This study
Cre:: $TraI_{RP4}$1-531	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 1-531 AA Primer: FTW151/FTW152	Sm	This study
Cre:: $TraI_{RP4}$243-732	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 243-732 AA Primer: CreTraIfw02/ CreTraIrev01	Sm	This study
Cre:: $TraI_{RP4}$625-732	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 625-732 AA Primer: FTW149/FTW150	Sm	This study
Cre:: $TraI_{RP4}$599-679	*pMB1*	CFPBSm: Truncation $TraI_{RP4}$ 599-679 AA Primer: FTW147/FTW148	Sm	This study
Cre:: $TraI_{RP4}$532-632	*pMB1*	CFPBSm: truncation $TraI_{RP4}$ 532-632 AA Primer: FTW145/FTW146	Sm	This study
pBFFCre	*pBBR1*	ZN245; full-length Cre with N-terminal fused FLAG tag and SV40 NLS	Gm	van Kregten *et al.*, 2009
pBFFCre(N)	*pBBR1*	ZN239; parental plasmid for cloning; $virF_{Ti}$ promoter; contains full-length Cre with fused N-terminal FLAG tag and SV40; NLS; polylinker; $VirF_{Ti}$ 166-202 AA translocation signal	Gm	van Kregten *et al.*, 2009
pBFFCre::$VirF_{Ti}$TS	*pBBR1*	Translocation signal $VirF_{Ti}$ 166-202 AA fused to Cre, with polylinker XhoI fragment removed	Gm	BJ van der Zaal, unpublished
pBFFCre::$TraI_{RP4}$625-732	*pBBR1*	Full-length Cre fused to truncation $TraI_{RP4}$ 625-732 AA Primer: TraI625FW and TraI732RV	Gm	This study
RP4	*Inc*Pα	mob_{RP4}; tra_{RP4}; $oriT_{RP4}$; isolated from strain *A. tumefaciens* LBA2210	Km, Tc, Cb	Pansegrau *et al.*, 1994
RP4Δ*traI*	*Inc*Pα	Δ*traI*	Km, Tc, Cb	This dissertation, Chapter 5
pKD46	*pSC101*	contains the λ-Red recombinase; r*epA101ts* is temperature sensitive	Amp	Datsenko & Wanner, 2000
pKD-*cre*	*pSC101*	based on pKD46; encoding Cre recombinase	Amp	Lang *et al.*, 2014
p2G2L	*ColE1, F1*	Target DNA plasmid for λ-Red knock-out; pBSSK based; cloning site XhoI and SpeI flanked by *loxP* sites	Amp	S Lang, unpublished
p2G2LApra	*ColE1, F1*	apramycine *aac(3)IV* cassette flanked with *loxP* sites	Amp	This study
pSET152	*pMB1*	lacZαMCS; $oriT_{RP4}$; $Int^{\Phi C31}$	Apra	Bierman, Logan & O'Brian, 1992
pGZ119EH	*ColD*	$lacI^{Q}$ regulated *pTAC* promoter	Cam	Lessl *et al.*, 1992
pGZ119EHTraI	*ColD*	$lacI^{Q}$ regulated *pTAC* promoter; full-length $traI_{RP4}$	Cam	This study
pJET1.2	*pMB1*	*bla* (Ap^{R}); *eco47IR*; P_{lacUV5}; *pT7*; supplied as cloning vector blunt linearized with EcoRV		Fermentas, GenBank EF694056

oriV: Origin of vegetative replication (synonymously with origin of replication *oriRep* or *ori*); Amp: Ampicilin; Apra: Apramycin; Cam: Chloramphenicol; Cb: Carbenicillin; Gm: Gentamycin; Km: Kanamycin; Sm: Streptomycin; Tc: Tetracycline

Table 3. Oligonucleotide sequences used in this study.

FTW145 GCAGGTACCATTCGACCTGCCGCAGCCAGAAC
FTW146 GCAGTCGACTCATCTTCCTGGCCCAGAAACTC
FTW147 GCAGGTACCGCTGAAATGCTATTGCCGCGTGA
FTW148 GCAGTCGACTCAGACGTTCCGGGTGCCTGCATA
FTW149 GCAGGTACCGGAGTTTCTGGGCCAGGAAGAGG
FTW150 GCAGTCGACTCATCTACTCCTACCTCGGGTAG
FTW151 GCAGGTACCGTGATTGCCAAGCACGTCC
FTW152 GCAGTCGACTCAGCCTCCATCGCCGCCGCGA
CreTraIfw01 GCAGGTACCGTGATTGCCAAGCACGTCC
CreTraIrev01 GCAGTCGACTCATCTACTCCTACCTCGGGTAGTTTTAAGGG
CreTraIfw02 GCAGGTACCCCGAAGCTCGAAGCCCGATTCGG
TraI625FW AATTCTCGAGGGAGTTTCTGGGCCAGGAA
TraI732RV AATTGTCGACTCATCTACTCCTACCTCGGG
ApraFW: GGAAGTCGACTACGCGCAGAAAAAAAGGATCT
ApraRV: GGAAACTAGTCGGTGAGTTCAGGCTTTTTCAT
TraIFW1:CAAAGGTGATTGCCAGACGATAACTTCGTATAGCATACATTATACGAAGTTATCTCGA-CTACGCGCAGAAAAAAAGGATCTCAA
TraIFW2: CTCCTTCGTCAGTAGTTCTTGACGGCGGCGCTCAAGGGCGGCGTCGTCAAAGGTGATTGCCAGACG
TraIRV1: ACGAACTGCGAGGCAAACATAACTTCGTATAATGTATGCTATACGAAGTTATACTAGTC-GGTGAGTTCAGGCTTTTTCATATCTCATT
TraIRV2: ATCACCGACGAGCAAGGCAAGACCGAGCGCCTGGGTCACGTGCGCGTCACGAACTGCGAGGCAAAC
TraICheckFW: GTGGCGTTGAGGGTGGTTCTG
TraICheckRV: GGCGAAGATCGAAGAGAAGCAG
ApraCheckFW: ACGAATGGCGAAAAGCCGAG
ApracCheckRV: GGCGGATGCAGGAAGATCAAC

*Underscored bases are opal stop codons; all oligonucleotides are depicted in 5′-3′direction

Acknowledgements

This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Alperi A, Larrea D, Fernández-González E, Dehio C, Zechner EL & Llosa M (2013) A translocation motif in relaxase TrwC specifically affects recruitment by its conjugative type IV secretion system. *J. Bacteriol.* **195:** 4999–5006

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* **73:** 775–808

Babic A, Guérout AM & Mazel D (2008) Construction of an improved RP4 (RK2)-based conjugative system. *Res. Microbiol.* **159:** 545–549

Balzer D, Pansegrau W & Lanka E (1994) Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.* **176:** 4285–4295

Bates S, Cashmore AM & Wilkins BM (1998) IncP plasmids are unusually effective in mediating conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*: involvement of the Tra2 mating system. *J. Bacteriol.* **180:** 6538–6543

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Bierman M, Logan R & O'Brien K (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. Gene **116**: 43–49

Brencic A & Winans SC (2005) Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. *Microbiol. Mol. Biol. Rev.* **69:** 155–194

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

Cascales E & Christie PJ (2004) Definition of a bacterial type IV secretion pathway for a DNA substrate. *Science* **304:** 1170–1173

Cole SP, Lanka E & Guiney DG (1993) Site-directed mutations in the relaxase operon of RP4. *J. Bacteriol.* **175:** 4911–4916

Cole SP & Guiney DG (1995) Site-specific mutations in the TraI relaxase and upstream region of plasmid RP4. *Plasmid* **34:** 236–239

Cox KEL & Schildbach JF (2017) Sequence of the R1 plasmid and comparison to F and R100. *Plasmid* **91:** 53–60

Datsenko KA & Wanner BL (2000) One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc. Natl. Acad. Sci. U.S.A.* **97:** 6640–6645

de Groot MJA, Bundock P, Hooykaas PJJ & Beijersbergen AGM (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat. Biotechnol.* **16:** 839–842

Ditta G, Stanfield S, Corbin D & Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **77:** 7347–7351

Draper O, Cesar CE, Machon C, de la Cruz F & Llosa M (2005) Site-specific recombinase and integrase activities of a conjugative relaxase in recipient cells. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 16385–16390

den Dulk-Ras AD & Hooykaas PJJ (1995) Electroporation of *Agrobacterium tumefaciens*. *Plant Cell Electroporation Electrofusion Protoc.* **55:** 63–72

Escobar MA & Dandekar AM (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends Plant Sci.* **8:** 380–386

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Fernández-López R, Garcillán-Barcia MP, Revilla C, Lázaro M, Vielva L & de la Cruz F (2006) Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. *FEMS Microbiol. Rev.* **30:** 942–966

Garcillán-Barcia MP, Francia MV & de La Cruz F (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol. Rev.* **33:** 657–687

Gordon JE & Christie PJ (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Grant SG, Jessee J, Bloom FR & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 4645–4649

Grohmann E, Christie PJ, Waksman G & Backert S (2018) Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol. Microbiol.* **107:** 455–471

Haase J, Lurz R, Grahn AM, Bamford DH & Lanka E (1995) Bacterial conjugation mediated by plasmid RP4: RSF1010 mobilization, donor-specific phage propagation, and pilus production require the same Tra2 core components of a proposed DNA transport complex. *J. Bacteriol.* **177:** 4779–4791

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Hooykaas PJJ & Beijersbergen AGM (1994) The virulence system of *Agrobacterium tumefaciens*. *Annu. Rev. Phytopathol.* **32:** 157–181

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Inoue H, Nojima H & Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:** 23–28

Jakubowski SJ, Kerr JE, Garza I, Krishnamoorthy V, Bayliss R, Waksman G & Christie PJ (2009) *Agrobacterium* VirB10 domain requirements for type IV secretion and T pilus biogenesis. *Mol. Microbiol.* **71:** 779–794

Jain A & Srivastava P (2013) Broad host range plasmids. *FEMS Microbiol. Lett.* **348:** 87–96

Kyte J & Doolittle RF (1982) A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157:** 105–132

Lang S, Gruber CJ, Raffl S, Reisner A & Zechner EL (2014) Common requirement for the relaxosome of plasmid R1 in multiple activities of the conjugative type IV secretion system. *J. Bacteriol.* **196:** 2108–2121

Lang S, Gruber K, Mihajlovic S, Arnold R, Gruber CJ, Steinlechner S, Jehl MA, Rattei T, Fröhlich KU & Zechner EL (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol. Microbiol.* **78:** 1539–1555

Lang S & Zechner EL (2012) General requirements for protein secretion by the F-like conjugation system R1. *Plasmid* **67:** 128–138

Lanka E & Fürste JP (1984) Function and properties of RP4 DNA primase. In *Proteins Involved in DNA Replication* pp 265–280. Boston, MA: Springer US

Mantis NJ & Winans SC (1992) The *Agrobacterium tumefaciens vir* gene transcriptional activator VirG is transcriptionally induced by acid pH and other stress stimuli. *J. Bacteriol.* **174:** 1189–1196

Melchers LS, Maroney MJ, den Dulk-Ras AD, Thompson DV, van Vuuren HAJ, Schilperoort RA & Hooykaas PJJ (1990) Octopine and nopaline strains of *Agrobacterium tumefaciens* differ in virulence; molecular characterization of the *virF* locus. *Plant Mol. Biol.* **14:** 249–259

Melchers LS, Regensburg-Tuïnk AJG, Bourret RB, Sedee NJ, Schilperoort RA & Hooykaas PJJ (1989) Membrane topology and functional analysis of the sensory protein VirA of *Agrobacterium tumefaciens*. *EMBO J.* **8:** 1919–1925

Meyer R (2015) Mapping type IV secretion signals on the primase encoded by the broad-host-range plasmid R1162 (RSF1010). *J. Bacteriol.* **197:** 3245–3254

Miele L, Strack B, Kruft V & Lanka E (1991) Gene organization and nucleotide sequence of the primase region of IncP plasmids RP4 and R751. *DNA Seq.* **2:** 145–162

Nagai H & Roy CR (2003) Show me the substrates: modulation of host cell function by type IV secretion systems. *Cell. Microbiol.* **5:** 373–383

Paget MSB, Chamberlin L, Atrih A, Foster SJ & Buttner MJ (1999) Evidence that the extracytoplasmic function sigma factor σ(E) is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J. Bacteriol.* **181:** 204–211

Pansegrau W & Lanka E (1996A) Mechanisms of initiation and termination reactions in conjugative DNA processing. *J. Biol. Chem.* **271:** 13068–13076

Pansegrau W & Lanka E (1996B) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncPα plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Pansegrau W, Schröder W & Lanka E (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90:** 2925–2929

Pansegrau W, Schröder W & Lanka E (1994) Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.* **269:** 2782–2789

Parker C & Meyer RJ (2007) The R1162 relaxase/primase contains two, type IV transport signals that require the small plasmid protein MobB. *Mol. Microbiol.* **66:** 252–261

Rees CED & Wilkins BM (1990) Protein transfer into the recipient cell during bacterial conjugation: studies with F and RP4. *Mol. Microbiol.* **4:** 1199–1205

Sambrook J & Russell DW (2001) Molecular cloning: a laboratory manual 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

Samuels AL, Lanka E & Davies JE (2000) Conjugative junctions in RP4-mediated mating of *Escherichia coli*. *J. Bacteriol.* **182:** 2709–2715

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

Schrammeijer B, den Dulk-Ras AD, Vergunst AC, Jácome EJ & Hooykaas PJJ (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. *Nucleic Acids Res.* **31:** 860–868

Schröder G & Lanka E (2005) The mating pair formation system of conjugative plasmids - a versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54:** 1–25

Simon R (1984) High frequency mobilization of Gram-negative bacterial replicons by the *in vitro* constructed Tn 5-Mob transposon. *MGG Mol. Gen. Genet.* **196:** 413–420

Strand TA, Lale R, Degnes KF, Lando M & Valla S (2014) A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One* **9:** 1–6

Thomas J & Hecht DW (2007) Interaction of *Bacteroides fragilis* pLV22a relaxase and transfer DNA with *Escherichia coli* RP4-TraG coupling protein. *Mol. Microbiol.* **66:** 948–960

Trieu-Cuot P, Carlier C, Martin P & Courvalin P (1987) Plasmid transfer by conjugation from *Escherichia coli* to Gram-positive bacteria. *FEMS Microbiol. Lett.* **48:** 289–294

Turk SCJ, Melchers LS, den Dulk-Ras H, Regensburg-Tuïnk AJG & Hooykaas PJJ (1991) Environmental conditions differentially affect *vir* gene induction in different *Agrobacterium* strains. Role of the VirA sensor protein. *Plant Mol. Biol.* **16:** 1051–1059

van Kregten M, Lindhout BI, Hooykaas PJJ & van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Mol. Plant-Microbe Interact.* **22:** 1356–1365

Vergunst AC, Schrammeijer B, den Dulk-Ras AD, De Vlaam CMT, Regensburg-Tuink AJG & Hooykaas PJJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science.* **290:** 979–982

Vergunst AC, van Lier MCM, den Dulk-Ras AD, Grosse Stuve TA, Ouwehand A & Hooykaas PJJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 832–837

Vogel AM & Das A (1992) Mutational analysis of *Agrobacterium tumefaciens* VirD2: tyrosine 29 is essential for endonuclease activity. *J. Bacteriol.* **174:** 303–308

Wilkins BM & Thomas AT (2000) DNA-independent transport of plasmid primase protein between bacteria by the I1 conjugation system. *Mol. Microbiol.* **38:** 650–657

Ziemienowicz A, Merkle T, Schoumacher F, Hohn B & Rossi L (2001) Import of *Agrobacterium* T-DNA into plant nuclei: two distinct functions of VirD2 and VirE2 proteins. *Plant Cell* **13:** 369

Supplemental information

Supplemental Table S1. Tables showing the average translocation frequencies of CRAfT assays within figures.

Figure 1 protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(4.6 \pm 0.8) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-531	$(8.4 \pm 1.0) \times 10^{-8}$
RP4 / Cre::$TraI_{RP4}$625-732	$(9.5 \pm 1.7) \times 10^{-5}$
RP4 / Cre::$TraI_{RP4}$599-697	$(6.4 \pm 0.3) \times 10^{-9}$
RP4 / Cre::$TraI_{RP4}$532-732	$(8.3 \pm 1.9) \times 10^{-9}$

Figure 3A protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(8.5 \pm 1.2) \times 10^{-3}$
RP4Δ*traI* / Cre::$TraI_{RP4}$1-732	$(1.9 \pm 0.5) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$243-732	$(3.3 \pm 0.8) \times 10^{-3}$
RP4Δ*traI* / Cre	0

Figure 3B protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Constructs used	Average translocation efficiency (n = 3)
RP4 / Cre::$TraI_{RP4}$1-732	$(9.6 \pm 4.8) \times 10^{-5}$
RP4Δ*traI* / Cre::$TraI_{RP4}$1-732	$(2.2 \pm 0.7) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$243-732	$(7.8 \pm 1.2) \times 10^{-4}$
RP4Δ*traI* / Cre::$TraI_{RP4}$625-732	$(4.5 \pm 1.2) \times 10^{-4}$
RP4Δ*traI* / Cre	0

Figure 4 protein translocation frequencies from the *A. tumefaciens* LBA288 and LBA1100 donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Strains and constructs used	Average translocation efficiency (n = 3)
LBA288 / RP4Δ*traI* / pBFFCre::$TraI_{RP4}$625-732	$(1.6 \pm 0.8) \times 10^{-2}$
LBA288 / RP4Δ*traI* / pBFFCre	0
LBA288 / RP4Δ*traI* / pBFFCre::$VirF_{Ti}$166-202	$(1.8 \pm 0.1) \times 10^{-8}$
LBA1100 / RP4Δ*traI* / pBFFCre::$TraI_{RP4}$625-732	$(5.4 \pm 1.9) \times 10^{-3}$
LBA1100 / RP4Δ*traI* / pBFFCRE::$VirF_{Ti}$166-202	$(3.4 \pm 2.1) \times 10^{-3}$

Table continued on the next page.

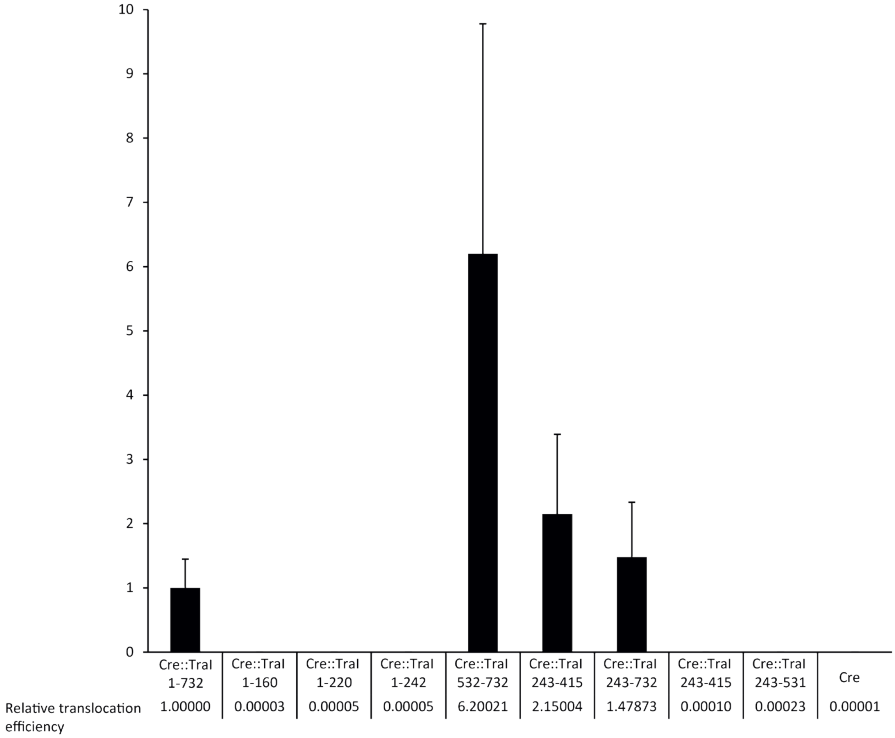
Figure S2 protein translocation frequencies from the *E. coli* DH5α donor to recipient *E. coli* CSH26Cm:LTL of the co-cultivation experiments. Error is SEM.

Experiment #1

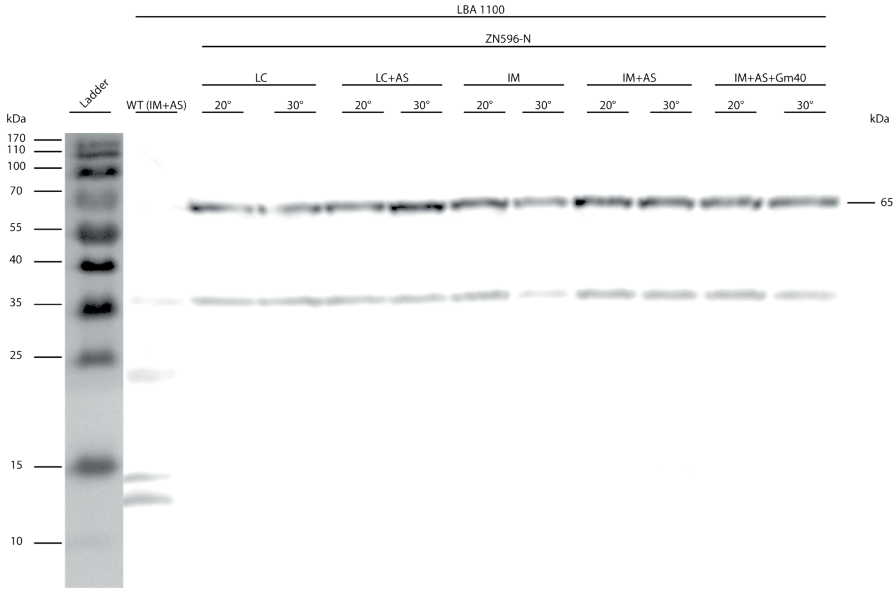
Constructs used	**Average translocation efficiency (n = 3)**
RP4 / Cre::$TraI_{RP4}$1-732	$(5.3 \pm 1.3) \times 10^{-3}$
RP4 / Cre::$TraI_{RP4}$1-160	$(1.4 \pm 0.2) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-220	$(2.9 \pm 0.4) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$1-242	$(2.5 \pm 0.3) \times 10^{-7}$
RP4 / Cre::$TraI_{RP4}$532-732	$(3.3 \pm 1.0) \times 10^{-2}$
RP4 / Cre	$(5.9 \pm 2.5) \times 10^{-8}$

Experiment #2

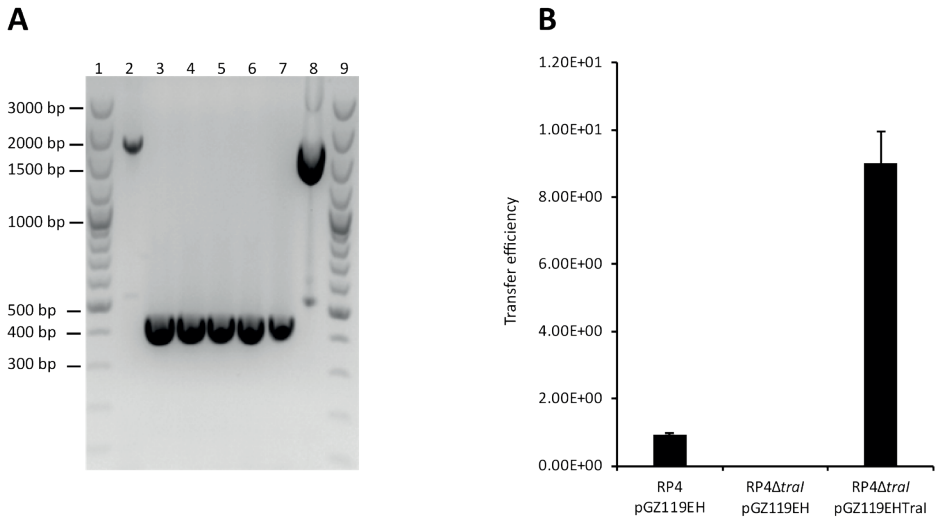
Constructs used	**Average translocation efficiency (n = 3)**
RP4 / Cre::TraIRP41-732	$(3.1 \pm 3.1) \times 10^{-3}$ (n=2)
RP4 / Cre::TraIRP4243-415	$(1.0 \pm 0.1) \times 10^{-2}$
RP4 / Cre::TraIRP4243-415	$(6.9 \pm 1.8) \times 10^{-3}$
RP4 / Cre::TraIRP4243-531	$(4.7 \pm 2.4) \times 10^{-7}$
RP4 / Cre	$(1.6 \pm 0.1) \times 10^{-8}$



Supplemental Figure S2. Preliminary characterization of the TS of the relaxase $TraI_{RP4}$ using a CRAfT assay. This CRAfT assay showed the normalized relative averaged translocation efficiency of a Cre fusion proteins compared to the full-length $TraI_{RP4}$ translocation efficiency from *E. coli* DH5α to *E. coli* CSH26Cm:LTL. The constructs used were as follows: Cre::TraI (full-length $TraI_{RP4}$ sequence 1-732 AA); Cre::TraI 1-160 ($TraI_{RP4}$ truncation from 1-160 AA); Cre::TraI 1-220 ($TraI_{RP4}$ truncation from 1-220 AA), Cre::TraI 1-242 ($TraI_{RP4}$ truncation from 1-242 AA); Cre::TraI 532-732 ($TraI_{RP4}$ truncation from 532-732 AA); Cre::TraI 243-415 ($TraI_{RP4}$ truncation from 243-415 AA); Cre::TraI 243-732 ($TraI_{RP4}$ truncation from 243-732 AA); Cre::TraI 243-415 ($TraI_{RP4}$ truncation from 243-415 AA); Cre::TraI 243-531 ($TraI_{RP4}$ truncation from 243-531 AA) and Cre (full-length Cre protein). In this experiment, there was an issue detected with constructs Cre::$TraI_{RP4}$ 1-160 AA, $TraI_{RP4}$ 1-220 AA, $TraI_{RP4}$ 1-242 AA, $TraI_{RP4}$ 243-415 AA, and $TraI_{RP4}$ 243-531 AA, as there was no stop codon present to terminate translation at the desired position within $TraI_{RP4}$ domains; translational stop occurred after incorporation of an additional and arbitrary C-terminal 60 AA. All other construct used had the intended stop codon present at the end. This series of CRAfT assays clearly demonstrated that the TS was most likely located in the last C-terminal 200 AA. Two independent experiments were normalized to the translocation efficiency of full-length $TraI_{RP4}$ fused to the Cre protein that was averaged and set to 1 for the basal translocation rate. All other translocation efficiencies were averaged and divided by the Cre::$TraI_{RP4}$ basal rate of its independent experiment. The y-axis depicts the normalized relative translocation frequency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The Error bars depict SEM (n = 3; except for Cre::TraI with n = 5; and Cre with n = 6). The transformation efficiencies of the experiments can be found in **figure S1**.



Supplemental Figure S3. Protein expression of *A. tumefaciens* expression plasmid pBFF under the control of $virF_{Ti}$ promoter in different conditions. This western blot shows the ZN596-N protein expression in different temperature and media conditions. The only Cre construct with a FLAG tag available for testing at the time point was the ZN596-N protein. ZN696-N is a Cre recombinase, fused with a 6 zinc-finger DNA binding motif and a $VirF_{Ti}$ TS as well as a FLAG tag, encoded on a pBFF plasmid (van Kregten *et al.*, 2009) which is present in *A. tumefaciens* LBA1100 and cultured in induction medium (IM; **Chapter 2**) or Luria broth (LB), either with or without 200 μM acetosyringone (AS) and 40 μg/mL gentamycin (Gm40). For detection, a monoclonal FLAG tag antibody was used (Sigma). The size of the ZN596-N protein is around 65 kDa. The bands detected around 36 kDa could be due to degradation of the recombinant protein and/or due to unspecific binding of the antibody. *A. tumefaciens* LBA1100 not containing the construct, in lane 2, shows minimal unspecific binding around those sizes as well.



Supplemental Figure S4. **A**: Gel electrophoresis of PCR products indicative of RP4Δ*traI* deletion plasmids. All plasmids were isolated from *E. coli* DH5α. Primers TraICheckFW and TraICheckRV were used. The expected fragment size of $traI_{RP4}$ with those primers was to be expected around 1700 bp and with a deletion of $traI_{RP4}$ the fragment size was to be expected around 400 bp. Lanes 1 and 9: contain the 100 bp plus DNA ladder (Thermo Scientific); Lane 2: colony preparation of wild-type RP4; Lane 3: RP4Δ*traI* clone 1; Lane 4: RP4Δ*traI* clone 2; Lane 5: RP4Δ*traI* clone 3; Lane 6: RP4Δ*traI* clone 4; Lane 7: RP4Δ*traI* clone 5; Lane 8: midi-prep of wild-type RP4. **B**: Conjugation experiment to determine the transfer of the RP4Δ*traI* knock-out plasmids with and without the complementation of the relaxase $TraI_{RP4}$. As donor strain *E. coli* DH5α and as recipient strain *E. coli* MS614 was used. The wild-type RP4 was used as a positive control to show self-transfer in the presence of the plasmid pGZ119EH. RP4Δ*traI* was not able to transfer itself in the presence of the plasmid pGZ119EH. RP4Δ*traI* was able to transfer itself when the complementing $traI_{RP4}$ was present on plasmid pGZ119EHTraI. Conjugation time was 2 h. Transfer efficiency (on y-axis) was expressed as the fraction of transconjugant cells with kanamycin (50 μg/mL) and streptomycin (25 μg/mL) resistance over the number of donor cells with kanamycin (50 μg/mL) and chloramphenicol (10 μg/mL) resistance. In transconjugant cells, the kanamycin resistance marker would be present on the transferred RP4 plasmid, and the streptomycin resistance marker is present in the genome of the recipient strain *E. coli* MS614. The donor cells contain a kanamycin marker on the RP4 plasmid and the chloramphenicol marker on the plasmid pGZ119EH or pGZ119EHTraI. The error bar depicts the SEM (n = 3).

Chapter 5

Exploration into the modularity of DNA processing proteins of conjugative systems RP4 and Ti of *Agrobacterium tumefaciens*

Felix T. Wittleben, Silvia Lang, Paul J.J. Hooykaas and Bert J. van der Zaal

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Abstract

Conjugative DNA transfer systems are able to transfer genetic material from their host organisms into recipient cells. The conjugative plasmid RP4 can transfer itself between a variety of bacterial species within short timeframes. *Agrobacterium tumefaciens,* via its Ti plasmid, is able to transfer a defined sequence of transfer-DNA (T-DNA) to plant cells. Both RP4 and the Ti plasmid utilize a type-IV secretion system (T4SS) for the DNA transfer sharing functional similarities. This study focused on a key element of the transfer machinery, the DNA transfer and replication (Dtr) system. The Dtr system employs a relaxase and its auxiliary proteins for the nicking of the DNA substrate to be transferred. Transfer is facilitated by a translocation peptide in the C-terminal part of the relaxase. For this study, the relaxase domain of $VirD2_{Ti}$ was fused with the translocation signal (TS) of the RP4 relaxase $TraI_{RP4}$, intending to find out whether transfer of $VirD2_{Ti}$ substrates would now be transferred via the $T4SS_{RP4}$. In order to aid the hybrid relaxase Dtr_{Ti}, auxiliary proteins ($VirC1_{Ti}$ $VirC2_{Ti}$ and $VirD1_{Ti}$) were supplied by expression from a newly constructed operon. The recognition site for the hybrid relaxase (border) was positioned on the substrate DNA to be transferred. Some of the individual functions of the elements created for this study were observed. The hybrid relaxase itself was shown to be transferred via the $T4SS_{RP4}$, and the operon for auxiliary protein expression was functional. However, all elements combined and inserted into a bacterial system did not lead to transfer of the substrate DNA, neither from *A. tumefaciens* to *Saccharomyces cerevisiae* nor from *Escherichia coli* to *E. coli*. A hybrid-relaxase combining the relaxase function of $TraI_{RP4}$ with the translocation signal of $VirD2_{Ti}$ to be transferred by the $T4SS_{Ti}$ was also created. Nevertheless, also in this case no DNA transfer could be detected between *A. tumefaciens* and *S. cerevisiae*. This study explored the possibility of combining the Dtr functions of different T4SS conjugative systems but also highlighted the complexity of systems and their interplay of their intricate processes.

Introduction

Bacterial conjugative systems transfer DNA between cells, and can also function in specialized environmental as well as in species-specific niches. In general, conjugational systems consist out of three elements. The actual transfer structure, the mating pair formation (Mpf), consists of a transmembrane spanning multi-protein complex and a pilus (Alvarez-Martinez & Christie 2009). This physical structure, bridging between the donor and the recipient cell allows DNA and proteins (substrates) to be transferred to the recipient cell. It is nowadays referred to as the type-IV secretion system (T4SS). The entry of the substrate to the T4SS is regulated by the coupling protein (CP). The so-called transfer and replication (Dtr) genes encode a relaxase protein as well as auxiliary proteins supporting the DNA processing necessary for transfer. The DNA to be transferred contains a specific *cis*-acting region that is recognized by the Dtr proteins, the origin of transfer (*oriT*) or sometimes also called basis of mobility (*bom*). The relaxase is presented to this site with the aid of auxiliary proteins, and is able to nick a single-strand of the *oriT* region called *nic* site, during which it becomes covalently attached at the 5'-end. The nicked single-strand is then displaced, and the single-stranded DNA-relaxase complex is presented to the CP. The CP in turn allows entry into the T4SS only if the relaxase has the required transfer signal (TS; van Kregten *et al.*, 2009). Once passed on through the T4SS into the recipient cell, by largely unknown mechanisms the single-stranded DNA is converted into a double-stranded form again, circularized and maintained as a plasmid or integrated into the recipient genome depending on the type of conjugative system and the recipient cell involved (Lanka & Wilkins, 1995; Pansegrau & Lanka, 1996A; Parker & Meyer, 2005; Fernández-López *et al.*, 2014).

The transfer system of conjugative plasmids such as RP4 can not only mediate transfer via the CP_{RP4}/$T4SS_{RP4}$ of the RP4 plasmid itself, but also of other DNA molecules present in the same cell, if these carry the corresponding single *cis*-acting $oriT_{RP4}$ sequence that can be recognized by the Dtr_{RP4} system (Guiney & Yakobson, 1983). This leads to efficient transfer of such mobilizable plasmids, or also to chromosomal mobilization, if an $oriT_{RP4}$ is present in the chromosome (Yakobson & Guiney, 1984).

The plant-pathogenic bacterium *Agrobacterium tumefaciens* carries a transfer system comprised of virulence genes (vir_{Ti}) carried on its Ti (tumor-inducing) plasmid, which can transfer part of the Ti plasmid, the T-DNA, into plant cells. The T-DNA produced from the Ti plasmid is derived from the T-region of the plasmid. The T-DNA is surrounded by 24 bp border repeats, which are recognized by the relaxase $VirD2_{Ti}$ of the Dtr_{Ti} encoded by the virulence system. These two *cis*-acting border sequences are called right border (RB_{Ti}) and left border (LB_{Ti}). Transfer of a plasmid containing only one RB_{Ti} sequence is also possible (**Chapter 3**; Jen & Chilton, 1986; Rolloos *et al.*, 2014). A single-stranded T-DNA (T-strand) is then displaced from the plasmid by the Dtr_{Ti} system (Stachel, Timmerman & Zambryski 1987; Howard *et al.*, 1989). Ensuing the processing, transfer is mediated by the CP_{Ti}/$T4SS_{Ti}$ encoded by the $virB_{Ti}$ operon of the virulence region (Hooykaas *et al.*, 1980; Christie & Gordon, 2014). After transfer to plants the T-strand can integrate into the recipient cells genome, a process called T-DNA integration. Following the T-DNA integration and expression within the plant genome, the plant cells physiological function is transformed.

The transfer systems of RP4 and the Ti plasmid share a close evolutionary development, as is evident from the genetic homology of their components, architecture and function (Farrand, Hwang & Cook, 1996; Pansegrau & Lanka, 1996A; Christie & Gordon, 2014), but also possess unique features on their own. In particular, although having a similar net function and sharing the same overall makeup, key details such as substrate transfer requirements (e.g. *oriT* or border sequences) for both Dtr systems and host range of both T4SSs are seemingly different. The host ranges vary for example, with the Ti plasmid transfer to plants and with the RP4 plasmid transfer to bacteria. If these differences would be directly

attributable to specific characteristics of homologous modules of the Dtr systems of both types of plasmids, it could be hypothesized that certain parts of the Dtr systems could keep their specific traits while functioning within another Dtr, CP and T4SS context. In other words, different modules should be exchangeable, but convey different specifics to a given Dtr system.

The highly promiscuous conjugational plasmid RP4 resides in a variety of Gram-negative host species (Pansegrau *et al.*, 1994). RP4 is able to efficiently transfer itself and also mobilizable plasmids, into other prokaryotic recipient cells. RP4 also exhibits *trans*-kingdom conjugation as in certain cases (laboratory conditions) plasmids can be transferred from *E. coli* to yeast cells (Nishikawa, Suzuki & Yoshida, 1990; Suzuki, Moriguchi & Yamamoto, 2015). The initial displaced single-strand (Pansegrau & Lanka, 1996A), once inside the recipient cell, can be circularized and further processed to form double-stranded plasmid that can be stably maintained, if possessing an origin of vegetative replication, or *oriV* (synonymously with origin of replication *oriRep* or *ori*) that is active within the new host cell. The Dtr_{RP4}-, CP_{RP4}- and $T4SS_{RP4}$-mediated conjugation requires physical association of the donor and recipient cells, but no induction of the system itself.

The Ti plasmid is endogenous to the phytopathogen *A. tumefaciens*. The natural recipients are plants cells, to which transfer occurs efficiently. Transfer to other eukaryotes, like yeast and fungi, is possible but the transfer is less efficient (Bundock *et al.*, 1995). T-DNA transfer and integration, also called *Agrobacterium*-mediated transformation (AMT), is one of the few *trans*-kingdom conjugational processes known (Suzuki, Moriguchi & Yamamoto, 2015). Induction of the AMT process is required and is highly dependent on factors such as pH level, sugar concentrations, inducing factors and temperature (Gao & Lynn, 2005; Lin, Binns & Lynn, 2008). It is to note that the Ti plasmid encodes a second conjugative transfer system, the tra_{Ti}/trb_{Ti} system that mediates intra-bacterial self-transfer after induction by specific opines (Farrand, Hwang & Cook, 1996; Hamilton *et al.*, 2000). This self-transfer system is not considered for this research. Therefore, in this study from here onwards, the Dtr_{Ti}/CP_{Ti}/$T4SS_{Ti}$ vir_{Ti} system will be referred to as Dtr_{Ti}/CP_{Ti}/$T4SS_{Ti}$.

Both transfer systems of Ti and RP4 have been widely studied and almost all individual elements are known. However, the differences in transfer efficiencies to recipient cells are not fully understood. It is not well known what factors influence the change in transformation efficiencies to non-natural recipient cells, or why certain recipients are recalcitrant for conjugation. Apart from differences in the transfer system components, it is postulated for this research, that cues for variation in transformation efficiencies could be attributable to differences within the Dtr systems. In addition to that, if the conserved Dtr components of both Ti and RP4 systems do indeed contain all essential factors for DNA transfer, and if their functional characteristics would be understood, dedicated exchange of key elements between the two related systems should be possible. Previous studies already indicated the possibility of exchange of different transfer system factors. For example the CP of different systems could be functionally switched (Cabezón, Lanka & Cruz, 1994; Hamilton *et al.*, 2000), and specific TS sequences of Dtr proteins could be used to transfer different non-natural substrates via the CP/T4SS (Vergunst *et al.*, 2000; Schrammeijer *et al.*, 2003; van Kregten *et al.*, 2009). This work showed that TSs of T4SSs can be translationally fused to proteins to allow transfer via the CP/T4SS. The $VirF_{Ti}$ TS when fused to Cre recombinase can mediate transport via the CP_{Ti}/$T4SS_{Ti}$ into plant cells (Vergunst *et al.*, 2000). The fusion of the full-length proteins of $VirE2_{Ti}$, $VirE3_{Ti}$ and $VirF_{Ti}$, all containing a $T4SS_{Ti}$ TS to Cre, also showed transfer via the CP_{Ti}/$T4SS_{Ti}$ to *S. cerevisiae* (Schrammeijer *et al.*, 2003; Vergunst *et al.*, 2005). Furthermore, a variety of other proteins, such as SceI endonuclease, zinc-fingers (van Kregten *et al.*, 2009), TALENs (This study; data not shown) or Cas9 enzymes (Schmitz *et al.*, 2020) can be translocated by adding a $VirF_{Ti}$ TS. In like manner, the relaxase function was also shown to be modular. The minimal $VirD2_{Ti}$ relaxase domain (1-204 AA) fused to the $VirF_{Ti}$ translocation signal was able

to mediate transfer of T-DNA to plant cells (van Kregten *et al.*, 2009).

The Dtr_{RP4} system is made up of one key DNA-nicking enzyme and its auxiliary proteins. The $TraI_{RP4}$ relaxase protein is made up of 733 AA. See **figure 1A** for a graphical representation. The relaxase domain is situated in the N-terminus (1-210 AA) of the protein. The TS of $TraI_{RP4}$ is localized in the C-terminus (**Chapter 4**) and is likely to be recognized by the CP_{RP4} $TraG_{RP4}$ for transfer via the $T4SS_{RP4}$. The relaxase $TraI_{RP4}$ was shown to nick the single-strand DNA on the nic_{RP4} site within the $oriT_{RP4}$, aided by the scaffold protein $TraH_{RP4}$, the recruiting protein $TraJ_{RP4}$ and the helicase protein $TraK_{RP4}$ (Ziegelin, Fürste & Lanka, 1989; Ziegelin *et al.*, 1992; Pansegrau, Schröder & Lanka, 1993; Pansegrau, Schröder & Lanka, 1994; Balzer, Pansegrau & Lanka, 1994). During the nicking reaction $TraI_{RP4}$ becomes covalently bound at the 5´-end of the single-stranded DNA. This binding is supposed to last during DNA transfer and its subsequent processing in the recipient cell (Pansegrau & Lanka, 1996A).

The Dtr_{Ti} system is comprised of the relaxase $VirD2_{Ti}$ and its auxiliary proteins. The $VirD2_{Ti}$ protein is made up of 424 AA. $VirD2_{Ti}$ contains the relaxase domain N-terminally (1-204 AA) with a C-terminal TS that is likely to be recognized by the CP_{Ti} $VirD4_{Ti}$ for transfer via the $T4SS_{Ti}$ (van Kregten *et al.*, 2009). See **figure 1A** for a graphical representation. It nicks the T-DNA at the RB_{Ti} and LB_{Ti} sequences and remains covalently bound to the RB_{Ti} at the 5´-end and moves with the T-strand to the recipient cell (Tinland, Hohn & Puchta, 1994; Zupan & Zambryski, 1995). $VirD2_{Ti}$ nicks a single-strand on the nic_{Ti} site present within the border sequences (Yadaf *et al.*, 1982; Wang *et al.*, 1987). The function of $VirD2_{Ti}$ is aided by three auxiliary proteins. Although $VirD2_{Ti}$ can nick single stranded DNA *in vitro*, $VirD1_{Ti}$ is essential for nicking of double-stranded DNA molecules in the Ti plasmid (Scheiffele, Pansegrau & Lanka, 1995). $VirC1_{Ti}$ binds to the overdrive OD_{Ti} sequence adjacent to the right border sequence and associates with $VirC2_{Ti}$ (Toro *et al.*, 1988; Relić *et al.*, 1998;). The $VirC2_{Ti}$ protein has a ribbon-helix-helix (RHH) DNA-binding fold by which it binds to double-stranded DNA molecules (Lu *et al.*, 2009). The $VirC1_{Ti}$/$VirC2_{Ti}$ forms a complex with the $VirD1_{Ti}$/$VirD2_{Ti}$/T-strand and further associates this network to the CP_{Ti} (Atmakuri *et al.*, 2007).

Despite the RB_{Ti} and LB_{Ti} nic_{Ti} sites bearing similarity to the nic_{RP4} site of the $oriT_{RP4}$ (**Figure 1C**), neither Dtr_{Ti} nor Dtr_{RP4} should be able to utilize access each other $oriT_{RP4}$ or border sequences *in vivo*, as it was shown with the Dtr_{RP4} not able to utilize the RB_{Ti} (**Chapter 3**). Yet *in vitro* $VirD2_{Ti}$ is able to cleave a single-strand containing the nic_{RP4} (Pansegrau, Schröder & Lanka, 1993; Pansegrau *et al*, 1993). The speed of conjugation differs as well between both transfer systems. RP4 is able to mediate plasmid transfer between *E. coli* within minutes and from *A. tumefaciens* to yeast after 30 min to hours (This study, data not shown; Babic *et al.*, 2008). The Ti plasmid on the other hand was shown to require several days for successful T-DNA transfer and AMT to yeast (Bundock *et al.* 1995) and plant cells (**Chapter 2**; Sakalis, 2013).

This study explores the possibility to change characteristics of the related Dtr transfer systems of RP4 and Ti. It was investigated whether it would be possible to exchange the DNA recognition parts of the relaxase proteins of both systems by making translational fusions. When such hybrid relaxase proteins could recognize either nic_{Ti} on a RB_{Ti} or nic_{RP4} on $oriT_{RP4}$ while being transferred via the other CP/T4SS, it could provide indications about elements specifying host range, transfer and modularity. Different hybrid relaxase proteins were created together with specific plasmids able to be recognized as substrate for those hybrid proteins. It was shown that individual parts of the modified Dtr systems worked *in vitro* and *in vivo*. But no transfer was detected via the hybrid relaxases mediated by CP/T4SS transfer from a donor to a recipient cell. However a platform to test the hybrid relaxases and the modified Dtr systems was created, yet further research is required to test if transfer actually can occur.

Results

DNA transfer mediated by $T4SS_{RP4}$ from *Agrobacterium tumefaciens* to plants was not detected

The promiscuous conjugative plasmid RP4 can transfer DNA into different bacterial species (**Chapter 3**) as well as into yeast cells (**Chapter 4**), but has not been shown yet to be able to transfer plasmids to plant cells. For studying Dtr_{RP4}-, CP_{RP4}-, $T4SS_{RP4}$-mediated DNA transfer to plants, an appropriate donor strain of *A. tumefaciens* was chosen. The strain LBA2210 contained the RP4 plasmid but lacked a Ti plasmid. In addition, a plasmid was constructed to function as substrate for the CP_{RP4}/$T4SS_{RP4}$ well as being detectable in plant cells. This mobilizable vector, was based on the plant transformation plasmid pCAMBIA2201 (Cambia legacy plasmid based on Hajdukiewicz, Svab & Maliga, 1994). An $oriT_{RP4}$ that can be processed by the Dtr_{RP4} system was introduced into the backbone of pCAMBIA2201 (**Figure 2**). Furthermore, pCAMBIA2201 already contained also RB_{Ti}/LB_{Ti} sequences, normally used for transfer to plants in conjunction with the Ti plasmid, not recognized by the Dtr_{RP4} (**Chapter 3**). For detection in plants, the plasmid pCAMBIA2201 contains an expression cassette for the *gus* (*gusA*) gene. GUS protein activity is easily detectable via a histochemical detection assay (Jefferson, Kavanagh & Becan, 1987). A derivative of *A. tumefaciens* strain LBA2210, harboring besides RP4 the pCAMBIA2201$oriT_{RP4}$ plasmid, was injected into *Nicotiana benthamiana* leaves (**Figure 2**). This injection method is also called agroinfiltration (Sakalis, 2013). After the agroinfiltration, no GUS protein expression could be measured in the collected leaf tissue. This lack of measurement indicated no plasmid transfer could be detected from *A. tumefaciens* to *N. benthamiana*. As functional control, to check for the possibility of DNA transfer to plants during the experimental conditions, *A. tumefaciens* LBA1100 containing a Ti plasmid and the control plasmid p14 carrying a *gus* reporter gene sequence flanked by a RB_{Ti} and LB_{Ti} sequence was used (**Chapter 3**). GUS protein activity could be measured in *N. benthamiana* leaves after co-cultivation with the control strain. In the next paragraphs, specific parts of the Dtr system of plasmids RP4 and Ti to investigate the requirements for DNA and protein transfer via these overall analogous yet heterologous transfer system particular the CP/T4SS are combined.

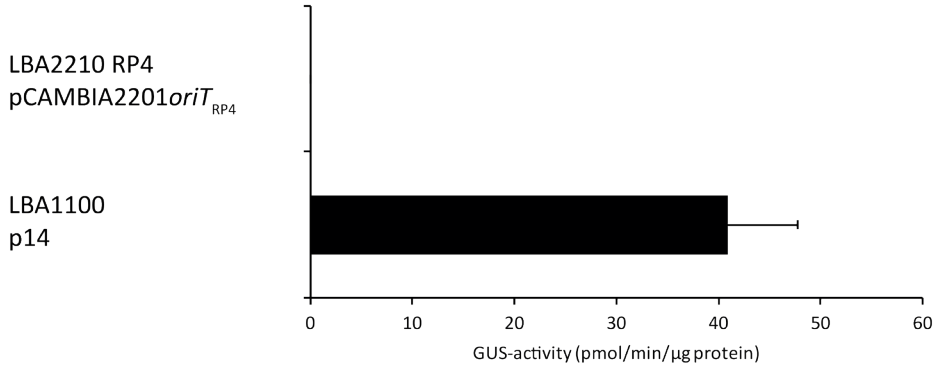


Figure 2. DNA transfer from *A. tumefaciens* to *N. benthamiana* mediated by conjugative plasmid RP4. Agroinfiltration of *A. tumefaciens* strain LBA2210 containing pCAMBIA2201*oriT*RP4 and LBA1100 containing the control plasmid p14 into *N. benthamiana* leaves. Co-cultivation time was 72 h. GUS expression was determined using a MUG assay. Error bar is SEM (pCAMBIA2201*oriT*RP4: n = 8; p14p: n = 3). Plasmids used were p14P (p14) and pCAMBIA2201*oriT*RP4 (pCAMBIA2201$oriT_{RP4}$). The GUS-activity for pCAMBIA2201$oriT_{RP4}$ was 0 and for p14 it was 40 ± 6.8.

Construction of hybrid relaxases

First the combination of elements of the Dtr systems of plasmids RP4 and Ti were explored to create a new functional hybrid Dtr system. A hybrid relaxase was created comprising of an N-terminal $VirD2_{Ti}$ relaxase domain fused to the $TraI_{RP4}$ C-terminal domain. The C-terminal domain of $TraI_{RP4}$ lacks the $TraI_{RP4}$ relaxase function, but contains the translocation signal TS (**Chapter 4**) recognized by the CP_{RP4} $TraG_{RP4}$ which controls transfer via the $T4SS_{RP4}$. Two versions of the $VirD2_{Ti}$ relaxase domain were employed. One version contained the $VirD2_{Ti}$ 1-204 AA sequence. The other version contained the $VirD2_{Ti}$ sequence 1-274 AA containing besides the relaxase domain also the DUF-60 domain (van Kregten *et al.*, 2009). The role of the DUF-60 domain in the DNA transfer process is not yet fully understood (van Kregten *et al.*, 2009). For the $TraI_{RP4}$ C-terminal domain the sequence from 243-732 AA was chosen, just lacking the proposed N-terminal relaxase domain (Pansegrau, Schröder & Lanka, 1993; Cole & Guiney, 1995). This created the hybrid relaxases $VirD2_{Ti}$ 1-204 AA::$TraI_{RP4}$ 243-732 AA and $VirD2_{Ti}$ 1-274 AA::$TraI_{RP4}$ 243-732 AA or in short HD204I and HD274I respectively (**Figure 1B**).

Hybrid relaxases $VirD2_{Ti}$ 1-204/274 AA::$TraI_{RP4}$ 243-732 AA are transferred via the CP_{RP4}/$T4SS_{RP4}$

The ability of the $TraI_{RP4}$ TS present in the hybrid-relaxase HD204/274I to mediate transfer via the CP_{RP4}/$T4SS_{RP4}$ was tested. For this a Cre Reporter Assay for Translocation (CRAfT) was used to check transfer of the hybrid-relaxase HD204I and HD274I fused with a Cre recombinase reporter protein via the CP_{RP4}/$T4SS_{RP4}$ (Vergunst *et al.*, 2000; Vergunst *et al.*, 2005). As protein donor strain *Escherichia coli* MS411, and as recipient strain *E. coli* MS614 was used (Lang *et al.*, 2010). The Cre protein was fused N-terminally to the hybrid-relaxase to arrange functionality of the recombinase domain (van Duyne, 2015), and to ensure a C-terminal location of the TS. Cre by itself cannot actively be translocated via the $T4SS_{RP4}$. The recipient strain contained a plasmid encoded chloramphenicol resistance cassette (*cat*) that was disrupted by a tetracycline resistance cassette (*tet*) flanked by *loxP* sites. In case Cre recombinase proteins are transported via the CP_{RP4}/$T4SS_{RP4}$ into the recipient strain, the *tet* cassette will be excised via recombination and thence be lost, hereby rendering the *cat* cassette functional. Using the same expression plasmid in the same donor strain *E. coli* MS411, the expression of the hybrid-relaxase fusion proteins during experimental conditions, was verified by western blotting using Cre antibodies (**Figure S4**). The results obtained in the CRAfT experiment testing the translocation of the Cre fusion proteins are given in **figure 3** and **table 1**. The CRAfT assay showed that the full-length $TraI_{RP4}$ 1-732 AA, as well as the C-terminal part of $TraI_{RP4}$ 243-732 AA, fused to the Cre protein were transferred via the CP_{RP4}/$T4SS_{RP4}$ into the recipient cell. Hybrids in which the relaxase part of $VirD2_{Ti}$, (either the short 1-204 AA domain or the 1-274 AA domain including DUF-60) was fused to $TraI_{RP4}$ 243-732 AA, were also translocated into recipient cells by the CP_{RP4}/$T4SS_{RP4}$ (**Figure 3**). The negative controls for the experiment were as follows: Cre without any TS and Cre fused with $VirD2_{Ti}$ were not translocated by CP_{RP4}/$T4SS_{RP4}$. Also a donor strain harboring only the RP4 plasmid and lacking the Cre expression plasmid was used to evaluate the background level of false positive colonies. With the negative controls, only very few of *cat* positive colonies were observed. They could be considered background level due to occasional spontaneous recombination at the homologous *loxP* sites with a frequency ranging from a 10^{-6} to 10^{-8} (**Table 1**). As each donor strain contained RP4, the self-transfer into the recipient cells was determined for each set-up. Self-transfer of RP4 should be proof of the production of a functional Dtr_{RP4}/CP_{RP4}/$T4SS_{RP4}$ system in the donor strain. The RP4 transfer was shown to occur in all strains used in this CRAfT assay (**Table 1**).

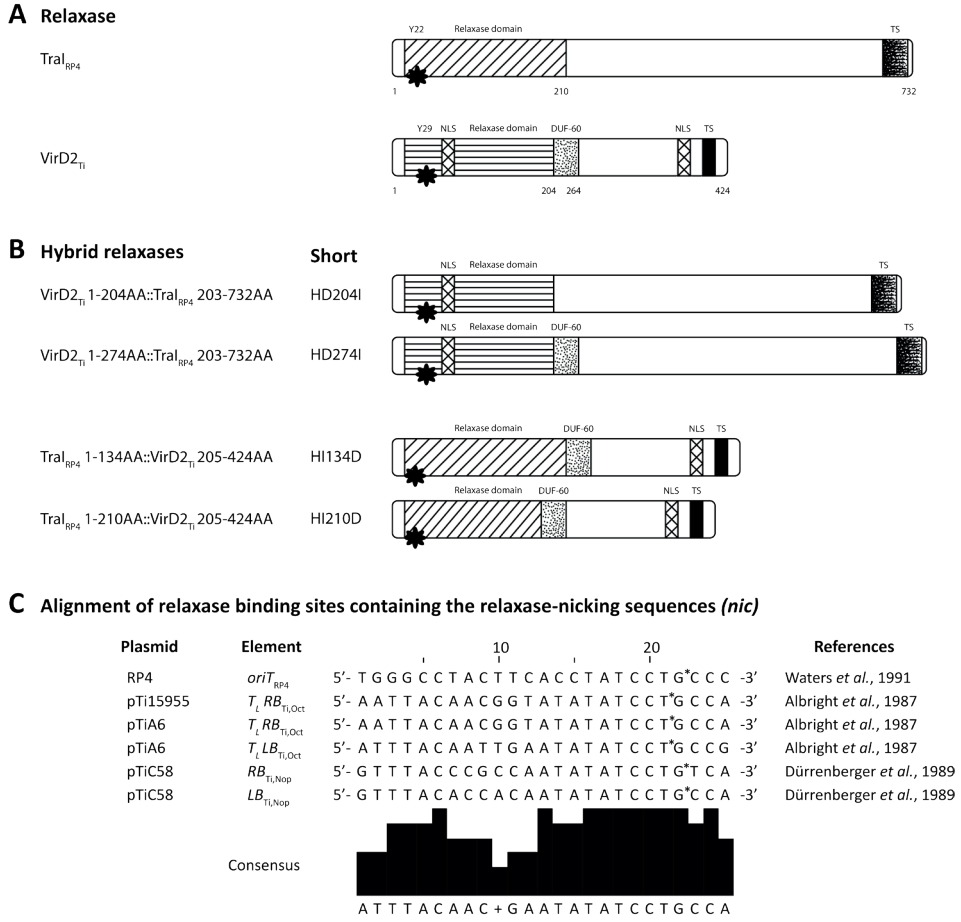


Figure 1. A: Overview of the proposed functional domains of the relaxase proteins of the Dtr systems of RP4 and Ti. The schematic representations do not represent the actual size of proteins or functional domains. The relaxase $TraI_{RP4}$ contains an N-terminal relaxase domain (hatched box) and a C-terminal translocation domain (shaded box). Auxiliary protein binding domains (e.g. for $TraH_{RP4}$) are not characterized yet. The active tyrosine at AA position 22 that is essential for the transesterification process at the $oriT_{RP4}$ is marked with a black asterisk (Pansegrau, Schröder & Lanka, 1993; Cole & Guiney, 1995). The relaxase $VirD2_{Ti}$ also contains an N-terminal relaxase domain (horizontal lined box) and a C-terminal translocation domain (black box) contained in the last 40AA (Vergunst *et al.*, 2005). $VirD2_{Ti}$ contains two nuclear localization sequence (NLS; cross hatched box): one located in the N-terminus, embedded in the relaxase domain, and one in the C-terminus close to the translocation signal (TS; Rossi, Hohn & Tinland, 1993). The active tyrosine at AA position 29, essential for the transesterification process at the border sequence of the T-DNA, is marked with a black asterisk (Vogel & Das, 1992; Lessl & Lanka, 1995). The DUF-60 domain is depicted with a dotted box (van Kregten *et al.*, 2009). **B:** Overview of the hybrid relaxase versions created for this study. The schematic representations do not represent the actual size of proteins or functional domains. **C**: Alignment of the relaxase binding sites containing the nick sequences (*nic*) between $oriT_{RP4}$ of the RP4 plasmid and border sequences of Ti plasmid T-regions. The relaxase *nic* sites are marked with an asterisk between the bases of the DNA single-strand. Alignment between sequences was performed using the Jalview (Version 2) software (Waterhouse *et al.*, 2009). Sequences of RP4, and the octopine (Oct) pTiA6 and nopaline (Nop) pTiC58 Ti plasmids were taken from Pansegrau & Lanka (1991). The sequence for octopine Ti plasmid pTi15955 was taken from GenBank NZ_CP032920.1. The octopine T_L OD_{Ti} and T_L RB_{Ti} sequences used in this work in all pSET152 derivatives were taken from pTi15955.

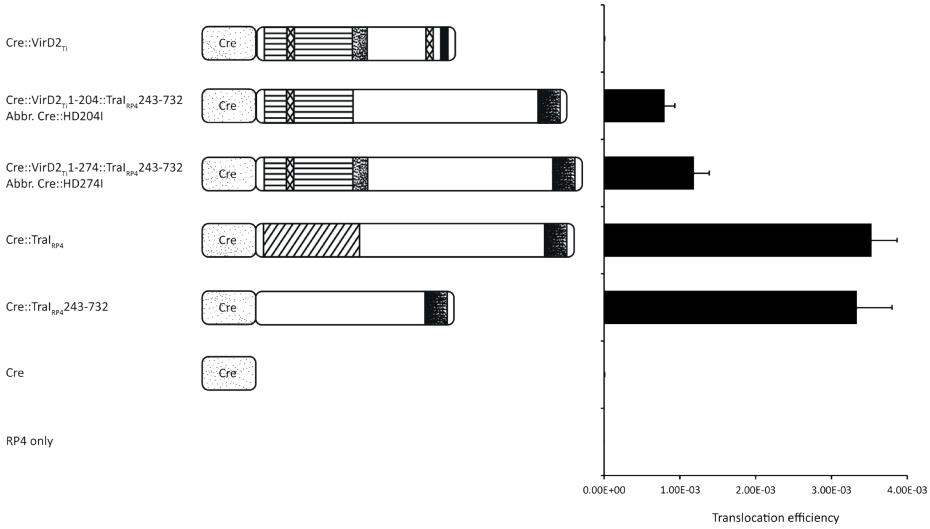


Figure 3. Protein translocation activity mediated by hybrid relaxase using a CRAfT assay. The donor strain was *E. coli* MS411 always containing plasmid RP4, and the recipient strain was *E. coli* CSH26Cm:LTL. Translocation frequency of Cre via the $CP_{RP4}/T4SS_{RP4}$, mediated by fusion to $TraI_{RP4}$ fragments containing TSs or hybrid-relaxase containing the C-terminal $TraI_{RP4}$ fragment was determined. As control no construct, Cre, or Cre fused to $VirD2_{Ti}$ all not containing a $CP_{RP4}/T4SS_{RP4}$ TS were used. The x-axis depicts the translocation efficiency expressed as the fraction of transconjugant cells per total number of recipient cells after co-cultivation. The left side depicts a graphical representation of the fusion proteins and indicates the used plasmids. The Cre protein box is not in size relation to the other protein depictions. The relaxase, their truncation, and hybrid relaxases are depicted in relative size to each other. The translocation frequencies are provided in detail in **table 1.** The translocation frequency is expressed as transformants per total number of donor cells. Co-cultivation time was 2.5 h. The error bar depicts SEM (n = 3). Abbr.: Abbreviation for the hybrid relaxase descriptions. Plasmids used were pCreVirD2 (Cre::$VirD2_{Ti}$), pCreHD204I (Cre::$VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA), pCreHD274I (Cre::$VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA), pCreTraI (Cre::$TraI_{RP4}$1-732AA), pCreTraI243-732 (Cre::$TraI_{RP4}$243-732AA), and CFPBSm (Cre).

Table 1. Translocation frequency of Cre fusion proteins via the $CP_{RP4}/T4SS_{RP4}$ in the CRAfT assay**.** The donor strain was *E. Coli* MS411 containing the constructs detailed below and recipient strain was *E. coli* CSH26CM::LTL. Translocation frequency of Cre and Cre-hybrid proteins was expressed as the fraction of transformants cells per total donor cells. Conjugation frequency of plasmid RP4 was expressed as fraction of transconjugant cells per total donor cells. Co-cultivation time was 2.5 h. The error determined is SEM (n = 3). Plasmids used were pCreVirD2 (Cre::$VirD2_{Ti}$), pCreHD204I (Cre::$VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA), pCreHD274I (Cre::$VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA), pCreTraI (Cre::$TraI_{RP4}$1-732AA), pCreTraI243-732 (Cre::$TraI_{RP4}$243-732AA), and CFPBSm (Cre).

Dtr/CP/T4SS plasmid	Constructs used	Cre translocation frequency	RP4 conjugation frequency
RP4	Cre::$VirD2_{Ti}$	$(4.9 \pm 1.1) \times 10^{-6}$	2.0 ± 0.3
RP4	Cre::$VirD2_{Ti}$1-204AA:$TraI_{RP4}$243-732AA	$(7.9 \pm 1.4) \times 10^{-4}$	2.7 ± 0.4
RP4	Cre::$VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA	$(1.2 \pm 0.2) \times 10^{-3}$	2.8 ± 0.4
RP4	Cre::$TraI_{RP4}$	$(3.5 \pm 0.5) \times 10^{-3}$	2.2 ± 0.3
RP4	Cre::$TraI_{RP4}$243-732	$(3.3 \pm 0.4) \times 10^{-3}$	3.4 ± 0.8
RP4	Cre	$(3.4 \pm 0.8) \times 10^{-6}$	2.2 ± 0.0
RP4	n.a.	$(6.5 \pm 2.0) \times 10^{-8}$	1.9 ± 0.3

n.a.: not applicable

Construction of a vector substrate to measure transfer from *Agrobacterium* to yeast by the hybrid relaxase

For the DNA transfer experiments, derivatives of plasmid pSET152Y*RB* were used as substrate. This plasmid contained the relaxase recognition sites $oriT_{RP4}$ as well as separately the minimal 25 bp RB_{Ti} in combination with an adjacent 38 bp OD_{Ti}. To be able to test transfer to yeast, *ARS6/Cen4* for maintenance in yeast cells and a *KanMX* resistance cassette for selection of the transformed yeast cells using the selection agent geneticin (G418) were added. For clarity, the resulting vector was renamed to pSET152Y*ODRBoriT* (Y: yeast) to indicate that the OD_{Ti} and RB_{Ti} are separately present from the $oriT_{RP4}$ (**Figure S1A**). As controls, further modifications were created, which were either the deletion of the OD_{Ti}/RB_{Ti} (pSET152Y*oriT*Δ*ODRB*) or with a deletion of the $oriT_{RP4}$ (pSET152Y*ODRB*Δ*oriT*). It was shown that pSET152Y*ODRBoriT* was able to be transferred via AMT from *A. tumefaciens* LBA1100 to *S. cerevisiae* (**Figure S1B**; **Chapter 3**), and via RP4 from *A. tumefaciens* LBA2210 to *S. cerevisiae* (**Figure S1B**). For the hybrid relaxase to be able to act optimally, the pSET152Y*ODRBoriT* relaxase binding sites were changed. First the OD_{Ti}/RB_{Ti} sequence was fully deleted. Then in order to provide substrate for the hybrid relaxases HD204/274I the original $oriT_{RP4}$ was replaced with a newly created origin of transfer called *ori*X (X as a placeholder for the modifications made). To this end, a 12 bp sequence containing the nic_{RP4} site in the $oriT_{RP4}$ was removed and replaced by the 25 bp minimal RB_{Ti} sequence of the octopine Ti plasmid, containing the nic_{Ti} site for the $VirD2_{Ti}$ relaxase; the plasmid with this modification was called pSET152Y*oriRB* (*FTWG1*; **Figure 4**). In an initial control experiment, it was tested whether this plasmid would still be mobilized by RP4. As a donor the strain LBA2556 was selected, containing a Ti plasmid with a Δ$virD2_{Ti}$ deletion and therefore lacking the relaxase that could act on the minimal RB_{Ti} in pSET152Y*oriRB*. After the introduction of RP4 and pSET152Y*oriRB* into this strain, it was co-cultivated with *S. cerevisiae* YPH250 yeast cells on IM following the standard AMT protocol. After co-cultivation and selection, G418 positive yeast cell were observed (n = 3; only a qualitative assessment with an estimated transformation efficiency of 10^{-8} to 10^{-9}; data not shown). The initial assumption was that the mobilizable vector could not be transferred, as no $VirD2_{Ti}$ was present to allow transfer via the CP_{Ti}/$T4SS_{Ti}$, and transfer via CP_{RP4}/$T4SS_{RP4}$ could not occur as there was no nic_{RP4} site present to be recognized by the relaxase $TraI_{RP4}$. This experiment, however, indicated that the relaxase $TraI_{RP4}$ aided by the Dtr_{RP4} auxiliary proteins could probably still act on the modified $oriT_{RP4}$ lacking the nic_{RP4} site. This may be due to the similarities between the *nic* sites present in the RB_{Ti} and in $oriT_{RP4}$ (**Figure 1C**). In another initial experiment it was shown that the $TraI_{RP4}$ together with the Dtr_{RP4}/CP_{RP4}/$T4SS_{RP4}$ was not able to transfer a plasmid containing only a OD_{Ti}/RB_{Ti} sequence. For this the plasmid pSET152*ODRB*Δ*oriT* was used in the donor *A. tumefaciens* LBA2556 containing Ti-Δ$virD2_{Ti}$ and RP4 in a co-cultivation with *S. cerevisiae* YPH250 (**Table S2**). To address this, two new origin of transfer substrate modifications were created. First, the 38 bp OD_{Ti} with the binding site for the auxiliary protein $VirC1_{Ti}$ adjacent to the 25 bp minimal RB_{Ti} to boost the nicking activity by $VirD2_{Ti}$ was inserted into the $oriT_{RP4}$. Here again, the minimal RB_{Ti} containing the nic_{Ti} was replacing the 12 bp sequence containing the nic_{RP4} site. This modification resulted in plasmid pSET152Y*oriODRB* (*FTWG2*; **Figure 4**). A second variant was created, that had the 13 bp Dtr_{RP4} $TraJ_{RP4}$ auxiliary protein binding site (*srj*) removed from the $oriT_{RP4}$, and also had the 38 bp OD_{Ti} inserted adjacent to the 25 bp minimal RB_{Ti} sequence replacing the nic_{RP4} site. This modification was called pSET152Y*oriODRB*Δ*srj* (*FTWG3*; **Figure 4**). A detailed sequence of the original $oriT_{RP4}$ and all modified origin of transfer sequences can be found in **table S3**. The *in vivo* function of the OD_{Ti} adjacent to the minimal RB_{Ti} to mediate transfer of a vector by the Dtr_{Ti}/CP_{Ti}/$T4SS_{Ti}$ has been shown earlier (Rolloos *et al.*, 2014; **Chapter 3**). The plasmids containing these new modifications were tested in a similar co-cultivation experiment as described above, with *A. tumefaciens* LBA2556 containing Ti-Δ$virD2_{Ti}$ and RP4 as donor and *S. cerevisiae* YPH250 as recipient. The co-cultivation was performed on IM following the standard AMT protocol to ensure induction and expression of the Dtr_{Ti}

auxiliary proteins. After co-cultivation and selection, no G418 positive yeast cell could be observed (n = 3; data not shown). This experiment revealed that already the extra insertion of the OD_{Ti} sequence was sufficient to prevent $TraI_{RP4}$ from nicking at the nic_{Ti} site.

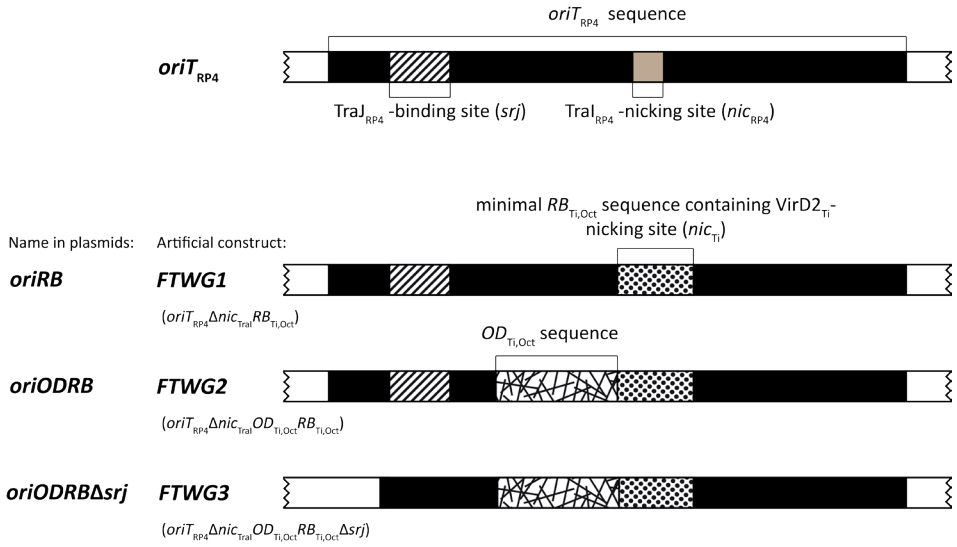


Figure 4. Graphical depiction of $oriT_{RP4}$ and *oriT* modifications used for this study. The structure of the $oriT_{RP4}$ is based on the sequence annotations of Pansegrau *et al.* (1994). The white ends depict the sequences of the plasmid RP4. The $oriT_{RP4}$ is represented as a black bar. The $oriT_{RP4}$ contains a $TraJ_{RP4}$–binding site (*srj*; Ziegelin, Fürste & Lanka, 1989; Pansegrau & Lanka, 1996B), a $TraI_{RP4}$-nicking site ($nic_{RP4;}$ Pansegrau & Lanka, 1991) as well as more poorly defined $TraK_{RP4}$-binding site (*srk*; Ziegelin *et al.*, 1992). The modification *FTWG1* has an insertion of the 25 bp $RB_{Ti,Oct}$ minimal sequence containing the $VirD2_{Ti}$-nicking site (nic_{Ti}; as present on pTi15955; GenBank NZ_CP032920.1) replacing the 12 bp nic_{RP4} site; this modification is called *oriRB* when inserted into plasmids. The modification *FTWG2* has an insertion of the 38 bp $OD_{Ti,Oct}$ sequence (as present on pTi15955; GenBank NZ_CP032920.1) in addition to the modification of *FTWG1*; this modification is called *oriODRB* when inserted into plasmids. The modification *FTWG3* has the 13 bp *srj* site deleted, but contains the $OD_{Ti,Oct}$ sequence (Fürste *et al.*, 1989) and the $RB_{Ti,Oct}$ sequence insertion in place of the nic_{RP4} site; this modification is called *oriODRBΔsrj* when inserted into plasmids. Full sequence details are provided in **table S3**.

Construction of an artificial operon, containing Dtr_{Ti} auxiliary proteins for support of the relaxase function of the hybrid-relaxases $VirD2_{Ti}$ 1-204/274 AA::$TraI_{RP4}$ 243-732 AA

In the previous CRAfT assays the hybrid relaxases HD204/274I were observed to be transferred via the CP_{RP4}/$T4SS_{RP4}$ (**Figure 3**). To test the biological activity of the hybrid relaxases *in vivo*, co-cultivations between *A. tumefaciens* and *S. cerevisiae* were required. In these experiments an *A. tumefaciens* donor strain would be used carrying the RP4 plasmid, but lacking a Ti plasmid. However, without the Ti plasmid and its virulence region present in the donor strain, the Dtr_{Ti} auxiliary proteins $VirC1_{Ti}$, $VirC2_{Ti}$ and $VirD1_{Ti}$ would still be required to support the relaxase function of the $VirD2_{Ti}$. Therefore, artificial operons with the sequences for the proteins $VirC1_{Ti}$, $VirC2_{Ti}$, and $VirD1_{Ti}$ as well as one of the two different hybrid relaxases, HD204I and HD274I or $VirD2_{Ti}$ under the control of the $virF_{Ti}$ promoter were created. The

activity of the artificial operon was first tested by means of a complementation experiment. In this set-up, the open reading frame encoding the natural VirD2$_{Ti}$ was put at the last position within the artificial operon. If VirD2$_{Ti}$ would be expressed, it could be inferred that upstream genes are expressed as well. The operon plasmid was placed in an avirulent *A. tumefaciens* LBA2556 strain containing a Ti plasmid with a deletion for *virD2*$_{Ti}$. The expression of the VirD2$_{Ti}$ protein encoded within the artificial operon, should then complement the function of the transfer of T-DNA from *A. tumefaciens* LBA2556 into plant cells. For T-DNA transfer into plant cells, the binary vector pCAMBIA2301 was used, containing a *gus* cassette flanked by border sequences. In plants, expression of the *gus* cassette would be detectable by histochemical staining. When the artificial operon was present on the expression plasmid, *N. benthamiana* leaf discs co-cultivated with this strain indeed showed X-Gluc staining (**Figure 5**). In leaf discs co-cultivated with the negative control strain with the expression plasmid lacking *virD2*$_{Ti}$, no X-Gluc staining was observed (**Figure 5**). As positive control, leaf discs were co-cultivated with LBA2556 containing an expression plasmid with *virD2*$_{Ti}$ under the control of the *virF*$_{Ti}$ promoter. Here, as expected, blue X-Gluc staining was visible on all leaf discs (**Figure 5**). This experiment consequently provided evidence that a functional VirD2$_{Ti}$ protein was produced in *A. tumefaciens* when its coding region was put at the most distal fourth position on the artificial operon construct using the *virF*$_{Ti}$ promoter. To provide further evidence that the operon was functional, it was attempted to prove expression of the auxiliary proteins using a western blot. For this experiment, VirC1$_{Ti}$, VirC2$_{Ti}$ and VirD2$_{Ti}$ were each additionally equipped with a single FLAG epitope. However, only the expression of VirC1$_{Ti}$ with the size of around 25 kDa (Srinivasan, Yanofsky & Nester, 1989) could be detected and thus provided only evidence that the position 1 of the operon was functional (**Figure S5**).

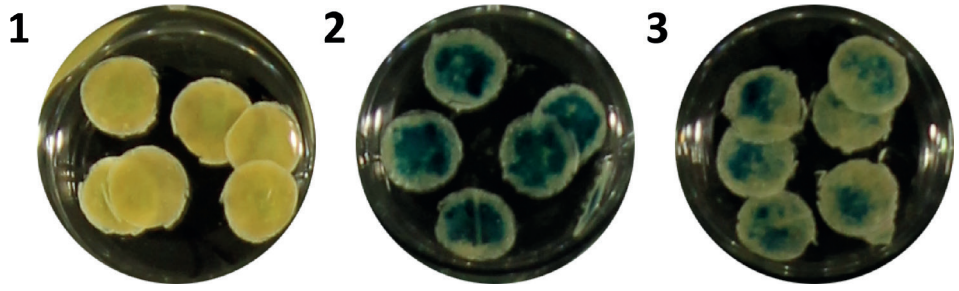


Figure 4. Plasmid transfer mediated by artificial *virC1*$_{Ti}$/*virC2*$_{Ti}$/*virD1*$_{Ti}$/*virD2*$_{Ti}$ operon expression. Agroinfiltration of *N. benthamiana* leaves and subsequent GUS protein expression detection using X-Gluc staining (3 d post-infection). **1.** *A. tumefaciens* LBA2556 (pAL1100Δ*virD2*) containing an empty expression plasmid (pBBRX; pCAMBIA2301) acting as negative control. **2.** *A. tumefaciens* LBA2556 (pAL1100Δ*virD2*) with *virD2*$_{Ti}$ coding complementation plasmid (pBFFVirD2; pCAMBIA2301) acting as positive control. **3.** *A. tumefaciens* LBA2556 (pAL1100Δ*virD2*) with full artificial operon expression plasmid coding for *virC1*$_{Ti}$, *virC2*$_{Ti}$, *virD1*$_{Ti}$ and *virD2*$_{Ti}$ (pBBRXOperon; pCAMBIA2301). For each strain, a representative picture is depicted, but several independent repeats were conducted (n = 3; full data not shown). For each independent repeat, a total of 18 leaf discs was sampled. All of the sampled leaf discs showed the same results as in the representative depiction shown in this figure.

No DNA transfer from *Agrobacterium* to yeast by hybrid-relaxase VirD2$_{Ti}$ 1-274 AA::TraI$_{RP4}$ 243-732 AA in the presence of Dtr$_{Ti}$ auxiliary proteins

DNA transfer by the hybrid relaxase HD274I was tested using a complex set-up to allow the presence of all different constructed components to work together. To this end, it was necessary to introduce the following plasmids into the *A. tumefaciens* LBA202 donor cell: RP4Δ*traI*$_{RP4}$, providing the CP$_{RP4}$ and T4SS$_{RP4}$, the plasmid expressing the operon providing the hybrid relaxase HD274I and the Dtr$_{Ti}$ auxiliary proteins, as well as the pSET152Y-based mobilizable vector containing the modified origin of transfer sequences (pSET152Y*oriODRB* or pSET152Y*oriODRB*Δ*srj*). While the *A. tumefaciens* strain LBA202, C58 wild-type cured from its Ti plasmid, was used as donor strain, the *S. cerevisiae* strain YPH250 was used as a recipient. The co-cultivations were performed on IM containing acetosyringone. Unfortunately, none of the co-cultivations resulted in any observable G418 positive yeast cells (**Table 2**), not only when pSET152Y*ODRB*Δ*oriT* was used, the vector containing an *OD*$_{Ti}$/*RB*$_{Ti,}$ but lacking the full *oriT*$_{RP4}$, but also when using the pSET152Y*oriODRB* or pSET152Y*oriODRB*Δ*srj*. Further, to exclude that transfer of substrate with the modified origin of transfer, containing each a *RB*$_{Ti}$ bound to VirD2$_{Ti}$, would be possible via the CP$_{RP4}$/T4SS$_{RP4}$, *A. tumefaciens* LBA202 as donor strain contained the plasmids: RP4Δ*traI*$_{RP4}$, the operon containing the wild-type *virD2*$_{Ti}$ and either the mobilizable vector pSET152Y*oriODRB* or pSET152Y*oriODRB*Δ*srj*. In this case the substrates containing an modified origin of transfer with an *RB*$_{Ti}$ may be recognized by the VirD2$_{Ti}$ relaxase and its Dtr$_{Ti}$ auxiliary proteins, but the T-strand complex would not be allowed entry into the T4SS$_{RP4}$ via the CP$_{RP4}$. No G418 positive colonies were observed in these control experiments as expected. Aa a positive control, *A. tumefaciens* LBA1100 containing a functional octopine Ti helper plasmid and pSET152Y*ODRBoriT*, containing both *OD*$_{Ti}$/*RB*$_{Ti}$ and *oriT*$_{RP4}$ as a mobilizable vector, was co-cultivated with yeast, to ensure that environmental conditions were suitable for DNA transfer to occur. In this case G418 positive yeast cells were observed after co-cultivation (**Table 2**).

No DNA transfer between *E.coli* strains by hybrid-relaxase VirD2$_{Ti}$ 1-274 AA::TraI$_{RP4}$ 243-732 AA in the presence of Dtr$_{Ti}$ auxiliary proteins

As analogous set-up to the previous co-cultivation experiment between *A. tumefaciens* and *S. cerevisiae*, the possibility of hybrid relaxase-mediated transfer from *E. coli* to *E. coli* was explored. For transfer in-between *E. coli*, the hybrid relaxase HD274I combined with the operon was introduced into to the *E. coli* specific expression plasmid pGZ119EH under the control of the inducible *tac* promoter. As donor strain *E. coli* DH5α was used, containing the following plasmids: RP4Δ*traI* providing the CP$_{RP4}$/T4SS$_{RP4}$, the expression plasmid with the operon coding for the Dtr$_{Ti}$ auxiliary proteins (VirC1$_{Ti}$, VirC2$_{Ti}$ and VirD1$_{Ti}$) and the HD274I, as well as mobilizable vectors pSET152Y*oriODRB* or pSET152Y*oriODRB*Δ*srj*. As recipient strain the streptomycin resistant *E. coli* MS614 was used. The co-cultivation was performed in time-periods ranging from 6 h, 24 h and 48 h. None of the co-cultivation experiments resulted in any apramycin and streptomycin resistant transconjugants, see **table 3**. As functional control *E. coli* DH5α containing RP4 plasmid expressing TraI$_{RP4}$ as well as the pSET152Y*ODRBoriT* mobilizable vector containing a *oriT*$_{RP4}$ was used. Streptomycin and apramycin positive colonies could be observed after the co-cultivation, indicating transfer of the shuttle vector via the CP/$_{RP4}$T4SS$_{RP4}$ into the recipient cells ruling out irregularities within the co-cultivation conditions.

Table 2. Conjugation experiment using hybrid-relaxase HD274I to test for DNA transfer between *A. tumefaciens* to *S. cerevisiae* via CP_{RP4}/$T4SS_{RP4}$ in the presence of Dtr_{Ti} auxiliary proteins. Transformation efficiency was determined as number of transformants per total number of yeast cells. Co-cultivation time was 7 d on IM with 200 µM acetosyringone. The recipient strain was *S. cerevisiae* YPH250. The error determined is SEM (n = 3). *A. tumefaciens* LBA201 contained the wild-type nopaline Ti plasmid pTiC58 and LBA1100 contained the octopine derivative Ti helper plasmid pAL1100.

#	Donor strain	CP/T4SS plasmid	Relaxase	Dtr auxiliary proteins to support relaxase	Vector transfer sequence(s)	Conjugation efficiency (7 d)
1	LBA202	RP4Δ*traI*	HD274I	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*	0
2	LBA202	RP4Δ*traI*	HD274I	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*Δ*srj*	0
3	LBA202	RP4Δ*traI*	HD274I	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	OD_{Ti}; RB_{Ti}; Δ$oriT_{RP4}$	0
4	LBA202	RP4Δ*traI*	$VirD2_{Ti}$	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*	0
5	LBA202	RP4Δ*traI*	$VirD2_{Ti}$	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*Δ*srj*	0
6	LBA201	pTiC58; RP4Δ*traI*	$VirD2_{Ti}$	Dtr_{RP4}; Dtr_{Ti}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	OD_{Ti}; RB_{Ti}; Δ$oriT_{RP4}$	0
7	LBA1100*	pAL1100	$VirD2_{Ti}$	Dtr_{Ti}	OD_{Ti}; RB_{Ti}; $oriT_{RP4}$	$(4.9 \pm 0.5) \times 10^{-5}$

*: functional control

Plasmids used per experiment:

#	Donor strain	CP/T4SS plasmid	Dtr expression plasmid	Transfer vector
1	LBA202	RP4Δ*traI*	pBBRXCmOperonHD274I	pSET152Y*oriODRB*
2	LBA202	RP4Δ*traI*	pBBRXCmOperonHD274I	pSET152Y*oriODRB*Δ*srj*
3	LBA202	RP4Δ*traI*	pBBRXCmOperonHD274I	pSET152Y*ODRB*Δ*oriT*
4	LBA202	RP4Δ*traI*	pBBRXCmOperon	pSET152Y*oriODRB*
5	LBA202	RP4Δ*traI*	pBBRXCmOperon	pSET152Y*oriODRB*Δ*srj*
6	LBA201	pTiC58; RP4Δ*traI*	pBBRXCmOperon	pSET152Y*ODRB*Δ*oriT*
7	LBA1100*	pAL1100	n.a.	pSET152Y*ODRBoriT*

*: functional control; n.a.: not applicable

Table 3. Conjugation experiment using the hybrid-relaxase HD274I to mediated DNA transfer between *E. coli* via the $CP_{RP4}/T4SS_{RP4}$ in the presence of Dtr_{Ti} auxiliary proteins. The used donor strain was *E. coli* DH5α and recipient strain *E. coli* MS614. Time-points were taken after 6 h, 24 h, and 48 h of co-cultivation (n = 3). All strains were cultivated with the addition of 1 mM IPTG. The functional control was only performed with n = 1 per time point. The conjugation efficiency was determined by the total number of donor cells divided by number of transconjugant cells.

#	CP/T4SS plasmid	Relaxase	Dtr auxiliary proteins to support relaxase	Vector transfer sequence(s)	Conjugation efficiency		
					6 h	24 h	48 h
1	RP4Δ*traI*	HD274I	Dtr_{RP4}; VirC1Ti; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*	0	0	0
2	RP4Δ*traI*	HD274I	Dtr_{RP4};$VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriODRB*Δ*srj*	0	0	0
3	RP4Δ*traI*	HD274I	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	RB_{Ti}; Δ$oriT_{RP4}$	0	0	0
4	RP4Δ*traI**	HD274I	Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	ΔOD_{Ti}; ΔRB_{Ti}; Δ$oriT_{RP4}$	0	0	0
5	RP4**	$TraI_{RP4}$	Dtr_{RP4}	OD_{Ti}; RB_{Ti}; $oriT_{RP4}$	n.d.	4.2×10^{-1}	6.9×10^{-1}

*: negative control; **: functional control; n.d.: not determined;

Plasmids used per experiment:

#	Donor strain	CP/T4SS plasmid	Dtr expression plasmid	Transfer vector
1	MS614	RP4Δ*traI*	pGZ119EHOperonHD274I	pSET152Y*oriODRB*
2	MS614	RP4Δ*traI*	pGZ119EHOperonHD274I	pSET152Y*oriODRB*Δ*srj*
3	MS614	RP4Δ*traI*	pGZ119EHOperonHD274I	pSET152Y*ODRB*Δ*oriT*
4	MS614	RP4Δ*traI**	pGZ119EHOperonHD274I	pSET152YΔ*ODRB*Δ*oriT*
5	MS614	RP4**	n.a.	pSET152Y*ODRBoriT*

*: negative control; **: functional control; n.a.: not applicable

Hybrid-relaxase $TraI_{RP4}$ 1-134/210 AA::$VirD2_{Ti}$ 205-424 AA was not able to mediate DNA transfer from *A. tumefaciens* to *S. cerevisiae* via the $CP_{Ti}/T4SS_{Ti}$

To investigate reversely whether $oriT_{RP4}$ containing plasmid transfer via the $CP_{Ti}/T4SS_{Ti}$ would be possible, another hybrid relaxase was created using the putative relaxase domain of $TraI_{RP4}$ combined with the C-terminal part of $VirD2_{Ti}$, containing the TS, but lacking the relaxase domain. The relaxase domain of $TraI_{RP4}$ was not yet described in detail in literature; therefore, two different fragments were used. First, a minimal $TraI_{RP4}$ 1-134 AA relaxase domain was defined; guided by a secondary structure prediction using the CFSSP server (Ashok Kumar, 2013), the further C-terminal domain was cut off right after a possible secondary helix structure (**Figure S6**). The idea behind including a putative minimal relaxase domain was to minimize the length of the folded structure of the domain that would possibly be in contact with the $oriT_{RP4}$. The second $TraI_{RP4}$ relaxase fragment was $TraI_{RP4}$ 1-210 AA, supposedly containing the relaxase domain as a more conservative approach. Both fragments were fused with the C-terminal $VirD2_{Ti}$ 205-424 AA fragment. This created the hybrid relaxases $TraI_{RP4}$ 1-134 AA::$VirD2_{Ti}$ 205-424 AA and $TraI_{RP4}$ 1-210 AA::$VirD2_{Ti}$ 205-424 AA or in short HI134D and HI210D, respectively (**Figure 1B**).

To test the hybrid relaxases HI134D and HI210D, *A. tumefaciens* LBA201 containing a Ti plasmid (nopaline pTiC58) was used as donor strain and *S. cerevisiae* YPH250 as recipient

strain. *A. tumefaciens* LBA201 was equipped with the following plasmids: RP4Δ*traI*, with a deletion for the relaxase $TraI_{RP4}$, the expression plasmid for the hybrid relaxases (HI134D and HI210D) and the mobilizable vector pSET152Y*oriT*Δ*ODRB* vector containing an $oriT_{RP4}$ but no T-DNA border sequences. In this set-up, the $TraI_{RP4}$ relaxase part would be supported in its function by the Dtr_{RP4} proteins provided by the RP4Δ*traI* plasmid. After a 7 d co-cultivation on IM, very few G418 positive yeast colonies could be observed for both HI134 and HI210D (**Table 4**). However, the control using *A. tumefaciens* LBA201 with RP4Δ*traI* combined with an empty expression plasmid pBBRXCm as well as the mobilizable vector pSET152Y*oriT*Δ*ODRB*, exhibited the same low numbers of G418 positive colonies (**Table 4**). This indicated that this transfer was not dependent on the hybrid relaxase. As a functional control to check if DNA transfer to *S. cerevisiae* was possible during the experimental conditions, *A. tumefaciens* LBA1100 containing an octopine helper Ti plasmid and the pSET152Y*ODRBoriT* mobilizable vector containing an RB_{Ti} and $oriT_{RP4}$ sequence, was used as a donor. This resulted in G418 positive colonies after co-cultivation (**Table 4**).

Table 4. Conjugation experiment mediated by the hybrid relaxases HI134D and HI210D between *A. tumefaciens* and *S. cerevisiae* via the $CP_{Ti}/T4SS_{Ti}$ in the presence of $CP_{RP4}/T4SS_{RP4}$ Dtr auxiliary proteins transformation efficiency was determined as transformants per total of yeast cells. Co-cultivation time was 7 d on IM with 200 μM acetosyringone. As recipient strain, *S. cerevisiae* YPH250 was used. The error determined is SEM (n = 3). *A. tumefaciens* LBA201 contained the wild-type nopaline Ti plasmid pTiC58 and LBA1100 contained the octopine derivative Ti helper plasmid pAL1100.

#	Donor strain	CP/T4SS plasmid	Relaxase	Dtr auxiliary proteins to support relaxase	Vector transfer sequence(s)	Conjugation efficiency (7 d)
1	LBA201	pTiC58; RP4Δ*traI*	HI134D	Dtr_{Ti}; Dtr_{RP4}	$oriT_{RP4}$; ΔRB_{Ti}	$(2.0 \pm 1.1) \times 10^{-7}$
2	LBA201	pTiC58; RP4Δ*traI*	HI210D	Dtr_{Ti}; Dtr_{RP4}	$oriT_{RP4}$; ΔRB_{Ti}	$(1.1 \pm 0.4) \times 10^{-7}$
3	LBA201	pTiC58; RP4Δ*traI*	$VirD2_{Ti}$	Dtr_{Ti}; Dtr_{RP4}	$oriT_{RP4}$; ΔRB_{Ti}	$(3.3 \pm 1.3) \times 10^{-7}$
4	LBA201	pTiC58; RP4Δ*traI*	$VirD2_{Ti}$	Dtr_{Ti}; Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	OD_{Ti}; RB_{Ti}; $\Delta oriT_{RP4}$	0
5	LBA1100*	pAL1100	$VirD2_{Ti}$	Dtr_{Ti}	OD_{Ti}; RB_{Ti}; $oriT_{RP4}$	$(1.3 \pm 0.6) \times 10^{-6}$

*: functional control

Plasmids used per experiment:

#	Donor strain	CP/T4SS plasmid	Dtr expression plasmid	Transfer vector
1	LBA201	pTiC58; RP4Δ*traI*	pBBRXCmHI134D	pSET152Y*oriT*Δ*ODRB*
2	LBA201	pTiC58; RP4Δ*traI*	pBBRXCmHI210D	pSET152Y*oriT*Δ*ODRB*
3	LBA201	pTiC58; RP4Δ*traI*	pBBRXCm	pSET152Y*oriT*Δ*ODRB*
4	LBA201	pTiC58; RP4Δ*traI*	pBBRXCmOperon	pSET152Y*ODRB*Δ*oriT*
5	LBA1100*	pAL1100	n.a.	pSET152Y*ODRBoriT*

*: functional control; n.a.: not applicable

Discussion

This study describes the first attempts to study the functional conservation and modularity of the Dtr_{Ti} and Dtr_{RP4} systems. It was reasoned that true understanding of both rather extensively studied systems should in principle allow for *in vivo* assembly of hybrid Dtr systems, with novel but predictable characteristics regarding environmental cues for induction of the systems as well as host range. Several tools and constructs were created in order to test functional interaction of the systems: hybrid-relaxes to be transferred between different CP/T4SS systems, a functional operon to provide Dtr_{Ti} proteins, and mobilizable vectors with specific transfer sequences to be used by the hybrid relaxases. The experimental set-ups were complex and ambitious. Currently, the *in vivo* functionality of hybrid relaxases in the envisioned way could not be proven, but with this study, a new approach is provided to explore this area further.

Initially, the focus was on the different functions of the hybrid relaxases. In particular the relaxase domains themselves and their C-terminal domains carrying a TS, as well as on the possibility to create artificial operons for expression of up to four proteins in *A. tumefaciens*. In principle, the relaxase domain of $VirD2_{Ti}$ should be functional when fused within the hybrids HD204I and HD274I. A truncation of the $VirD2_{Ti}$ relaxase domain using the minimal length of 1-204 AA was able to function as shown previously by a complementation assay by van Kregten *et al.* (2009). In this study, the translocation ability of the C-terminal fragment of $TraI_{RP4}$ was determined by a CRAfT assay (**Table 1**; **Figure 3**). Not only was the $TraI_{RP4}$ C-terminal fragment able to transfer an N-terminally fused Cre protein, but it was also able to translocate fusions with the $VirD2_{Ti}$ minimal relaxase domains (1-204 AA and 1-272 AA) preceded by an N-terminal Cre recombinase. The ability of the artificial operon under control of a $virF_{Ti}$ promoter for expression of its most distal open reading frame was functionally tested using a complementation assay (**Figure 5**). In addition, expression of the first open reading frame for $VirC1_{Ti}$ expression was confirmed via a western blot (**Figure S5**). In conclusion, the operon position 1 and 4 were proven functional, but no clear evidence was found yet for position 2 and 3. It could be possible that the FLAG tagged proteins on position 2 and 3 were expressed below the detection limit of the western blot assay. In some cases, GFP protein expression could be visible using a fluorescent microscope but could not be detected using a western blot assay (Lindhout, 2008). The ORFs present at all of the positions in the operon, as well as the bridges between position 2 and 3 were taken from wild-type sequences, and it was assumed for all of them to be functional.

Subsequently, the functional activity of the hybrid relaxases was tested *in vivo* in a more complex set-up. As discussed before, the individual parts of the experimental set-up were most likely functional. This top-down approach to test all presumably necessary elements was the best choice as a pilot experiment to test the modularity of Dtr systems. This approach was more time and resource saving than a bottom-up approach. A bottom-up approach would require having a larger number of presumably simpler experiments to test each individual component. In the bottom-up approach, each negative result would mean to rest the experimental set-up and adjust each individual tested component again. A downside of the top-down approach was the initial complex experimental set-up, thus for this the most compatible system encompassing all components was required. This created several practical issues, for example compatibility of all used plasmids as well as antibiotic resistance cassettes as well as a non-ideal donor strain. The first tests were aimed at constructing novel functional Dtr hybrid systems in *A. tumefaciens* cells, devoid of as many of unnecessary Dtr_{Ti}/CP_{Ti}/$T4SS_{Ti}$ traits as possible, whilst ensuring a similar chromosomal background for the supplied Dtr_{Ti} auxiliary proteins. Expression constructs for the hybrid relaxases and auxiliary proteins were introduced into the *A. tumefaciens* LBA202 strain. This strain was used to allow the maintenance and antibiotic selection of three different plasmids required for the experiment.

Other more commonly used strains such as *A. tumefaciens* LBA1100 for example and its derivatives for example lacking $virD2_{Ti}$, or the Ti plasmid altogether, already contained genomic antibiotic resistance markers that would have interfered with the plasmid selection. The strain LBA202 was cured of the wild-type nopaline Ti plasmid (pTiC58), but had the plasmid RP4Δ*traI* introduced which is non-transmissible due to lack of $TraI_{RP4}$ expression. It was thus hypothesized that hybrid relaxases, consisting of an N-terminal $VirD2_{Ti}$-derived relaxase domain and a C-terminal $TraI_{RP4}$ TS domain, might then be transported via the $CP_{RP4}/T4SS_{RP4}$ supplied by the RP4 plasmid. For optimal $VirD2_{Ti}$-mediated DNA substrate recognition, two specialized mobilizable vectors pSET152Y*oriODRB* and pSET152Y*oriODRB*Δ*srj* were used, containing no T-DNA border sequences other than the ones present within their differently modified $oriT_{RP4}$ (**Figure 4**). These modifications were necessary as an earlier experiment showed that the initially designed mobilizable vector pSET152Y*oriRB,* in which only the nic_{RP4} was replaced by the minimal $RB_{Ti,Oct}$ sequence containing nic_{Ti}, still showed transfer via $TraI_{RP4}$, and was therefore unsuitable as a substrate for the hybrid-relaxase. One vector had the nic_{RP4} site replaced by an OD_{Ti} to allow for binding of the $T4SS_{Ti}$ Dtr auxiliary proteins, followed by a minimal RB_{Ti} sequence (pSET152Y*oriODRB*). Such an OD_{Ti}/RB_{Ti} type border sequence was shown to be very effective to transfer a plasmid via the $CP_{Ti}/T4SS_{Ti}$ (Rolloos *et al.*, 2014; **Chapter 3**). The other vector had the same replacement of the nic_{RP4} site, but also had a deletion of the $TraJ_{RP4}$-binding site to prevent docking of RP4 specific auxiliary proteins (pSET152Y*oriODRB*Δ*srj*). Along with it, the $virF_{Ti}$ promoter on the hybrid relaxase expression plasmid, normally required to be induced via the regulatory system of the Ti plasmid, was shown to be active independently of the Ti plasmid induction in the context of this plasmid (**Chapter 4**). However, co-cultivation of LBA202/201 with *S. cerevisiae* YPH250 did not yield any DNA transfer (**Table 2**). It is to note that conjugational transfer mediated by RP4 or the related plasmid R751 from bacteria to yeast was reported with an efficiency ranging from 10^{-3} to 10^{-6} (Heinemann & Sprague, 1989; Nishikawa, Suzuki & Yoshida, 1990; Suzuki, Moriguchi & Yamamoto, 2015), this would be around 1000 to 1000000 times less efficient then RP4 transfer between bacterial cells. Considering that the complex experimental set-up may have interfered with the $CP_{RP4}/T4SS_{RP4}$ transfer process, and the transformation efficiency of RP4 to yeast was already low, this raised the question whether the detection limit of the used co-cultivation assay was enough to detect any positive transformants. To increase the sensitivity of the transfer assay and to circumvent any unknown inhibitory factors in the previsioned set-up for the *A. tumefaciens* system, an analogous set-up was tested within an *E. coli* model (**Table 3**). As mentioned before the individual parts, expression of the hybrid relaxase, operon and transfer of the hybrid relaxase via the $CP_{RP4}/T4SS_{RP4}$ were tested to be functional. However, the *E. coli* system also did not provide any evidence for DNA transfer either, suggesting a problem with the hybrid-relaxase transfer via the $CP_{RP4}/T4SS_{RP4}$. It may be possible that the presentation of the relaxase domain towards the substrate is somewhat impaired, or that the fusion of the C-terminal $TraI_{RP4}$ part is masking or inhibiting the relaxase activity of $VirD2_{Ti}$. Further, the activation of the transfer process when the relaxosome complex is presented to the CP_{RP4} could be inhibited due to a different conformation of the hybrid-relaxase. In the experimental set-up containing no Ti plasmid but only the operon plasmid, the Dtr_{Ti} auxiliary protein $VirC1_{Ti}$ could negatively interfering with the presentation of the hybrid relaxosome to the CP_{RP4} $TraG_{RP4}$. Normally, $VirC1_{Ti}$ binds to the OD_{Ti} ahead of the RB_{Ti} and is thus supposed to help to localize the wild-type relaxosome to *A. tumefaciens* poles (Atmakuri *et al.*, 2007), thereby maybe preventing association with the CP_{RP4} $TraG_{RP4}$; but no data is available so far specifying the cellular location of the $CP_{RP4}/T4SS_{RP4}$. In addition, $VirD1_{Ti}$ could potentially interfere with the hybrid relaxosome presentation to the CP_{RP4} $TraG_{RP4}$, shown by an experiment where the nopaline $VirD1_{pTiC58}$ (without $VirD2_{pTiC58}$) inhibited the octopine $VirD1/2_{pTIA6}$ from translocating DNA via the $CP_{Ti,pTiA6}/T4SS_{Ti,pTiA6}$, indicating a potential interaction with the CP_{Ti} $VirD4_{Ti}$ (Steck, Lin & Kado 1990). The auxiliary protein $VirD1_{Ti}$ is essential for the action of $VirD2_{Ti}$ on double-stranded DNA, but not for the nicking of

single-stranded DNA *in vitro* (Pansegrau *et al.*, 1993; Scheiffele, Pansegrau & Lanka 1995). It is still not clear what biochemical activity is determined by $VirD1_{Ti}$. An earlier report that $VirD1_{Ti}$ exhibited topoisomerase activity (Ghai & Das, 1989), was not confirmed with purified enzymes (Scheiffele, Pansegrau & Lanka 1995). An predicted helicase-like domain was shown to be associated with the relaxase protein of the conjugative plasmid R1-16 $TraI_{R1-16}$ (Lang *et al.*, 2010). For CP_{Ti} interactions, no study so far proved or disproved coupling of $VirD1_{Ti}$ with $VirD4_{Ti}$. A recent study (Whitaker *et al.*, 2016) indicated only that $VirD1_{Ti}$ is not able to bind to the elusive $AAD_{VirD4,Ti}$ domain of $VirD4_{Ti}$, but did not address binding to the entire $VirD4_{Ti}$ protein. So far, the function and interactions of $VirD1_{Ti}$ remain unclear.

Another reason for lack of hybrid relaxase transfer could have been caused by using the *A. tumefaciens* strain LBA201/202. As described earlier, it was not possible to use already tested *A. tumefaciens* strain for the co-cultivation due to conflicts of maintaining all the different plasmids for the hybrid relaxase experiment. At the time of experimentation, LBA201 and LBA202 were the only available strains for the co-cultivation. When the parental strain LBA201 still containing a Ti plasmid was used as control, it was not able to conjugate with yeast cells (**Table 2**). It could be that the pTiC58 plasmid in LBA201 would not be possible to transfer DNA to yeast cells. But it was shown that LBA202 derivative strains LBA1010, LBA1100 and LBA2210 were able to transform yeast and plant cells (This study; **Chapter 2**). Due to the complexity of the co-cultivation experiment using the hybrid relaxases, operon and different substrates, too many factors could have influenced and inhibited the transfer of the hybrid-relaxase and not one single cause could be identified. Using the same experimental set-up, using *E. coli* as donor and recipient cells, in order to increase the sensitivity of the assay did not result either in positive transformants. Further, the strain *A. tumefaciens* LBA201 contained a nopaline Ti plasmid (pTiC58), and the RB_{Ti} sequence of the mobilizable vectors were octopine (pTi15955). This should not have been an issue, as nopaline Ti plasmids could transfer octopine T-DNA (Hoekema, Hooykaas & Schilperoort, 1984).

Another approach to study the relaxase exchange, was utilizing the minimal relaxase domain of $TraI_{RP4}$ (1-134 AA or 1-210 AA) fused to the C-terminal part of Vir_{D2} (205-424 AA) containing the TS of $VirD2_{Ti}$ (van Kregten *et al.*, 2009). This hybrid-relaxase could likewise only be tested in *A. tumefaciens* LBA201 background as donor strain. The strain LBA201 contained wild-type nopaline Ti plasmid (pTiC58) to supply the $CP_{Ti}/T4SS_{Ti}$ for transfer of the hybrid relaxases HI134/210D, and it also contained RP4Δ*traI* for providing auxiliary proteins to assist the relaxase function of $TraI_{RP4}$. The substrate for the HI134/210D was a plasmid containing an $oriT_{RP4}$ and no T-DNA border sequences. In none of the experiments proof of hybrid relaxase-mediated DNA transfer was detected (**Table 4**). A low number of colonies were detected, but this was also the case with the negative control strain lacking a hybrid relaxase. However, the positive control in this set-up showed no transfer of plasmid pSET152Y*ODRB*Δ*oriT* containing the OD_{Ti}/RB_{Ti} via the Ti plasmid of *A. tumefaciens* strain LBA201 to yeast. This indicates a possible issue with the nopaline $CP_{Ti}/T4SS_{Ti}$ channel of pTiC58 and its conjugative ability to yeast, (**Tables 2** and **4)**. However, nopaline Ti plasmids are known to transfer also octopine T-DNA to plant cells (Hoekema, Hooykaas & Schilperoort, 1984). Why no transfer was observed, remains a mystery therefore.

Another point in the discussion about interfering factors for the hybrid-relaxases is the stochastic expression balance of proteins involved in a conjugative transfer system, that can also be called the translocation equilibrium; meaning the exact interplay of all involved proteins to ensure a successful and optimal transfer or translocation of DNA and proteins from the donor into the recipient cells. A disruption can have severe effects, for example, excessive *in trans* expression of $TraD_{R1}$ could rescue a negative effect of overexpression of $TraI_{R1}$ in an *E. coli* to *E. coli* conjugation, but not up to wild-type levels (Lang *et al.*, 2010). The authors of this study also pointed out the significance of expression levels for the translocation

equilibrium. In a much earlier study, the addition of overexpression of $VirD1_{pTiC58}$ and $VirD2_{pTiC58}$ led to an increase of plant transformation using a wild-type C58 Ti plasmid (Wang, Herrera-Estrella & van Montagu, 1990). Altogether, these notions might indicate that interfering with the expression stochastics of proteins involved in conjugative plasmid transfer can have profound effects, either negative or positive. However, since the actual expression stoichiometry of any of the transfer system protein assembly is unknown (Christie & Gordon, 2014), it will be impossible to make a dedicated choice for any expression system.

As a last point, the function of the $virF_{Ti}$ promoter on the expression plasmid family pBBRXCm could be questioned. The maternal plasmid pBBR contains a $virF_{Ti}$ promoter shown to be active without induction (**Chapter 4**). As the experimental set-up to check the hybrid-relaxase required the use of multiple plasmids, the gentamycin restriction cassette of pBBRX was exchanged with a chloramphenicol cassette to create the plasmid family pBBRXCm. In addition, no transfer was shown using the hybrid relaxase on *E. coli* expression plasmids with known functionality (**Table 3**), indicating that the expression of the hybrid relaxase would not be the major issue.

In summary, for the creation of a modular Dtr system, more knowledge about the individual proteins involved (of the different transfer systems) is required. Nevertheless, this study showed that the possibility of combining functional elements of different Dtr systems could work, as the hybrid relaxases HD204I and HD274I could be transferred via the CP_{RP4}/$T4SS_{RP4}$, and further introduced the possibility of a functioning operon to supply Dtr auxiliary proteins *in trans* to support the hybrid relaxase. These tools could provide methods for future exploration into this matter.

Material and methods

Organisms and constructs

All organisms used are listed in **table 5** and plasmid constructs in **table 6**. *E. coli* DH5α and DH10β were made competent and transformed using the protocol of Inoue *et al.* (1990). *A. tumefaciens* strains were created according to the electroporation protocol of (den Dulk-Ras & Hooykaas, 1995) or using tri-parental mating (Ditta *et al.*, 1980). All *E. coli* and *A. tumefaciens* strains used in this study were cultured on appropriate antibiotic selection (**Tables 5 and 6)** on LB agar or in LB medium (10 g/L bacto tryptone, 5 g/L bacto yeast extract, 8 g/L NaCl, pH 7.0). *E. coli* was cultured at 37° C, *A. tumefaciens* at 29° C. The yeast strain *S. cerevisiae* YPH250 was cultured on YPD agar or liquid medium (Zonneveld, 1986) with an appropriate selection at 29° C.

Molecular cloning

Standard molecular cloning procedures were performed according to standard laboratory methods (Sambrook & Russell, 2001). Plasmids were contained and amplified in *E. coli* DH5α and DH10β. The plasmid RP4 and its derivatives were isolated from *E. coli* using the plasmid DNA purification kit Nucleobond PC100 (Macherey-Nagel), and from *A. tumefaciens* using the plasmid isolation kit GeneJet Plasmid Midi-prep Kit (Fermentas).

Construction of hybrid-relaxase VirD2$_{Ti}$ 1-204/274 AA::TraI$_{RP4}$ 243-732 AA

To create the hybrid-relaxase, the relaxase domain of VirD2$_{Ti}$ and the C-terminal part of TraI$_{RP4}$ were translationally fused. See **table 7** for all used primers. Two versions of the VirD2$_{Ti}$ functional domain sequence were used to join to the TraI$_{RP4}$ C-terminus. One version utilized the minimal functional domain consisting from 1-204 AA and the second version utilized the minimal functional domain together with the DUF-60 domain from 1-274 AA (van Kregten *et al.*, 2009). Also, van Kregten *et al.* (2009), described the size of the minimal relaxase functioning domain together with the DUF-60 domain as ranging from to 1-264 AA, but for the present cloning procedure the naturally occurring *virD2*$_{Ti}$ sequence encoding the next 10 AA acids was added to the construct, allowing to make use of the BamHI restriction site for the fusion with TraI$_{RP4}$. The C-terminal part of *traI*$_{RP4}$ (containing 243-732 AA) was amplified from the RP4 plasmid, isolated from *A. tumefaciens* LBA2210. The primer combination FTW1, containing a NdeI site and FTW2, containing an EcoRI site was used for fusion with the 204 AA encoding *virD2*$_{Ti}$ minimal fragment and resulted in a 1513 bp PCR fragment. The combination of primer FTW3, containing a BamHI site, and primer FTW2, containing an EcoRI site, was used for the fusion with the 274 AA (including the DUF-60 domain) encoding *virD2*$_{Ti}$ fragment, which finally resulted in a 1489 bp fragment. Both fragments were agarose gel purified and subcloned into pJET1.2 (Fermentas). The NdeI-EcoRI fragment (for the 1-204 AA encoding *virD2*$_{Ti}$ sequence fusion) as well as the BamHI-EcoRI fragment (for the 1-274 AA encoding *virD2*$_{Ti}$ sequence fusion) of pJet1.2 were further cloned into the NdeI-EcoRI digested and BamHI-EcoRI digested pSDM3149 backbone (van Kregten *et al.*, 2009). This generated the plasmid pSDM3149HD204I (VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA) and pSDM3149HD274I (VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA). For nomenclature purpose, the plasmids were named as follows: H for hybrid; D for VirD2$_{Ti}$, followed by a number indicating the length of the VirD2$_{Ti}$ fragment; I for TraI$_{RP4}$. In the plasmid pSDM3149, the

hybrid-relaxase was under the control of the acetosyringone inducible *virD*$_{Ti}$ promoter (AC Vergunst, AD den Dulk-Ras & PJJ Hooykaas, unpublished). The plasmid pSDM3149 as used by van Kregten *et al.* (2009) was based on the modification of the pRL662 (Vergunst *et al.*, 2000), which was based on the pBBR1 plasmid (Kovach *et al.*, 1994). The *virD*$_{Ti}$ promoter and *virD2*$_{Ti}$ sequence were originally derived from octopine Ti plasmid pTiA6 (Rossi, Hohn & Tinland, 1993).

Hybrid-relaxase VirD2$_{Ti}$ 1-204/274 AA::TraI$_{RP4}$ 243-732 AA fusion with the *virF*$_{Ti}$ promoter

To gain non-inducible expression of the fusion proteins independent of the VirA$_{Ti}$/VirG$_{Ti}$ signaling system, the hybrid-relaxase was cloned into a plasmid containing the *virF*$_{Ti}$ promoter. Experimental evidence of the *virF*$_{Ti}$ promoter activity on the pBBR plasmid can be found in **chapter 4**. Both VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA and VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA were amplified via PCR (from plasmids pSDM3149HD204I and pSDM3149HD274I), with forward primer FTW55 and reverse primer FTW56 for both variants (**Table 7**), adding a BspHI and XhoI site and an additional ATG site before the start of *virD2*$_{Ti}$ to compensate for the BspHI cleavage site, to ensure that the ORF was cloned in front of the *virF*$_{Ti}$ promoter. The PCR products were subcloned into pJET1.2. Both the BspHI-XhoI fragments were inserted into the 4380 bp BspHI-XhoI fragment of the pBBR6ΔXho+VirF backbone. This generated the plasmids pBBRHD204I (VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA) and pBBRHD274I (VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA).

Construction of the artificial operon of VirD2$_{Ti}$ and its auxiliary proteins

To supplement VirD2$_{Ti}$ with the specific Dtr$_{Ti}$ auxiliary proteins VirC1$_{Ti}$, VirC2$_{Ti}$, VirD1$_{Ti}$, an artificial operon controlled via the *virF*$_{Ti}$ promoter was constructed. See **table 7** for all used primers. For expression of the operon, the plasmid pBBR6ΔXhoI+VirF was modified. In the 4380 bp BspHI-XhoI fragment, the double-stranded oligonucleotide fragment pBBSX consisting of primer FWpBBSX and RVpBBSX was inserted (Neuteboom *et al.*, 2006), adding a SpeI and XbaI site creating the plasmid pBBRX. The operon was constructed using DNA isolated from *A. tumefaciens* strain LBA1100, that contained the disarmed pAL1100 helper plasmid (octopine pTiB6 ΔT_L, ΔT_R, Δ*tra*, Δ*occ*; Beijersbergen *et al.*, 1992) as a template. For the fragment containing *virC1*$_{Ti}$ and *virC2*$_{Ti}$, the primer pair FTW83 and FTW84 was used. The forward primer FTW83 contained a BspHI site and the reverse primer FTW84 shared a 20 bp homology with the *virD1*$_{Ti}$ and *virD2*$_{Ti}$ fragment. The *virD1*$_{Ti}$ and *virD2*$_{Ti}$ fragment was amplified with the primer pair FTW85/FTW86. FTW85 contained the linker sequence between *virC2*$_{Ti}$ and *virD1*$_{Ti}$. The linker sequence was the natural occurring sequence between *virD2*$_{Ti}$ and *virD3*$_{Ti}$. For initial annealing of the primers to the Ti plasmid 15 cycles of 65° C were performed for all pairs, followed by 25 cycles with 72° C annealing temperature. Each fragment was gel purified. To fuse the *virC1*$_{Ti}$ and *virC2*$_{Ti}$ fragment with the *virD1*$_{Ti}$ and *virD2*$_{Ti}$ fragment equal amounts of each fragment were subjected to a PCR protocol that allowed for annealing of the overhang structures between both fragments and subsequent DNA synthesis without added primers. First, 10 cycles with an annealing temperature of 65° C were followed by 10 cycles with 72° C for the annealing step. Of this mixture, 1 μL of a 10^2 dilution was used for another PCR reaction containing primer pair FTW83 and FTW86 for a final amplification of the annealed fragments with 30 cycles of amplification at an annealing temperature of 72° C. The fragment isolated from an agarose gel, cloned into pJET1.2 and sequenced for verification. The correct operon sequence was removed as a BspHI-SpeI fragment and cloned

into the BspHI-SpeI backbone fragment of pBBRX. This generated pBBRXOperon with a *virF*$_{Ti}$ promoter driven *virC1*$_{Ti}$-*virC2*$_{Ti}$- *virD1*$_{Ti}$-*virD2*$_{Ti}$ artificial operon.

Insertion of hybrid-relaxase VirD2$_{Ti}$ 1-204/274 AA::TraI$_{RP4}$ 243-732 AA into the artificial auxiliary protein operon

The hybrid-relaxase was adjoined with the auxiliary protein operon. The SalI fragment of pBBRXOperon (*virC1*$_{Ti}$-*virC2*$_{Ti}$- *virD1*$_{Ti}$-*virD2*$_{Ti}$) including the *virF*$_{Ti}$ promoter, and the operon including the 379 bp backbone remnant downstream of stop codon of *virD2*$_{Ti}$ was used. This was ligated with the 5451 bp SalI fragment of pBBRHD204I (VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA) and the 5667 bp SalI fragment of pBBRHD274I (VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA), recreating the VirD2$_{Ti}$ relaxase domain variants. This generated the plasmids pBBRXOperonHD204I (VirC1$_{Ti}$-VirC2$_{Ti}$-VirD1$_{Ti}$-VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA) and pBBRXOperonHD274I (VirC1$_{Ti}$-VirC2$_{Ti}$-VirD1$_{Ti}$-VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA).

Plasmid pBBRX antibiotic cassette exchange

The gentamycin resistance cassette (*aacI*) of the pBBRX expression vector was replaced with a chloramphenicol resistance cassette (*cat*). The chloramphenicol resistance cassette was amplified from the plasmid pCAMBIA2200 with the primer pair Cm1 and Cm2 (**Table 7**), using a two-step protocol with 6 cycles using 60° C annealing temperature followed by 30 cycles with combined annealing and extension at 72° C for 2 min. The 1166 bp fragment was inserted into the pJet1.2 plasmid and the chloramphenicol cassette was subsequently removed using a BglII digest and ligated into the BglII digested pBBRX plasmid variations. This created the plasmids pBBRXCm, pBBRXCmOperonHD204I (VirC1$_{Ti}$-VirC2$_{Ti}$-VirD1$_{Ti}$-VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA), pBBRXCmOperonHD274I (VirC1$_{Ti}$-VirC2$_{Ti}$-VirD1$_{Ti}$-VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA) and pBBRXCmOperon (VirC1$_{Ti}$-VirC2$_{Ti}$-VirD1$_{Ti}$-VirD2$_{Ti}$).

Construction of pSET152 based mobilizable vectors with modified *oriT*$_{RP4}$ sequences

The mobilizable plasmid pSET152 was used to test the function of variations made to the *oriT*$_{RP4}$. For construction of pSET152 derivatives, see **chapter 3**. For this study, the plasmid pSET152 derivatives (**Table 6**) were equipped with the autonomous replication and centromeric region sequence *ARS6/Cen4* to allow for stable replication in yeast cells after transfer. The *ARS6/Cen4* sequence was amplified from vector pUC34CPF with primer pair FTW28 and FTW29 (**Table 7**), both containing KpnI sites. The KpnI digested PCR fragment was introduced in the KpnI sites of the following plasmids (Y: indicative of yeast compatible shuttle or mobilizable vector, was added to the names of the resulting plasmid; as well as the *OD*: containing the *OD*$_{Ti}$ sequence): pSET152*RB* (containing an *OD*$_{Ti,Oct}$, *RB*$_{Ti,Oct}$ and *oriT*$_{RP4}$) resulting in the plasmid SET152Y*ODRBoriT* (**Figure S1**). pSET152*oriT*Δ*RB* (containing the *oriT*$_{RP4}$, but no Ti border sequences or *OD*$_{Ti}$) resulting in plasmid pSET152Y*oriT*Δ*ODRB*. Moreover, pSET152Δ*RB*Δ*oriT* (containing no *oriT*$_{RP4}$ and no border sequences or *OD*$_{Ti}$) resulting in plasmid pSET152YΔ*ODRB*Δ*oriT*. The PCR fragment insert was in the same orientation in all plasmids, as checked by restriction digest. The *RB*$_{Ti,Oct}$ border sequences contained in the pSET152 plasmids used in this study, consisted of the 38 bp *OD*$_{Ti}$ sequence adjacent to the minimal 25 bp *RB*$_{Ti,Oct}$, initially taken from the sequence of the octopine Ti plasmid pTi15955 (GenBank NZ_CP032920.1; Rolloos *et al.*, 2014).

To test the function of the hybrid relaxase several variants of the $oriT_{RP4}$ were created to see the influence of a minimal 25 bp $RB_{Ti,Oct}$ and its 38 bp $OD_{Ti,Oct}$ sequence as well as the 13 bp $traJ_{RP4}$ binding site. As $oriT_{RP4}$ reference, the 840 bp ApaI-SphI fragment present on pSET152 was used. The following modifications were created with a detailed sequence description in **figure 4** and **table S3**: *FTWG1* has an insertion of a minimal 25 bp $RB_{Ti,Oct}$ sequence containing the nic_{Ti} site concomitantly deleting a 12 bp sequence in the $TraI_{RP4}$ relaxase binding site removing the nic_{RP4} site. The minimal 25 bp $RB_{Ti,Oct}$ border sequence were taken from the sequence of the octopine Ti plasmid pTi15955 (GenBank NZ_CP032920.1; Rolloos *et al.*, 2014). *FTWG2* has an insertion of the minimal $RB_{Ti,Oct}$ sequence containing the nic_{Ti} site flanked by an $OD_{Ti,Oct}$ sequence (sequence taken from the octopine Ti plasmid pTi15955, GenBank NZ_CP032920.1; Rolloos *et al.*, 2014) with the same 12 bp deletion of the $TraI_{RP4}$ recognition site containing the nic_{RP4} site. *FTWG3* has the same insertion and deletion as *FTWG2,* but additionally has a deletion of the upstream $TraJ_{RP4}$-binding site (*srj*; Fürste *et al.*, 1989; Pansegrau *et al.*, 1994; Pansegrau & Lanka, 1996). The three modified *oriT* variants were synthesized by ShineGene Molecular Biotech (China) and were supplied in the pUC57 plasmid (Fermentas; Accession number: Y14837). The modified $oriT_{RP4}$ sequences were all flanked by ApaI/SphI restriction sites. *FTWG* constructs were removed from the pUC57 backbone using ApaI and SphI, and then inserted into the larger ApaI/SphI fragment of the pSET152Y*oriT*Δ*RB* plasmid, as the ApaI/SphI digestion of the pSET152Y*oriT*Δ*RB* plasmid removed the 840 bp fragment containing the original $oriT_{PR4}$ sequence. This generated the plasmids pSET152Y*oriRB* (from *FTWG1*; *oriRB*; modified $oriT_{RP4}$ with insertion of $RB_{Ti,Oct}$ into the nic_{RP4} site), pSET152Y*oriODRB* (from *FTWG2*; *oriODRB*: modified $oriT_{RP4}$ with insertion of $OD_{Ti,Oct}$ and $RB_{Ti,Oct}$ into the nic_{RP4} site), pSET152Y*oriODRB*Δ*srj* (from *FTWG3*; *oriODRB*Δ*srj*: modified $oriT_{RP4}$ with insertion of $OD_{Ti,Oct}$ and $RB_{Ti,Oct}$ into the nic_{RP4} site, and deletion of the $TraJ_{RP4}$-binding site *srj*).

Introduction of the artificial auxiliary protein operon containing the hybrid-relaxase into an *E. coli* expression vector and FLAG tag insertion

To use the artificial operon containing the ORFs of auxiliary proteins and a relaxase or hybrid relaxase variants in *E. coli* together with RP4 plasmid variants, the operon was transferred into a different expression vector. The *E. coli* expression vector pGZ119EH (Lessl *et al.*, 1992) was used containing an inducible *tac* promoter and *cat* resistance gene. The promoter was induced with IPTG for each experiment. The operon $virC1_{Ti}$-$virC2_{Ti}$-$virD1_{Ti}$-$virD2_{Ti}$ was made available as a SmaI-XbaI fragment and was inserted in similarly digested vector pGZ119EH, creating pGZ119EHOperon (ORFs: $virC1_{Ti}$-$virC2_{Ti}$-$virD1_{Ti}$-$virD2_{Ti}$). Fragments encoding the hybrid relaxase variants were amplified from pBBROperonHD204I and pBBROperonHD274I, using the primer pair FTW117 and FTW118 (**Table 7**). Both PCR fragments were digested with SalI and inserted into the 6044 bp vector fragment of SalI digested pGZ119EHOperon. The correct orientation of the insert was determined by restriction enzyme digestion. These steps created the plasmids pGZ119EHOperonHD204I ($VirC1_{Ti}$-$VirC2_{Ti}$-$VirD1_{Ti}$-$VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA) and pGZ119EGOperonHD274I ($VirC1_{Ti}$-$VirC2_{Ti}$-$VirD1_{Ti}$-$VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA). To check the expression of each of the operon positions, FLAG epitopes were added C-terminally to the auxiliary helper $virC1_{Ti}$, $virC2_{Ti}$ and $virD2_{Ti}$ encoded on the expression plasmid pGZ119EHOperon. No feasible restriction site for an oligo insertion was present on the $virD1_{Ti}$ sequence. The FLAG tags were inserted by double-strand oligo insertion (Neuteboom *et al.*, 2006). For $virC1_{Ti}$ the oligos C1FLAG1 and C1FLAG2 were inserted into the NheI site. For $virC2_{Ti}$ the oligos C2FLAGFW and C2FLAGRV were inserted into the SphI site. For $virD2_{Ti}$ the oligos D2FLAGFW and D2FLAGRV were inserted into the SalI site. This created the artificial operon encoding the following proteins $VirC1_{Ti}$::FLAG, $VirC2_{Ti}$::FLAG, $VirD1_{Ti}$ without a FLAG tag and $VirD2_{Ti}$::FLAG called pGZ119EHOperonFLAG.

Construction of the hybrid-relaxase variants TraI$_{RP4}$ 1-134 AA::VirD2$_{Ti}$ 205-424 AA and TraI$_{RP4}$ 1-210 AA::VirD2$_{Ti}$ 205-424 AA expression plasmids

The relaxase domain of TraI$_{RP4}$ was fused with the TS of VirD2$_{Ti}$. Two versions of the TraI$_{RP4}$ relaxase domain were tested, a minimal version consisting of only 134 AA and a longer version with 210 AA. For both versions, the initial GTG codon of TraI$_{RP4}$ (Pansegrau *et al.*, 1994) was changed into ATG for optimized expression (Pansegrau *et al.* 1990). The TraI$_{RP4}$ fragments were then fused with the C-terminal part of VirD2$_{Ti}$, ranging from 205-424 AA, removing the relaxase domain of VirD2$_{Ti}$ (van Kregten *et al.*, 2009), while in principle allowing for transfer of the hybrid protein by means of the native VirD2$_{Ti}$ TS via the CP$_{Ti}$/T4SS$_{Ti}$. See **table 7** for all used primers. To fuse both halves of the hybrid-relaxase, a two-step PCR method was used with fragments containing homologous overlaps of each other to obtain the final construct. The TraI$_{RP4}$ fragments were amplified using forward primer FTW127 containing a BspHI site, and the reverse primer FTW128 for the 134 AA fragment and FTW129 for the 210 AA fragment from an RP4 plasmid template isolated of *A. tumefaciens* LBA2210. Both reverse primers contain a 25 bp homology with the C-terminal part of the 205-424 AA encoding *virD2*$_{Ti}$ fragment. The *virD2*$_{Ti}$ sequence encoding the 205-424 AA was amplified using forward primer FTW130 with a 25 bp homology to the 1-134 AA encoding *traI*$_{RP4}$ fragment, or with the forward primer FTW131 containing a 25 bp homology to the 1-210 AA encoding *traI*$_{RP4}$ fragment, and both with the reverse primer FTW136 introducing a double-stop codon and an SpeI site. The plasmid pBBRXCmOperon was used as a template. After amplification of the separate fragments, the *traI*$_{RP4}$ and *virD2*$_{Ti}$ parts were fused together via a PCR reaction. Equal amounts of the corresponding fragments were mixed and were combined with primer pair FTW127 and FTW136. Five initial cycles with 60° C annealing temperature were followed by 27 cycles combined annealing and extension in one step (90 s at 72° C), followed by a final 10 min extension at 72° C. For all PCR reactions Fusion polymerase (Thermo Scientific) was used unless indicated otherwise, with cycle times and temperatures according to the manufacturer. That generated the fragments encoding TraI$_{RP4}$1-134AA::VirD2$_{Ti}$205-424AA and TraI$_{RP4}$1-210AA::VirD2$_{Ti}$205-424AA. Both fragments were cleaned up using gel electrophoresis. Afterwards, both fragments were BspHI-SpeI digested and inserted into the BspHI-SpeI digested pBBRXCm. That created the expression plasmids pBBRXCmHI134D (TraI$_{RP4}$1-134AA::VirD2$_{Ti}$205-424AA) and pBBRXCmHI210D (TraI$_{RP4}$1-210AA::VirD2$_{Ti}$205-424AA).

CRAfT-assay plasmids

The plasmids and primers used are described in **table 6** and **table 7**. The Cre fusion plasmids were constructed by inserting *traI*$_{RP4}$ PCR-amplified from RP4 using primer pair CreTraIfw01/CreTraIrev01 between the KpnI/SalI sites of the CFPBSm plasmid (Lang *et al.*, 2010). This created the plasmid pCreTraI, encoding the full-length *traI*$_{RP4}$ sequence (TraI$_{RP4}$1AA-732AA). The N-terminal truncation of TraI$_{RP4}$ (TraI$_{RP4}$243AA-732AA) was created first via the PCR-amplification of RP4 DNA with the primer pair CreTraIfw02/CreTraIrev01 followed by the insertion of the fragment between the KpnI/SalI sites of the CFPBSm creating the plasmid pCreTraI243-732 (TraI$_{RP4}$243AA-732AA). A two-step PCR protocol was used to generate all *traI*$_{RP4}$/*virD2*$_{Ti}$ hybrid plasmids. For both pBBRH204I and pBBRH274I, the primer pair CreVirD2fw01/VirD2KpnIexchangerev was used to amplify the *virD2*$_{Ti}$ fragment removing an unwanted second N-terminal KpnI restriction site within *virD2*$_{Ti}$, and primer pair CreTraIrev01/VirD2KpnIexchangefw to amplify the *traI*$_{RP4}$ fragment. A second PCR was performed on both *traI*$_{RP4}$/*virD2*$_{Ti}$ fragments to be annealed and amplified with the primer pair CreVirD2fw01/CreTraIrev01. The fragments were then ligated into the KpnI digested plasmid CFPBSm creating the plasmids pCreHD204I (Cre::VirD2$_{Ti}$1-204AA::TraI$_{RP4}$243-732AA) and

pCreHD274I (Cre::VirD2$_{Ti}$1-274AA::TraI$_{RP4}$243-732AA). To generate plasmid pCreVirD2 (Cre::VirD2$_{Ti}$1AA-425AA), the SalI digested fragment from pCreHybrid204 was replaced with the SalI/XhoI digested fragment of the plasmid pBBROperon.

CRAfT-assay

To detect translocation of the C-terminal part of TraI$_{RP4}$ via the T4SS$_{RP4}$, a Cre Reporter Assay for Translocation (CRAfT) system (Vergunst *et al.*, 2000) modified by Lang *et al.* (2010) was used. The modifications to the original CRAfT assay by Lang *et al.*, (2010) were only minimal: a single recipient strain was used and different antibiotic selections due to the different selection markers present in the strains. The donor strain *E. coli* MS411 contained the RP4 plasmid as well as the Cre-fusion expression vector CFPBSm (**Table 6**). As the recipient strain, *E. coli* CSH26Cm::LTL was used. See **table 5** for details of the strains used in this assay.

Both strains were cultured in liquid LB medium with appropriate antibiotic selection at 37° C for 24 h. After the 24 h growth, 100 µL of donor strains were transferred into 3 mL fresh LB medium without antibiotics and grown at 37° C for 3 h until reaching an OD_{600} of 0.2. Then an equal amount of donor strain with an OD_{600} of 0.2 was mixed with a recipient strain volume containing a total OD_{600} of 2, to create a ratio of 10 recipient cells per 1 donor cell. The mixture was centrifuged, the pellet was re-suspended in 30 µL LB and transferred onto nitrocellulose filters (0.48 µm pore size) on solid LB medium (no antibiotic selection). The co-cultivation was carried out 37° C for 2.5 h and the filter was subsequently transferred into a micro-reaction tube containing 1 mL LB medium, and vortexed to re-suspend the bacterial biofilm. A dilution series was prepared and plated onto medium containing the respective antibiotics. Donor cells (carbenicillin-positive as they contained RP4) were identified by growth with carbenicillin (100 µg/mL), and X Gal (100 µg/mL) showing blue discoloration of colonies. Transconjugant cells, recipient cells containing RP4 after successful transfer from donor cells, were identified by showing no discoloration on the plate. Recombinants created by uptake of the Cre-fusion proteins were selected on plates containing chloramphenicol (10 µg/mL). The transconjugant and protein translocation frequencies were calculated as recombinants per donor.

Co-cultivation

The co-cultivation with *S. cerevisiae* was performed based on the protocol of Bundock *et al.* (1995), with the following modifications as described in **chapter 2**. *E. coli* to *E. coli* conjugation was performed according to the protocol of Lang *et al.* (2010). Plasmids pGHZ119EH, pGHZ119EHOperon, pGHZ119EHOperonHD204I, pGHZ119EHOperonHD274I were grown with the addition of 1 mM isopropyl-β-d-thiogalactopyranoside (IPTG) to the medium. *S. cerevisiae* selection was performed on YPD plates containing geneticin (G418; 150 µg/ml) and cefotaxamine (200 µg/mL). Selection after conjugation for *E. coli* MS614 was done on plates containing kanamycin (100 µg/mL) and streptomycin (25 µg/ml).

For the co-cultivation experiment (**Table 2**), using the hybrid relaxase HD274I, a specific donor strain *A. tumefaciens* LBA202 cured from its Ti plasmid (octopine pTiC58; **Table 5**) was required to be able to harbor all the different plasmids and their antibiotic selections. As the following plasmids needed to be maintained in the host cell: the RP4 plasmid (containing the resistance markers Km, Tc and Cb), the operon expression plasmid (containing the resistance marker Cm) and the shuttle or mobilizable vector (containing the resistance marker Apra). Using the LBA1100/LBA2210 based strains for this set-up was not

suitable as there were interferences between the plasmids and the already existing antibiotic resistance cassettes. The same issue about resistance marker compatibility was occurring when testing the hybrid relaxases HI134D and HI210D (**Table 4**). For this, the *A. tumefaciens* LBA201 containing a Ti plasmid (nopaline pTiC58) was used (**Table 5**). LBA201 was the precursor of LBA202. The resistance markers of the used plasmids were the same as described above.

pCAMBIA2201 *oriT*$_{RP4}$ insertions

The *oriT*$_{RP4}$ sequence was removed from the plasmid pSET152*RBoriT* using XbaI and NheI. The *oriT*$_{RP4}$ fragment was then inserted into the XbaI and NheI digested pCAMBIA2201 backbone. This generated plasmid pCAMBIA2201*oriT*$_{RP4}$. The insertion was checked by restriction digest analysis.

Agroinfiltration in *Nicotiana benthamiana* and transformation detection

The protocol for agroinfiltration into leaves of 3 weeks old *N. benthamiana* plants and the subsequent detection using quantitative MUG analysis is described in detail in **chapter 2** and was based on Sakalis (2013).

Table 5. Organisms used in this study.

Strain	Description	Resistance	Reference
A. tumefaciens			
LBA201	Wild type C58; nopaline pTiC58, pAtC58		Hooykaas *et al.*, 1980
LBA202	C58 cured of pTiC58; contains cryptic plasmid pAtC58		Hooykaas *et al.*, 1980
LBA1100	LBA202; Rif^R; Nal^R; pAL1100 or octopine pTiB6: ΔT_L, ΔT_R, Δ*tra*, Δ*occ*; pAtC58	Rif, Nal, Spec	Beijersbergen *et al.*, 1992
LBA2556	LBA202; Rif^R; Nal^R; pAL1100Δ$virD2_{Ti}$; pAtC58	Rif, Nal, Spec	Jurado-Jácome, 2011
LBA2210	LBA202; Rif^R; Nal^R; RP4; pAtC58	Rif, Km, Tc, Cb, Nal	PJJ Hooykaas (unpublished)
E. coli			
DH5α	*F-*; *Φ80lacZΔM15*; *Δ(lacZYA-argF); U169*; *recA1*; *endA1*; *hsdR17 (rK–, mK+)*; *phoA*; *supE44*; *λ– thi-1*; *gyrA96*; *relA1*		Grant *et al.*, 1990
DH10β	*F-*; *araDJ39; A(ara, leu)7697*; *AlacX74*; *galU*; *galK*; *rpsL*; *deoR*; *Φ80dlacZAM15*; *endAI*; *nupG*; *recA1*; *mcrA*; *A(mrr hsdRMS, mcrBC)*		Grant *et al.*, 1990
MS411	*ilvG*; *rfb-50*; *thi*		Lang *et al.*, 2010
MS614	*ilvG*; *rfb-50*; *thi*; *rpsL*	Sm	Lang *et al.*, 2010
CSH26CM::LTL	CSH26 galK::cat::loxP-Tet-loxP	Tc	Lang *et al.*, 2010
S. cerevisiae			
YPH250	*MATa*; *ade2-101^{ochre}*; *his3-Δ200*; *leu2-Δ1*; *lys2-801^{amber}*; *trp1-Δ1*; *ura3-52*	n.a.	Sikorski & Hieter, 1989

n.a.: not applicable; Cb: Carbenicillin; Km: Kanamycin; Nal: Nalidixic acid; Rif: Rifampicin; Sm: Streptomycin; Spec: Spectinomycin; Tc: Tetracycline

Table 6. Genetic constructs used in this study.

Plasmid	*oriV*	Description	Resistance	Reference
CFPBSm	*pMB1*	Cre fusion plasmid	Sm	Lang *et al.*, 2010
pCreTraI	*pMB1*	CFPBSm; Cre::$TraI_{RP4}$1-732AA	Sm	This study
pCreTraI243-732	*pMB1*	CFPBSm: Cre::$TraI_{RP4}$243-732AA	Sm	This study
pCreHD204I	*pMB1*	CFPBSm; Cre::$VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA	Sm	This study
pCreHD274I	*pMB1*	CFPBSm; Cre::$VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA	Sm	This study
pCreVirD2	*pMB1*	CFPBSm; Cre::$VirD2_{Ti}$1-425AA	Sm	This study
pBFFVirD2	*pBBR1*	*virD2*$_{Ti}$	Gm	van Kregten *et al.*, 2009
pBBR6ΔXhoI+VirF	*pBBR1*	*virF*$_{Ti}$ promoter; deletion of the second XhoI site of pBBR6	Gm	van Kregten *et al.*, 2009
pBBRX	*pBBR1*	*virF*$_{Ti}$ promoter; oligo insertion adding an additional SpeI and XbaI site	Gm	This study
pSDM3149	*pBBR1*	*virD*$_{Ti}$ promoter; *virD2*$_{Ti}$, both from octopine Ti plasmid pTiA6	Gm	van Kregten *et al.*, 2009
pSDM3149HD204I	*pBBR1*	*virD*$_{Ti}$ promoter; $VirD2_{Ti}$1-204AA:$TraI_{RP4}$243-732AA	Gm	This study
pSDM3149HD274I	*pBBR1*	*virD*$_{Ti}$ promoter; $VirD2_{Ti}$1-274AA::$TraI_{RP4}$243-732AA	Gm	This study
pBBRHD204I	*pBBR1*	*virF*$_{Ti}$ promoter; $VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA	Gm	This study
pBBRHD274I	*pBBR1*	*virF*$_{Ti}$ promoter; $VirD2_{Ti}$1-204AA::$TraI_{RP4}$243-732AA	Gm	This study
pBBRXOperon	*pBBR1*	*virF*$_{Ti}$ promoter; *virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; *virD2*$_{Ti}$; all from octopine pAL1100	Gm	This study
pBBRXOperonHD204I	*pBBR1*	*virF*$_{Ti}$ promoter; *virC1*$_{Ti}$,*virC2*$_{Ti}$,*virD1*$_{Ti}$,$VirD2_{Ti}$1-204AA::$TraI_{RP4}$203-732AA	Gm	This study
pBBRXOperonHD274I	*pBBR1*	*virF*$_{Ti}$ promoter; *virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; $VirD2_{Ti}$1-274AA::$TraI_{RP4}$203-732AA	Gm	This study
pBBRXCm	*pBBR1*	pBBRX with Cam inserted into Gm	Cam	This study
pBBRXCmOperon	*pBBR1*	Cm: *cat*; *virF*$_{Ti}$ promoter; *virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; *virD2*$_{Ti}$	Cam	This study
pBBRXCmOperonHD204I	*pBBR1*	Cm: *cat;* *virF*$_{Ti}$ promoter; *virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; $VirD2_{Ti}$1-204AA::$TraI_{RP4}$203-732AA	Cam	This study
pBBRXCmOperonHD274I	*pBBR1*	Cm: *cat*; *virF*$_{Ti}$ promoter; *virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; $VirD2_{Ti}$1-274AA::$TraI_{RP4}$203-732AA	Cam	This study
pBBRXCmHI134D	*pBBR1*	Cm: *cat*; *virF*$_{Ti}$ promoter; $TraI_{RP4}$1-134AA::$VirD2_{Ti}$205-424AA	Cam	This study
pBBRXCmHI210D	*pBBR1*	Cm: *cat*; *virF*$_{Ti}$ promoter; $TraI_{RP4}$1-210AA::$VirD2_{Ti}$205-424AA	Cam	This study
pGZ119EH	*ColD*	*lacI*^{Q} regulated *pTAC* promoter (IPTG induction required)	Cam	Lessl *et al.*, 1992
pGZ119EHOperon	*ColD*	*virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; *virD2*$_{Ti}$	Cam	This study
pGZ119EHOperonFLAG	*ColD*	*virC1*$_{Ti}$::*FLAG*; *virC2*$_{Ti}$::*FLAG*; *virD1*$_{Ti}$; *virD2*$_{Ti}$::*FLAG*	Cam	This study
pGZ119EHOperonHD204I	*ColD*	*virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; $VirD2_{Ti}$1-204AA::$TraI_{RP4}$203-732AA	Cam	This study

Table continued on the next page.

pGZ119EHOperonHD274I	*ColD*	*virC1*$_{Ti}$; *virC2*$_{Ti}$; *virD1*$_{Ti}$; VirD2$_{Ti}$1-274AA::TraI$_{RP4}$203-732AA	Cam	This study
pSET152*RBoriT*	*pMB1*	lacZαMCS; *oriT*$_{RP4}$; Int$^{\Phi C31}$; T_L *OD*$_{Ti,Oct}$; T_L *RB*$_{Ti,Oct}$ (of pTi15955; GenBank NZ_CP032920.1); *KanMX*;	Apra	This dissertation, Chapter 3
pSET152*RB*Δ*oriT*	*pMB1, pVS1*	*OD*$_{Ti}$; *RB*$_{Ti,Oct}$; Δ*oriT*$_{RP4}$; *KanMX*	Apra	This dissertation, Chapter 3
pSET152*oriT*Δ*RB*	*pMB1, pVS1*	Δ*OD*$_{Ti}$; Δ*RB*$_{Ti,Oct}$; *oriT*$_{RP4}$; *KanMX*	Apra	This dissertation, Chapter 3
pSET152Δ*RB*Δ*oriT*	*pMB1, pVS1*	Δ*OD*$_{Ti}$; Δ*RB*$_{Ti,Oct}$; Δ*oriT*$_{RP4}$; *KanMX*	Apra	This dissertation, Chapter 3
pSET152Y*ODRBoriT*	*pMB1, pVS1*	*OD*$_{Ti}$; *RB*$_{Ti,Oct}$; *oriT*$_{RK2}$, *ARS6/Cen4*, *KanMX*	Apra	This study
pSET152Y*oriT*Δ*ODRB*	*pMB1, pVS1*	Δ*OD*$_{Ti}$; Δ*RB*$_{Ti,Oct}$; *oriT*$_{RP4}$; *ARS6/Cen4*; *KanMX*	Apra	This study
pSET152YΔ*ODRB*Δ*oriT*	*pMB1, pVS1*	Δ*OD*$_{Ti}$; Δ*RB*$_{Ti,Oct}$; Δ*LB*$_{Ti,Oct}$; Δ*oriT*$_{RP4}$; *ARS6/Cen4, KanMX*	Apra	This study
pUC57FTWG1	*pMB1*	*FTWG1*: *oriT*$_{RP4}$Δ*nic*$_{TraI}$*RB*$_{Ti,Oct}$	Cb	This study
pUC57FTWG2	*pMB1*	*FTWG2*: *oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$	Cb	This study
pUC57FTWG3	*pMB1*	*FTWG3*: *oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$Δ*srj*	Cb	This study
pUG34CFP	*pMB1*	pRUL1001; contains *ARS6/Cen4*	Cb	Sakalis, 2013
pSET152Y*oriRB*	*pMB1, pVS1*	*FTWG1*; *ARS6/Cen4*; *KanMX*; *oriT*$_{RP4}$Δ*nic*$_{TraI}$*RB*$_{Ti,Oct}$,	Apra	This study
pSET152Y*oriODRB*	*pMB1, pVS1*	*FTWG2*; *ARS6/Cen4*; *KanMX*; *oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$	Apra	This study
pSET152Y*oriODRB*Δ*srj*	*pMB1, pVS1*	*FTWG3*; *ARS6*/Cen4; KanMX; *oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$Δ*srj*	Apra	This study
RP4	*Inc*Pα	*Mob*$_{RP4}$; *tra*$_{RP4}$; *oriT*$_{RP4}$; isolated from *A. tumefaciens* LBA2210	Km, Tc, Cb	Pansegrau et *al.*, 1994 GenBank BN000925.1
RP4Δ*traI*	*Inc*Pα	Δ*traI*	Km, Tc, Cb	This dissertation, Chapter 3
pBCAMBIA2301	*pBR322, pVS1*	*gusA*; *nptII*; *RB*$_{Ti,Nop}$; *LB*$_{Ti,Nop}$	Km	GenBank AF234316.1
pCAMBIA2200	*pBR322, pVS1*	*cat*; *RB*$_{Ti,Nop}$; *LB*$_{Ti,Nop}$	Cam	GenBank AF234313.1
pCAMBIA2201	*pBR322, pVS1*	*gusA*; *cat*; *RB*$_{Ti,Nop}$; *LB*$_{Ti,Nop}$	Cam	GenBank AF234314.1
pCAMBIA2201*ori*TRP4	*pBR322, pVS1*	*gusA*; *cat; RB*$_{Ti,Nop}$; *LB*$_{Ti,Nop}$; *oriT*$_{RP4}$	Cam	This study
p14P	*IncP*	*RB*$_{Ti,Oct}$; *LB*$_{Ti,Oct}$; *PDC6*; *gusA* (*CaMV 35S* promoter)	Km	This dissertation, Chapter 2
pJET1.2	*pMB1*	*bla (Ap^R)*; *eco47IR*; *P*$_{lacUV5}$; *pT7*; supplied as cloning vector blunt linearized with EcoRV		Fermentas, GenBank EF694056

oriV: Origin of vegetative replication (synonymously with origin of replication *oriRep* or *ori*); *oriT*: Origin of transfer; Amp: Ampicillin; Apra: Apramycin; Cam: Chloramphenicol; Cb: Carbenicillin; Km: Kanamycin; Sm: Streptomycin; Tc: Tetracycline; *oriT*$_{RP4}$ is synonymously with *oriT*$_{RK2}$; Nop: nopaline from pTiC58; Oct: octopine from pTi15955; Y: *ARS6/Cen4*; *RB*$_{Ti,Oct}$: Right Border pTi$_{Oct}$; *LB*$_{Ti,Oct}$: Left Border pTi$_{Oct}$; *OD*$_{Ti,Oct}$: Overdrive pTi$_{Oct}$ (Rolloos *et al.*, 2014); *srj*: TraJ$_{RP4}$-binding site

Table 7. Oligonucleotide sequences used in this study.

FTW1: ACATATGCTGAACATCGCCGCCTTGAGCGGATGCCGAAGCTCGAAGCCCGATTCGGGGCATTCACGC (underlined NdeI restriction site)
FTW2: TGAATTCTCATCTACTCCTACCTCGGGTAGTTTTAA (underlined EcoRI restriction site)
FTW3: AGGATCCAACCGAAGCTCGAAGCCCGATTCGGGGC (underlined BamHI restriction site)
FTW28: AGGTACCCCACCTGGGTCCTTTTCATCACGTG (underlined KpnI restriction site)
FTW29: AGGTACCTAATGGTTTCTTAGGACGGATCGCTTGC (underlined KpnI restriction site)
FTW55: ATCATGATGCCCGATCGCGCTCAAGTAATCAT (underlined is the BspHI restriction site)
FTW56: GCTCGAGACCTTATCATCATCTACTC**CTA**CCTCGGGT (underlined is the XhoI restriction site, bold is a stop codon)
FWpBBSX: CATGATAATAG**ACTAGT***CTAGA*C (underlined overlap new BspHI site, bold is SpeI site, Italics Is XbaI site, wave underlined double stop codon)
RVpBBSX: TCGAGTCTAGACTAGTCTATTAT (see FWpBBSX for description)
FTW83 FW C1C2: CGATA**TCATGA**CGCAACTTTTGACGTTTTG (in bold BspHI restriction site, underlined *virC1*$_{Ti}$ homology sequence)
FTW84 RV C1C2: **ATTATTCGGTCCTTCCTGTCT**CTACCAATTCCTCGATGGCT (bold is overlap with *virD1*$_{Ti}$::*virD2*$_{Ti}$ fragment, underlined *virC2*$_{Ti}$ homology)
FTW85 RV C1C2: **TAG**AGACAGGAAGGACCGAATAATGTCAAAACACACCAGAGTCAC (bold modified stop codon for *virC2*$_{Ti}$, wave underlined is the linker sequence of *virD2*$_{Ti}$ and *virD3*$_{Ti}$, underlined *virD1*$_{Ti}$ homology)
FTW86 RV D1D2: : CCTTCTAG**ACTAGT**CTATTACTAGGTCCCCCCGCG (bold SpeI restriction site, underlined *virD2*$_{Ti}$ homology)
FTW117: GCCGCTATAACTATCTGACAGCCTA (contains no restriction site, but is upstream of the VirD2$_{Ti}$ SalI restriction site)
FTW118: AGATGTCGACTCTAGATTATCATCATCTACTCCTAC (underlined is the SalI restriction site)
FTW127: ATCATGATTGCCAAGCACGTCCCCATGCGCTC (underlined is the BspHI restriction site)
FTW128: **CTTCGAATTGAATCTTTTGAGCCTG**CGGCTCATGGATGGTGTTTCGGGTC (in bold is the 25bp homology VirD2$_{Ti}$ and underlined is the 25bp homology TraI$_{RP4}$)
FTW129: **CTTCGAATTGAATCTTTTGAGCCTG**GTTTTCCCGCAGGACGCGGTGCAGG (in bold is the 25bp homology VirD2$_{Ti}$ and underlined is the 25bp homology TraI$_{RP4}$)
FTW130: **GACCCGAAACACCATCCATGAGCCG**CAGGCTCAAAAGATTCAATTCGAAG (in bold is the 25bp homology TraI$_{RP4}$ and underlined is the 25bp homology of VirD2$_{Ti}$)
FTW131: **CCTGCACCGCGTCCTGCGGGAAAAC**CAGGCTCAAAAGATTCAATTCGAAG (in bold is the 25bp homology TraI$_{RP4}$ and underlined is the 25bp homology of VirD2$_{Ti}$)
FTW136: CGTACTAGTTCATCAGGTCCCCCCGCGCCCATCGTCGC (underlined is the homology with VirD2$_{Ti}$)
CreTraIfw01: GCAGGTACCGTGATTGCCAAGCACGTCC (underlined KpnI restriction site)
CreTraIfw02: GCAGGTACCCCGAAGCTCGAAGCCCGATTCGG (underlined KpnI restriction site)
CreTraIrev01: GCAGTCGACTCATCTACTCCTACCTCGGGTAGTTTTAAGGG (underline SalI restriction site)
CreVirD2fw01: GCAGGTACCATGCCCGATCGAGCTCAAGTTATCATTCGC (underlined KpnI restriction site)
Cm1: ATTACCATGGATATATCTCCCAATTTGTGTAGGGC
Cm2: TAATCCATGGTGATCAATTCGGGCACGAAC
C1FLAGFW: GCCGACTACAAGGACGACGACGATAAGCGCATG
C1FLAGRV: CGCTTATCGTCGTCGTCCTTGTAGTCGGCCATGC
C2FLAGFW: GCTAGGATGGCCGACTACAAGGACGACGACGATAAGCGGG
C2FLAGRV: CTAGCCCGCTTATCGTCGTCGTCCTTGTAGTCGGCCAT
D2FLAGFW: TCGACATGGCCGACTACAAGGACGACGACGATAAGCGG
D2FLAGRV: TCGACCCGCTTATCGTCGTCGTCCTTGTAGTCGGCCAT

All oligonucleotides are depicted in 5´-3´directions

Acknowledgements

This work is part of the research programme Green Genetics with project number 828.11.001, which is (partly) financed by the Dutch Research Council (NWO).

References

Albright LM, Yanofsky MF, Leroux B, Ma DQ & Nester EW (1987) Processing of the T-DNA of *Agrobacterium tumefaciens* generates border nicks and linear, single-stranded T-DNA. *J. Bacteriol.* **169:** 1046–1055

Alvarez-Martinez CE & Christie PJ (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol. Mol. Biol. Rev.* **73:** 775–808

Ashok Kumar T (2013) CFSSP: Chou and Fasman secondary structure prediction server. *Wide Spectr.* **1:** 15–19

Atmakuri K, Cascales E, Burton OT, Banta LM & Christie PJ (2007) *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *EMBO J.* **26:** 2540–2551

Babic A, Lindner AB, Vulic M, Stewart EJ & Radman M (2008) Direct visualization of horizontal gene transfer. *Science* **319:** 1533–1536

Balzer D, Pansegrau W & Lanka E (1994) Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.* **176:** 4285–4295

Beijersbergen AGM, den Dulk-Ras AD, Schilperoort RA & Hooykaas PJJ (1992) Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256:** 1324–1327

Bierman M, Logan R & O'Brien K (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces spp*. *Gene* **116:** 43–49

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

Cabezón E, Lanka E & de la Cruz F (1994) Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: *trwB* and RP4 *traG* are interchangeable. *J. Bacteriol.* **176:** 4455–4458

Chou PY & Fasman GD (1974A) Prediction of protein conformation. *Biochemistry* **13:** 222–245

Chou PY & Fasman GD (1974B) Conformational parameters for amino acids in helical, β-Sheet, and random coil regions calculated from proteins. *Biochemistry* **13:** 211–222

Christie PJ & Gordon JE (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Christie PJ, Whitaker N & González-Rivera C (2014) Mechanism and structure of the bacterial type IV secretion systems. *Biochim. Biophys. Acta - Mol. Cell Res.* **1843:** 1578–1591

Cole SP & Guiney DG (1995) Site-specific mutations in the *traI* relaxase and upstream region of plasmid RP4. *Plasmid* **34:** 236–239

Ditta G, Stanfield S, Corbin D & Helinski DR (1980) Broad host range DNA cloning system for Gram-negative bacteria: construction of a gene bank of *Rhizobium meliloti*. *Proc. Natl. Acad. Sci. U.S.A.* **77:** 7347–7351

den Dulk-Ras AD & Hooykaas PJJ (1995) Electroporation of *Agrobacterium tumefaciens*. *Methods Mol. Biol.* **55:** 63–72

Dürrenberger F, Crameri A, Hohn B & Koukolikova-Nicola Z (1989) Covalently bound VirD2 protein of *Agrobacterium tumefaciens* protects the T-DNA from exonucleolytic degradation. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 9154–9158

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Fernández-López C, Bravo A, Ruiz-Cruz S, Solano-Collado V, Garsin DA, Lorenzo-Díaz F & Espinosa M (2014) Mobilizable rolling-circle replicating plasmids from Gram-positive bacteria: a low-cost conjugative transfer. *Microbiol. Spectr.* **2:** 257–276

Fürste JP, Pansegrau W, Ziegelin G, Kroger M & Lanka E (1989) Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 1771–1775

Gao R & Lynn DG (2005) Environmental pH sensing: resolving the VirA/VirG two-component system inputs for *Agrobacterium* pathogenesis. *J. Bacteriol.* **187:** 2182–2189

Ghai J & Das A (1989) The virD operon of *Agrobacterium tumefaciens* Ti plasmid encodes a DNA-relaxing enzyme. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 3109–3113

Grant SG, Jessee J, Bloom FR & Hanahan D (1990) Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. *Proc. Natl. Acad. Sci. U.S.A.* **87:** 4645–4649

Guiney DG & Yakobson EA (1983) Location and nucleotide sequence of the transfer origin of the broad host range plasmid RK2. *Proc. Natl. Acad. Sci. U.S.A.* **80:** 3595–3598

Hajdukiewicz P, Svab Z & Maliga P (1994) The small, versatile pPZP family of *Agrobacterium* binary vectors for plant transformation. *Plant Mol. Biol.* **25:** 989–994

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Heinemann JA & Sprague GF (1989) Bacterial conjugative plasmids mobilize DNA transfer between bacteria and yeast. *Nature* **340:** 205–209

Hoekema A, Hooykaas PJJ & Schilperoort RA (1984) Transfer of the octopine T-DNA segment to plant cells mediated by different types of *Agrobacterium* tumor- or root-inducing plasmids: generality of virulence systems. *J. Bacteriol.* **158:** 383–385

Hooykaas PJJ, den Dulk-Ras H, Ooms G & Schilperoort RA (1980) Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. *J. Bacteriol.* **143:** 1295–1306

Howard EA, Winsor BA, De Vos G & Zambryski P (1989) Activation of the T-DNA transfer process in *Agrobacterium* results in the generation of a T-strand-protein complex: tight association of VirD2 with the 5' ends of T-strands. *Proc. Natl. Acad. Sci. U.S.A.* **86:** 4017–4021

Inoue H, Nojima H & Okayama H (1990) High efficiency transformation of *Escherichia coli* with plasmids. *Gene* **96:** 23–28

Jefferson RA, Kavanagh TA & Bevan MW (1987) GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *EMBO J.* **6:** 3901–3907

Jen GC & Chilton MD (1986) Activity of T-DNA borders in plant cell transformation by mini-T plasmids. *J. Bacteriol.* **166:** 491–499

Jurado-Jácome E (2011) *Agrobacterium* infection: translocation of virulence proteins and role of VirF in host cells. Doctoral dissertation, Chapter **2**:39-60, Leiden University

Kovach ME, Phillips RW, Elzer PH, Roop RM & Peterson KM (1994) pBBR1MCS: a broad-host-range cloning vector. *Biotechniques* **16:** 800–802

Lang S, Gruber CJ, Raffl S, Reisner A & Zechner EL (2014) Common requirement for the relaxosome of plasmid R1 in multiple activities of the conjugative type IV secretion system. *J. Bacteriol.* **196:** 2108–2121

Lang S, Gruber K, Mihajlovic S, Arnold R, Gruber CJ, Steinlechner S, Jehl MA, Rattei T, Fröhlich KU & Zechner EL (2010) Molecular recognition determinants for type IV secretion of diverse families of conjugative relaxases. *Mol. Microbiol.* **78:** 1539–1555

Lanka E & Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* **64:** 141–169

Lessl M, Balzer D, Lurz R, Waters VL, Guiney DG & Lanka E (1992) Dissection of IncP conjugative plasmid transfer: definition of the transfer region *tra2* by mobilization of the *tra1* region *in trans*. *J. Bacteriol.* **174:** 2493–2500

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

Lin Y, Binns AN & Lynn DG (2008) The initial steps in *Agrobacterium tumefaciens* pathogenesis: Chemical Biology of Host Recognition. In *Agrobacterium: from biology to biotechnology* pp 221–241. New York, NY: Springer New York

Lindhout BI (2008) Regulatory DNA binding peptides as novel tools for plant functional genomics. Doctoral dissertation, Chapter **3**:55-73, Leiden University

Lu J, den Dulk-Ras AD, Hooykaas PJJ, and J. N. Mark Glover JNM (2009) *Agrobacterium tumefaciens* VirC2 enhances T-DNA transfer and virulence through its C-terminal ribbon–helix–helix DNA-binding fold. *Proc. Natl. Acad. Sci. U.S.A.* **106**: 9643-9648

Neuteboom LW, Lindhout BI, Saman IL, Hooykaas PJJ & van der Zaal BJ (2006) Effects of different zinc finger transcription factors on genomic targets. *Biochem. Biophys. Res. Commun.* **339:** 263–270

Nishikawa M, Suzuki K & Yoshida K (1990) Structural and functional stability of IncP plasmids during stepwise transmission by *trans*-kingdom mating: promiscuous conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*. *Japanese J. Genet.* **65:** 323–334

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

Pansegrau W & Lanka E (1996A) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W & Lanka E (1996B) Mechanisms of initiation and termination reactions in conjugative DNA processing. J. Biol. Chem. **271**: 13068–13076

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Pansegrau W, Schoumacher F, Hohn B & Lanka E (1993) Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proc. Natl. Acad. Sci. U.S.A.* **90:** 11538–11542

Pansegrau W, Schröder W & Lanka E (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci. U.S.A.* **90:** 2925–2929

Pansegrau W, Schröder W & Lanka E (1994) Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.* **269:** 2782–2789

Parker C & Meyer R (2005) Mechanisms of strand replacement synthesis for plasmid DNA transferred by conjugation. *J. Bacteriol.* **187:** 3400–3406

Relić B, Andjelković M, Rossi L, Nagamine Y & Hohn B (1998) Interaction of the DNA modifying proteins VirD1 and VirD2 of *Agrobacterium tumefaciens*: analysis by subcellular localization in mammalian cells. *Proc. Natl. Acad. Sci. U.S.A.* **95:** 9105–9110

Rolloos M, Dohmen MHC, Hooykaas PJJ & van der Zaal BJ (2014) Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. *Mol. Microbiol.* **91:** 1240–1251

Rossi L, Hohn B & Tinland B (1993) The VirD2 protein of *Agrobacterium tumefaciens* carries nuclear localization signals important for transfer of T-DNA to plant. *Mol. Gen. Genet.* **239:** 345–353

Sakalis PA (2013) Visualizing virulence proteins and their translocation into the host during *Agrobacterium*-mediated transformation. Doctoral dissertation, Chapter **4**:125-150, Leiden University

Sambrook J & Russell DW (2001) Molecular cloning: a laboratory manual 3rd ed. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

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

Schmitz DJ, Ali Z, Wang C, Aljedaani F, Hooykaas PJJ, Mahfouz M & de Pater S (2020) CRISPR/Cas9 mutagenesis by translocation of Cas9 protein into plant cells via the *Agrobacterium* type IV secretion System. *Front. Genome Ed.* **2:** 1–11

Schrammeijer B, Beijersbergen AGM, Idler KB, Melchers LS, Thompson D V & Hooykaas PJJ (2000) Sequence analysis of the *vir*-region from *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *J. Exp. Bot.* **51:** 1167–1169

Schrammeijer B, den Dulk-Ras AD, Vergunst AC, Jurado-Jácome E & Hooykaas PJJ (2003) Analysis of Vir protein translocation from *Agrobacterium tumefaciens* using *Saccharomyces cerevisiae* as a model: evidence for transport of a novel effector protein VirE3. *Nucleic Acids Res.* **31:** 860–868

Sikorski RS & Hieter P (1989) A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122:** 19–27

Srinivasan, Yanofsky MF & Nester EW (1989) Structural genes *virC1* and *virC2* of the host range determining *virC* operon are determinants of virulence in *Agrobacterium tumefaciens*. *J. Genet.* **68:** 23–28

Stachel SE, Timmerman B & Zambryski P (1987) Activation of *Agrobacterium tumefaciens vir* gene expression generates multiple single-stranded T-strand molecules from the pTiA6 T-region: requirement for 5′ *virD* gene products. *EMBO J.* **6:** 857–863

Steck TR, Lin TS & Kado CI (1990) *virD2* gene product from the nopaline plasmid pTiC58 has at least two activities required for virulence. *Nucleic Acids Res.* **18:** 6953–6958

Strand TA, Lale R, Degnes KF, Lando M & Valla S (2014) A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PLoS One* **9:** 1–6

Suzuki K, Moriguchi K & Yamamoto S (2015) Horizontal DNA transfer from bacteria to eukaryotes and a lesson from experimental transfers. *Res. Microbiol.* **166:** 753–763

Toro N, Datta A, Yanofsky M & Nester EW (1988) Role of the overdrive sequence in T-DNA border cleavage in *Agrobacterium*. *Proc. Natl. Acad. Sci. U.S.A.* **85:** 8558–8562

Tinland B, Hohn B & Puchta H (1994) *Agrobacterium tumefaciens* transfers single-stranded transferred DNA (T-DNA) into the plant cell nucleus. *Proc. Natl. Acad. Sci. U.S.A.* **91:** 8000–8004

van Duyne GD (2015) Cre recombinase. *Microbiol. Spectr.* **3:** MDNA3-0014–2014

van Kregten M, Lindhout BI, Hooykaas PJJ & van der Zaal BJ (2009) *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Mol. Plant-Microbe Interact.* **22:** 1356–1365

Vergunst AC, Schrammeijer B, den Dulk-Ras AD, De Vlaam CMT, Regensburg-Tuink AJG & Hooykaas PJJ (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* **290:** 979–982

Vergunst AC, van Lier MCM, den Dulk-Ras AD, Grosse Stüve TA, Ouwehand A & Hooykaas PJJ (2005) Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proc. Natl. Acad. Sci. U.S.A.* **102:** 832–837

Vogel AM & Das A (1992) Mutational analysis of *Agrobacterium tumefaciens* virD2: tyrosine 29 is essential for endonuclease activity. *J. Bacteriol.* **174:** 303–308

Wang K, Herrera-Estrella A & van Montagu M (1990) Overexpression of *virD1* and *virD2* genes in *Agrobacterium tumefaciens* enhances T-complex formation and plant transformation. *J. Bacteriol.* **172:** 4432–4440

Waterhouse AM, Procter JB, Martin DMA, Clamp M & Barton GJ (2009) Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25:** 1189–1191

Waters VL, Hirata KH, Pansegrau W, Lanka E & Guiney DG (1991) Sequence identity in the nick regions of IncP plasmid transfer origins and T-DNA borders of *Agrobacterium* Ti plasmids. *Proc. Natl. Acad. Sci. U.S.A.* **88:** 1456–1460

Whitaker N, Berry TM, Rosenthal N, Gordon JE, Gonzalez-Rivera C, Sheehan KB, Truchan HK, VieBrock L, Newton ILG, Carlyon JA & Christie PJ (2016) Chimeric coupling proteins mediate transfer of heterologous type IV effectors through the *Escherichia coli* pKM101-encoded conjugation machine. *J. Bacteriol.* **198:** 2701–2718

Yadav NS, Vanderleyden J, Bennett DR, Barnes WM & Chilton M-D (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. U.S.A.* **79:** 6322–6326

Yakobson EA & Guiney DG (1984) Conjugal transfer of bacterial chromosomes mediated by the RK2 plasmid transfer origin cloned into transposon Tn5. *J. Bacteriol.* **160:** 451–453

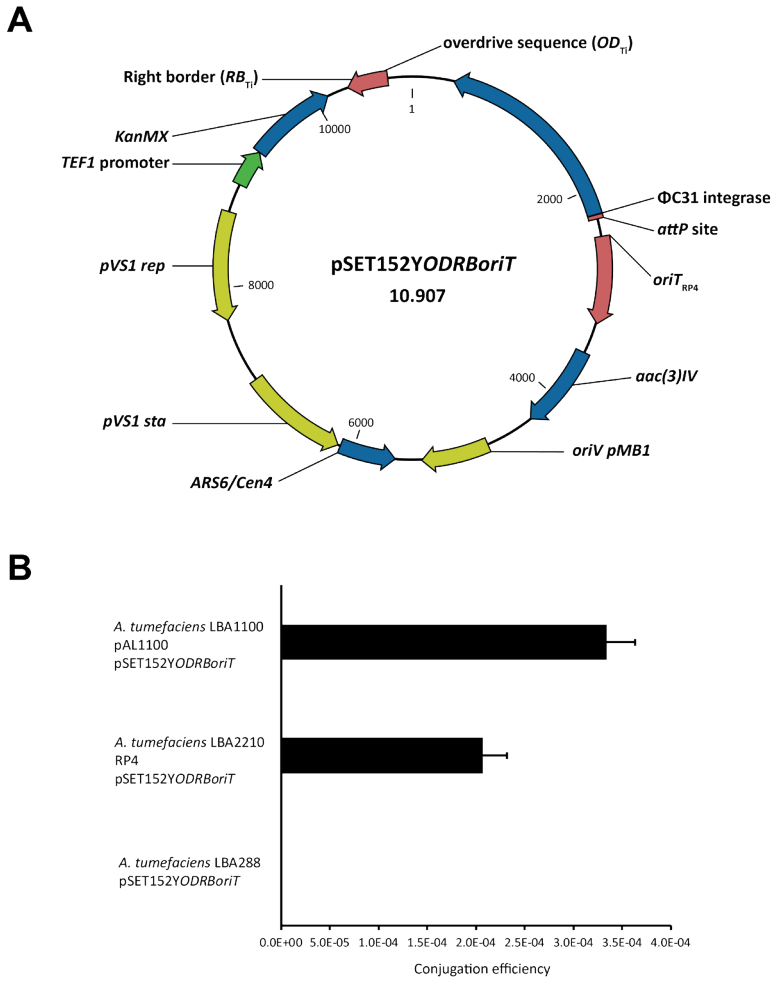
Ziegelin G, Fürste JP & Lanka E (1989) TraJ protein of plasmid RP4 binds to a 19-base pair invert sequence repetition within the transfer origin. *J. Biol. Chem.* **264:** 11989–94

Ziegelin G, Pansegrau W, Lurz R & Lanka E (1992) TraK protein of conjugative plasmid RP4 forms a specialized nucleoprotein complex with the transfer origin. *J. Biol. Chem.* **267:** 17279–17286

Zonneveld BJM (1986) Cheap and simple yeast media. *J. Microbiol. Methods* **4:** 287–291

Zupan JR & Zambryski P (1995) Transfer of T-DNA from *Agrobacterium* to the plant cell. *Plant Physiol.* **107:** 1041–1047

Supplemental information



Supplemental Figure S1. A: Shuttle vector pSET152Y*ODRBoriT*. The vector encodes a ΦC31 phage integrase, able to integrate the entire vector into the host genome via site-specific recombination. The integrase recognition sequence on the shuttle vector is *attP*, and in the *Streptomyces* genome *attB* (Kuhstoss & Rao, 1991). It further contains a *pVS1* minimal replicon (*sta* and *rep*) for *A. tumefaciens*, a replicon for *E. coli*, a *KanMX* resistance cassette for selection in yeast and the *ARS6/Cen4* autonomous replication and centromeric region sequence to allow for stable replication use in yeast, an apramycin resistance cassette (*aac(3)IV*), a 25 bp octopine-type right border (RB_{Ti}) recognition sequence for $VirD2_{Ti}$ combined with an upstream adjacent 38 bp overdrive (OD_{Ti}) sequence for Dtr_{Ti} auxiliary protein recognition (Rolloos *et al.*, 2014), and a RP4 based origin of transfer ($oriT_{RP4}$) for mobilization by RP4 (Mazodier, Petter & Thompson, 1989). The shuttle vector pSET152Y*oriT*Δ*ODRB* just lacks the OD_{Ti}/RB_{Ti} (vector map not shown). The shuttle vector pSET152*ODRB*Δ*oriT* contains a deletion of the $oriT_{RP4}$ region (vector map not shown). **B:** Transformation efficiencies of shuttle vector pSET152Y*ODRBoriT* by AMT or RP4-mediated transfer from *A. tumefaciens* to *S. cerevisiae*. For AMT, *A. tumefaciens* LBA1100 and for RP4-mediated transfer LBA2210 was used. As recipient *S. cerevisiae* YPH250 was used. For AMT the co-cultivation was performed at 21°C for 3 d on IM containing acetosyringone (transformation efficiency: $(2.7 \pm 0.3) \times 10^{-4}$) and for RP4-mediated transfer the co-cultivation was performed at 30° C for 1 d on YPD (transformation efficiency: $(2.1 \pm 0.3) \times 10^{-4}$). As control *A. tumefaciens* LBA288 cured of the Ti plasmid containing pSET152Y*ODRBoriT* the co-cultivation was performed at 30° C for 1 d on YPD (transformation efficiency: $(2.7 \pm 0.2) \times 10^{-8}$). The error bar is SEM (n = 3). The conjugation efficiency was determined by the total number of donor cells divided by number of transconjugant cells.

Supplemental Table S2. Conjugation experiment using the hybrid-relaxase HD204I and HD274I to mediated DNA transfer from *A. tumefaciens to S. cerevisiae* via the CP_{RP4}/$T4SS_{RP4}$ in the presence of Dtr_{Ti} auxiliary proteins. The used donor strain was *A. tumefaciens* LBA2556 and LBA2210, and the recipient strain *S. cerevisiae* YPH250. Co-cultivation was performed for 72 h at 21° C on IM with acetosyringone (n = 1). The conjugation efficiency was determined by the total number of donor cells divided by number of transconjugant cells. The total amount of colonies that counted per experiment were as follows: for #1: 101; #2: 79; #3: 1; #4: 0; and #5: 1224.

#	Donor strain	CP/T4SS plasmid	Relaxase	Dtr auxiliary proteins to support relaxase	Vector transfer sequence(s)	Conjugation efficiency (3 d)
1	LBA2556	pAL1100Δ*virD2*; RP4	HD204I; $TraI_{RP4}$	Dtr_{Ti}; Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriRB*	7.9×10^{-8}
2	LBA2556	pAL1100Δ*virD2*; RP4	HD274I; $TraI_{RP4}$	Dtr_{Ti}; Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	*oriRB*	8.7×10^{-8}
3	LBA2556	pAL1100Δ*virD2*; RP4	HD204I; $TraI_{RP4}$	Dtr_{Ti}; Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	OD_{Ti}; RB_{Ti}; Δ$oriT_{RP4}$	1.3×10^{-9}
4	LBA2556	pAL1100Δ*virD2*; RP4	HD274I; $TraI_{RP4}$	Dtr_{Ti}; Dtr_{RP4}; $VirC1_{Ti}$; $VirC2_{Ti}$; $VirD1_{Ti}$	OD_{Ti}; RB_{Ti}; Δ$oriT_{RP4}$	0
5	LBA2210	RP4	$TraI_{RP4}$	Dtr_{RP4}	OD_{Ti}; RB_{Ti}; $oriT_{RP4}$	2.3×10^{-5}

Plasmids used per experiment:

#	Donor strain	CP/T4SS plasmid	Dtr expression plasmid	Transfer vector
1	LBA2556	pAL1100Δ*virD2*; RP4	pBBRXCmHD204I	pSET152Y*oriRB*
2	LBA2556	pAL1100Δ*virD2*; RP4	pBBRXCmHD274I	pSET152Y*oriRB*
3	LBA2556	pAL1100Δ*virD2*; RP4	pBBRXCmHD204I	pSET152Y*ODRBΔoriT*
4	LBA2556	pAL1100Δ*virD2*; RP4	pBBRXCmHD274I	pSET152Y*ODRBΔoriT*
5	LBA2210	RP4	n.a.	pSET152Y*ODRBoriT*

n.a.: not applicable

Supplemental Table S3. Sequences of *oriT*$_{RP4}$ and modified *oriT*s (synthesized sequences FTWG1-3). All sequences are depicted in 5'-3' direction. The *oriT*$_{RP4}$ sequence annotation is based on Pansegrau *et al.* (1994), and is depicted underlined with the 13 bp *traJ*$_{RP4}$-binding site (*srj*; Fürste *et al.*, 1989; Pansegrau & Lanka, 1996B) in bold, and the TraI$_{RP4}$-nicking site marked with an asterisk in between G and C (*nic*$_{RP4}$; Pansegrau, Schröder & Lanka 1993), and marked in a grey box the 12 bp deletion sequence containing the TraI$_{RP4}$-nicking for later modifications. The *traK*$_{RP4}$-binding site (*srk*; Ziegelin *et al.*, 1992), situated upstream of the *nic*$_{RP4}$, is up to now poorly defined and therefore not marked in this sequence. The ***FTWG1*** sequence has the modified *oriT*$_{RP4}$ site underlined, the minimal 25 bp *RB*$_{Ti,Oct}$ insert (as present on pTi15955; GenBank NZ_CP032920.1) in bold-italics, the *srj* site in bold, the VirD2$_{Ti}$-nicking site (*nic*$_{Ti}$) marked with an asterisk between T and G, and the inserted *ApaI* site at the 5'-end and the *SphI* site on the 3'-end are marked with a waved underline. ***FTWG2*** sequence has the modified *oriT*$_{RP4}$ site underlined, the *srj* site in bold, the 38 bp *OD*$_{Ti}$ sequence (as present on pTi15955; GenBank NZ_CP032920.1) in a light grey box, the minimal 25 bp *RB*$_{Ti,Oct}$ insert in bold-italics, the *nic*$_{Ti}$ site marked with an asterisk between T and G, and the inserted *ApaI* site at the 5'-end and the *SphI* site on the 3'-end are marked with a waved underline. ***FTWG3*** sequence has the modified *oriT*$_{RP4}$ site underlined, the 38 bp *OD*$_{Ti}$ sequence in a light grey box, the minimal 25 bp *RB*$_{Ti,Oct}$ insert in bold-italics, the *nic*$_{Ti}$ site marked with an asterisk between T and G, and the inserted *ApaI* site at the 5'-end and the *SphI* site on the 3'-end are marked with a waved underline.

***oriT*$_{RP4}$ reference sequence**

TCGGTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCTCTTCTTGATGGAG-
CGCATGGGGACGTGCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGTCAT-
GGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCCTGCTTCTCTTCGATCT-
TCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGGTGAGCCA-
GAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGCCAGCTCGCGGACGTGCT-
CATAGTCCACGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGCCGACAGGCT-
CATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTGGAAGGCAGTACACCTTGA-
TAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCTCATCTGTTACGC-
CGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAGGTGCGAATAAGG-
GACAGTGAAGAAGGAACACCCGCTCGCG**GGTGGGCCTACTT**CACCTATCCTG*CCCGGCTGACGC-
CGTTGGATACACCAAGGAAAGTCTACACGAACCCTTTGGCAAAATCCTGTATATCGTGCGAAAAAG-
GATGGATATACCGAAAAAATCGCTATAATGACCCCGAAGCAGGGTTATGCAGCGGAAAAGCGC

FTWG1 (*oriT*$_{RP4}$Δ*nic*$_{TraI}$*RB*$_{Ti,Oct}$) / Name in plasmids: ***oriRB***

AGGGCCCTGGCCAGCTAGCTAGAGTCGACCTGCAGGTCCCCGGGGGATCG-
GTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCTCTTCTTGATGGAG-
CGCATGGGGACGTGCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGTCAT-
GGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCCTGCTTCTCTTCGATCT-
TCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGGTGAGCCA-
GAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGCCAGCTCGCGGACGTGCT-
CATAGTCCACGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGCCGACAGGCT-
CATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTGGAAGGCAGTACACCTTGA-
TAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCTCATCTGTTACGC-
CGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAGGTGCGAATAAGG-
GACAGTGAAGAAGGAACACCCGCTCGCG**GGTGGGCCTACTT**CA***AATTACAACGGTATATATCCT*GC-CA***GGCTGACGCCGTTGGATACACCAAGGAAAGTCTACACGAACCCTTTGGCAAAATCCTG-
TATATCGTGCGAAAAAGGATGGATATACCGAAAAAATCGCTATAATGACCCCGAAGCAGGGTTATG-
CAGCGGAAAAGATCCGTCGACCTGCAGGCATGCT

Table continued on the next page.

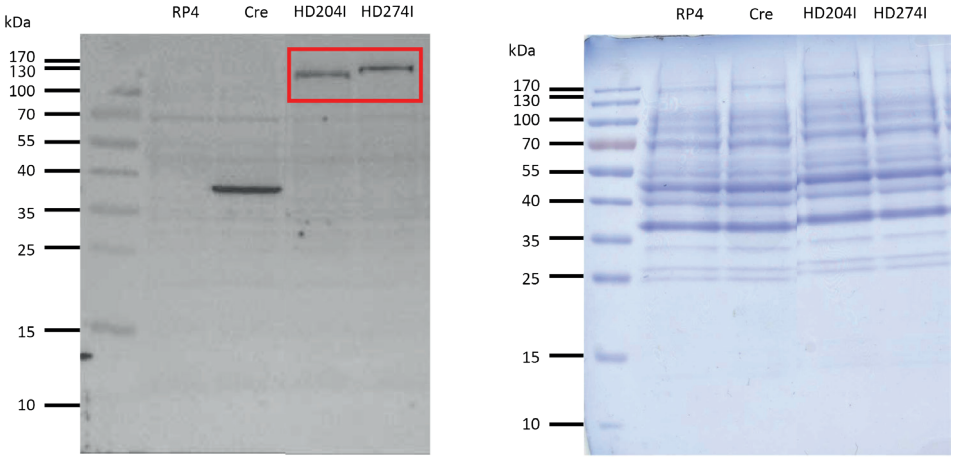
FTWG2 (*oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$) / Name in plasmids: ***oriODRB***

A GGGGCCCTGGCCCAGCTAGCTAGAGTCGACCTGCAGGTCCCCCGGGGGATCG-
GTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCTCTTCTTGATGGAG-
CGCATGGGGACGTGCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGT-
CATGGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCCTGCTTCTCT-
TCGATCTTCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGG-
TGAGCCAGAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGCCAGCTCGCG-
GACGTGCTCATAGTCCACGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGC-
CGACAGGCTCATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTGGAAGGCAG-
TACACCTTGATAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCT-
CATCTGTTACGCCGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAG-
GTGCGAATAAGGGACAGTGAAGAAGGAACACCCGCTCGCG**GGTGGGCCTACTT**CAAACAAA-
CATACACAGCGACTTATTCACACGAGCTCA***AATTACAACGGTATATATCCT*GCCA***GGCTGACGC-
CGTTGGATACACCAAGGAAAGTCTACACGAACCCTTTGGCAAAATCCTGTATATCGTGCGAAAAA-
GGATGGATATACCGAAAAAATCGCTATAATGACCCCGAAGCAGGGTTATGCAGCGGAAAAGATC-
CGTCGACCTGCAGGCATGCA

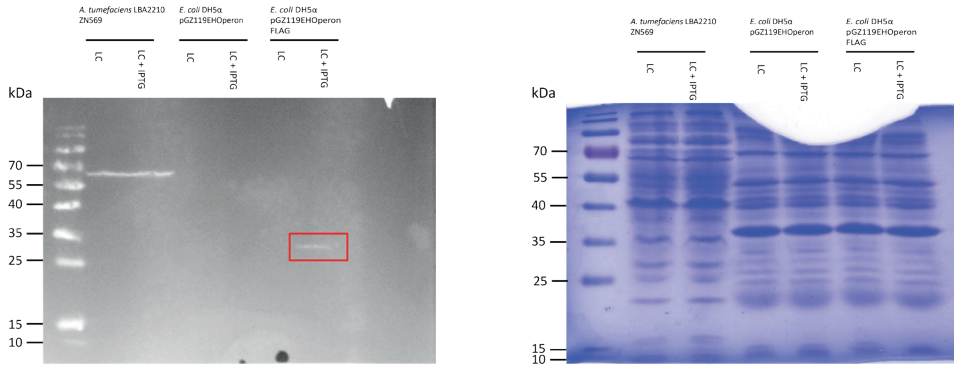
FTWG3 (*oriT*$_{RP4}$Δ*nic*$_{TraI}$*OD*$_{Ti,Oct}$*RB*$_{Ti,Oct}$Δ*srj*) / Name in plasmids: ***oriODRBΔsrj***

A GGGGCCCTGGCCCAGCTAGCTAGAGTCGACCTGCAGGTCCCCCGGGGGATCG-
GTCTTGCCTTGCTCGTCGGTGATGTACTTCACCAGCTCCGCGAAGTCGCTCTTCTTGATGGAG-
CGCATGGGGACGTGCTTGGCAATCACGCGCACCCCCCGGCCGTTTTAGCGGCTAAAAAAGT-
CATGGCTCTGCCCTCGGGCGGACCACGCCCATCATGACCTTGCCAAGCTCGTCCTGCTTCTCT-
TCGATCTTCGCCAGCAGGGCGAGGATCGTGGCATCACCGAACCGCGCCGTGCGCGGGTCGTCGG-
TGAGCCAGAGTTTCAGCAGGCCGCCCAGGCGGCCCAGGTCGCCATTGATGCGGGCCAGCTCGCG-
GACGTGCTCATAGTCCACGACGCCCGTGATTTTGTAGCCCTGGCCGACGGCCAGCAGGTAGGC-
CGACAGGCTCATGCCGGCCGCCGCCGCCTTTTCCTCAATCGCTCTTCGTTCGTCTGGAAGGCAG-
TACACCTTGATAGGTGGGCTGCCCTTCCTGGTTGGCTTGGTTTCATCAGCCATCCGCTTGCCCT-
CATCTGTTACGCCGGCGGTAGCCGGCCAGCCTCGCAGAGCAGGATTCCCGTTGAGCACCGCCAG-
GTGCGAATAAGGGACAGTGAAGAAGGAACACCCGCTCGCGCAAACAAACATACACAGCGACT-
TATTCACACGAGCTCA***AATTACAACGGTATATATCCT*GCCA***GGCTGACGCCGTTGGATACACCAA-
GGAAAGTCTACACGAACCCTTTGGCAAAATCCTGTATATCGTGCGAAAAAGGATGGATATAC-
CGAAAAAATCGCTATAATGACCCCGAAGCAGGGTTATGCAGCGGAAAAGATCCGTCGACCTG-
CAGGCATGCA

oriT$_{RP4}$ sequence based on the 840 bp ApaI-SphI fragment present on plasmid pSET152; the *oriT*$_{RP4}$ sequence on pSET152 was itself based on the 760 bp HaeII fragment of the RP4 plasmid (Guiney and Yakobson, 1983); the minimal 25 bp *RB*$_{Ti,Oct}$ and *OD*$_{Ti,Oct}$ sequences were taken from the octopine Ti plasmid pTi15955 sequence (GenBank NZ_CP032920.1)



Supplemental Figure S4. Western blot protein detection of Cre-fusion hybrid relaxase Cre::HD204I and Cre::HD274I constructs. The constructs used were pCreHD204I (Cre::HD204I) and pCreHD274I (Cre::HD274I). All plasmids were expressed in *E. coli* DH5α. As negative control, the lane RP4 shows protein detection in the presence of only plasmid RP4. As positive only control, lane Cre shows the protein detection of Cre (size of 38 kDA) from plasmid CFPBSm. Moreover, lane HD204I and HD274I shows protein detection of plasmid pCreHD204I and pCreHD274I. Visible and marked in a red rectangle are the predicted proteins of HD204I with around 118 kDA and HD274I with 126 kDA. For analysis after 16 h growth, 1 unit (of OD_{600} of 1) bacteria was used to isolate proteins. The acrylamide gel was composed of an 11% stacking gel and a 12% running gel, the protein ladder was PageRuler Prestained (Thermo Scientific). The left side depicts the western blot and the right side shows the CBB stained gel (both cropped). The gel and western blot were cropped to leave out non-relevant lanes containing constructs from different experiments.



Supplemental Figure S5. Western blot protein detection of FLAG-tagged $T4SS_{Ti}$ Dtr auxiliary proteins. Protein FLAG epitopes were added C-terminally to the auxiliary proteins $VirC1_{Ti}$, $VirC2_{Ti}$ and $VirD2_{Ti}$ encoded on the expression plasmid pGZ119EHOperon. The plasmids pGZ119EHOperon and pGZ119EHOperonFLAG were expressed in *E. coli* DH5α with the addition of 1 mM IPTG in the LC. For analysis after 16 hours growth 1 unit of OD_{600} of 1 bacteria were used to isolate proteins. The acrylamide gel was composed of an 11% stacking gel and a 12% running gel, the protein ladder was PageRuler Prestained (Thermo Scientific). FLAG epitope antibodies were used for this western blot. The left side depicts the western blot and the right side shows the CBB stained gel. Visible and marked with a red rectangle is the FLAG-tagged $VirC1_{Ti}$ with a predicted molecular weight of about 25 kDA (Srinivasen, Yanofsky & Nester, 1989). As a positive control, a construct yielding a 65 kDa FLAG-Cre-6Zincfinger-VirF protein was included on the expression plasmid ZN569 expressed in *A. tumefaciens* LBA2210. All constructs were either expressed in LC medium containing IPGT or just in LC medium.

Supplemental Figure S6. Secondary structure prediction of full-length TraI$_{RP4}$ amino acid sequence. This prediction was used to determine the minimal functional relaxase domain, to be used in the hybrid relaxase experiments. On the left side, the full secondary structure predication is shown with a zoom into the relevant TraI$_{RP4}$ sequence. On the right side, the TraIRP4 FASTA full amino acid sequence is provided. The end of the sequence for TraI$_{RP4}$1-134 AA and TraI$_{RP4}$1-210 AA is marked with a red arrow. Structure letter and color-coding: Red: Helix (H), Green: Sheet (S); Blue: Turn (T); Yellow: Coil (C). The secondary structure was predicted according to the Chou & Fasman algorithm (Chou & Fasman, 1974A; Chou & Fasman, 1974B) on the CFFSP server (Ashok Kumar, 2013).

Chapter 6

Summary

Felix T. Wittleben

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Trans-kingdom DNA transfer – Comparing the Ti and RP4 systems

Horizontal gene transfer (HGT) is a major driving force behind evolution and environmental adaption of microbes (Smillie *et al.*, 2010; Harrison & Brockhurst, 2012). One main category of HGT is the conjugative transfer, or also called bacterial conjugation (Zatyka & Thomas, 1998). Bacterial conjugation allows the transfer of DNA facilitated by a type-IV secretion machinery (T4SS) upon cell-to-cell contact of bacteria. The T4SS, or sometimes also called mating pair formation (Mpf) system, is a complex multi-protein apparatus present in a large variety of bacterial plasmids (Cascales & Christie, 2003). Conjugative DNA transfer via the T4SS is set in motion by a group of proteins responsible for DNA processing forming a Dtr (DNA transfer and replication) system (Schröder & Lanka, 2005). A specific protein forms the interface between the Dtr and the Mpf proteins, and is therefore responsible in determining the substrate to be transferred by the T4SS, called the coupling protein (CP; Schröder & Lanka, 2005). The bacterium *Agrobacterium tumefaciens* is special in that it uses a $T4SS_{Ti}$ to transfer DNA into plant cells, causing crown gall disease (Escobar & Dandekar, 2003). The $T4SS_{Ti}$ is located on a Ti (tumor inducting) plasmid in this bacterium. The conjugative plasmid RP4 (Pansegrau *et al.*, 1994) encodes an evolutionary related $T4SS_{RP4}$, by which it is able to transfer itself to other bacterial hosts. It was shown that in laboratory conditions, RP4 is also able to exhibit *trans*-kingdom transfer from *Escherichia coli* to *yeast* cells (Nishikawa, Suzuki & Yoshida, 1990; Suzuki, Moriguchi & Yamamoto, 2015). This dissertation studied and compared the DNA transfer systems of the Ti and the RP4 plasmid. An in depth view on HGT, bacterial conjugation and *A. tumefaciens*, its Ti plasmid as well as the plasmid RP4 was given in **chapter 1**.

Chapter 2 explored the fate of the $T4SS_{Ti}$ transferred DNA of the Ti plasmid in plant cells. When *A. tumefaciens* transforms plant cells, a specific part of the Ti plasmid called T-DNA, surrounded by two *cis*-acting sequences called border sequences (RB_{Ti} and LB_{Ti}, respectively) is processed in the bacterium. DNA processing is initiated by nicking of the border repeats by the bacterial endonuclease/relaxase VirD2, releasing a single-stranded copy of the T-DNA, or T-strand, which is transferred into the recipient cell's nucleus. In the nucleus the T-strand is converted into a double-strand form by the host endogenous DNA replication machinery (Liang & Tzfira, 2013). For this study a variety of hairpin structures were created adjacent to the border sequences that could fold back on itself in the single-stranded form, after T-strand formation. This fold theoretically could promote double-strand formation, and thereby (transient) expression of the T-DNA genes in the host cells before integration into the genome. A mutation in the left border sequence was discovered that initiated increased double-strand formation and transient expression of the used construct in *Nicotiana benthamiana* leaf cells compared to the non-mutated constructs. This mutation had an inhibiting effect on integration when the studied T-DNA construct was transferred to *Saccharomyces cerevisiae*. This was the first time a modification of a T-DNA border sequence had a profound impact on T-DNA processing, expression and integration.

In **chapter 3**, the ability of *A. tumefaciens* to transfer DNA via its $T4SS_{Ti}$ to the bacterium *Streptomyces lividans* was tested. To test T-DNA on a larger scale a high-throughput screen platform needs to be established. *S. lividans* with its peculiar biology, would have been a good candidate for such a screening platform. An initial study by Kelly & Kado (2002) showed the possibility of such DNA transfer, yet no transfer between *A. tumefaciens* and *S. lividans* mediated by the $T4SS_{Ti}$ could be detected in this study. Neither the original method, nor variations using different media and experimental conditions could show successful transfer. In this case, the originally used strains and constructs could not be used, and therefore a different set of strains and constructs had to be utilized. Nevertheless, in order to see whether DNA transfer between *A. tumefaciens* and *S. lividans* would be possible at all, the $T4SS_{RP4}$-

mediated DNA transfer by the RP4 plasmid was studied. Therefore it could be proven for the first time that $T4SS_{RP4}$-mediated DNA transfer does occur between *A. tumefaciens* and *S. lividans*.

In **chapter 4**, the conjugative DNA transfer of the RP4 plasmid was further examined, and in particular the translocation signal of the RP4 Dtr specific endonuclease/relaxase $TraI_{RP4}$ was identified. A single-strand of DNA bound to Dtr protein is transferred during the conjugative process (Pansegrau, Schröder & Lanka, 1993). This relaxase protein $TraI_{RP4}$ is the key element, that binds and nicks the RP4 plasmid at a specific site, the origin of transfer ($oriT_{RP4}$), after which it remains covalently bound to the RP4 DNA. The strand of RP4 DNA bound by $TraI_{RP4}$ is released and transferred via the $T4SS_{RP4}$ into the recipient cell. This complex only gains entry into the $T4SS_{RP4}$ when recognized by the CP_{RP4}, via a translocation signal (TS) present in $TraI_{RP4}$. In this study, the TS of $TraI_{RP4}$ was identified and found to be localized in the C-terminus. Further, this study found that the $T4SS_{RP4}$ is able to translocate proteins containing the isolated $TraI_{RP4}$ TS without accompanying DNA transfer. This was observed and described for the first time for the $T4SS_{RP4}$.

Chapter 5 looked at the specific interplay between the Dtr components of both the Ti and RP4 transfer systems. Both systems Dtr, CP and T4SS share a strong homology on a protein as well as functional level (Farrand, Hwang & Cook, 1996; Pansegrau & Lanka, 1996; Christie & Gordon, 2014). Other studies showed that in other related systems certain elements e.g. the CP could be functionally exchanged (Cabezón, Lanka & Cruz, 1994; Hamilton *et al.*, 2000). This study looked at the Dtr system, and in particular, investigated the modularity of the relaxases. A hybrid-relaxase containing the DNA processing function of the relaxase $VirD2_{Ti}$ was combined with the TS of $TraI_{RP4}$. The function was tested on a specifically designed DNA substrate to be transferred via the $T4SS_{RP4}$ in combination with specific Dtr_{Ti} support proteins. No DNA transfer could be detected, but it was shown that the hybrid-relaxase itself could be transferred by via the $T4SS_{RP4}$. Another hybrid-relaxase was created containing the DNA processing domain of $TraI_{RP4}$ combined with the TS of $VirD2_{Ti}$ to be transferred via the $T4SS_{Ti}$, but again no DNA transfer was detected. These negative results were discussed. However, this chapter provided a promising outlook and the toolbox for further investigation in the modularity of conjugative systems.

References

Cabezón E, Lanka E & de la Cruz F (1994) Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: trwB and RP4 traG are interchangeable. *J. Bacteriol.* **176:** 4455–4458

Cascales E & Christie PJ (2003) The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* **1:** 137–149

Christie PJ & Gordon JE (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Escobar M & Dandekar A (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends in Plant Science*. **8:** 380-386

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* **20:** 262–267

Kelly B & Kado CI (2002) *Agrobacterium*-mediated T-DNA transfer and integration into the chromosome of *Streptomyces lividans*. *Mol. Plant Pathol.* **3:** 125–134

Liang Z & Tzfira T (2013) *In vivo* formation of double-stranded T-DNA molecules by T-strand priming. *Nat. Commun.* **4:** 1–8

Nishikawa M, Suzuki K & Yoshida K (1990) Structural and functional stability of IncP plasmids during stepwise transmission by trans-kingdom mating: promiscuous conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*. *Japanese J. Genet*. **65:** 323–334

Pansegrau W & Lanka E (1996) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Pansegrau W, Schröder W & Lanka E (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci.* **90:** 2925–2929

Schröder G & Lanka E (2005) The mating pair formation system of conjugative plasmids - a versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54:** 1–25

Smillie C, Garcillán-Barcia MP, Francia M V., Rocha EPC & de la Cruz F (2010) Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74:** 434–452

Suzuki K, Moriguchi K & Yamamoto S (2015) Horizontal DNA transfer from bacteria to eukaryotes and a lesson from experimental transfers. *Res. Microbiol.* **166:** 753–763

Zatyka M & Thomas CM (1998) Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiol. Rev.* **21:** 291–319

Chapter 7

Nederlandse samenvatting

Felix T. Wittleben

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Horizontale DNA overdracht - Een vergelijking tussen het Ti en het RP4 systeem

Horizontale gen-overdracht (HGT) is een belangrijke drijvende kracht achter de evolutie van microben en hun aanpassing aan omgevingsfactoren (Smillie *et al.*, 2010; Harrison & Brockhurst, 2012). Een belangrijke vorm van HGT is conjugatieve overdracht, ook bekend als bacteriële conjugatie (Zatyka & Thomas, 1998). Tijdens bacteriële conjugatie wordt genetisch materiaal overgedragen door middel van een type-IV secretie-systeem (T4SS), ook wel het 'mating pair formation'-systeem (Mpf-systeem) genoemd. Het T4SS bestaat uit een eiwitcomplex dat gecodeerd wordt door een set genen aanwezig op zogenaamde conjugatieve bacteriële plasmiden (Cascales & Christie, 2003). De conjugatieve DNA overdracht wordt geïnitieerd/gestart door het zogeheten 'DNA transfer and replication'-systeem (Dtr-systeem) dat bestaat uit een groep eiwitten die het over te dragen DNA prepareren voor overdracht (Schröder & Lanka, 2005). Een specifiek eiwit, het koppelingseiwit (CP), zorgt voor de koppeling van het Dtr-systeem en de Mpf-eiwitten en bepaalt welke substraten door het T4SS worden getransporteerd (Schröder & Lanka, 2005).

Het $T4SS_{Ti}$ van *Agrobacterium tumefaciens* is bijzonder, omdat het wordt gebruikt om DNA te transporteren naar plantcellen in plaats van naar bacteriën. Die overdracht van DNA veroorzaakt de zogenaamde kroongalziekte in vele plantensoorten (Escobar & Dandekar, 2003). Het $T4SS_{Ti}$ wordt gecodeerd door een Ti (tumor inducerend) plasmide in *A. tumefaciens*. Het conjugatieve plasmide RP4 (Pansegrau *et al.*, 1994) codeert voor een evolutionair vergelijkbaar $T4SS_{RP4}$ waarmee het plasmide zichzelf naar andere bacteriële gastheren kan verplaatsen. Onder specifieke laboratoriumomstandigheden kan RP4 zich ook verplaatsen van *Escherichia coli* naar gist (Nishikawa, Suzuki & Yoshida, 1990; Suzuki, Moriguchi & Yamamoto, 2015). De focus van het onderzoek lag op het bestuderen en het vergelijken van de DNA-overdrachtsystemen van het Ti en het RP4 plasmide. In **hoofdstuk 1** wordt HGT uitvoerig besproken, met de nadruk op de bacteriële conjugatie van het RP4 plasmide en op de DNA overdracht via het Ti plasmide van *A. tumefaciens*.

In **hoofdstuk 2** wordt gekeken naar wat er in plantencellen gebeurt met het DNA dat overgedragen wordt via het $T4SS_{Ti}$. Als *A. tumefaciens* plantencellen transformeert, wordt in de bacterie allereerst een enkele streng van een specifiek gedeelte van het Ti plasmide, geflankeerd door twee *cis*-acterende bordersequenties, los gemaakt van de rest van het plasmide. Dit is mogelijk doordat het endonuclease/relaxase $VirD2_{Ti}$ enkelstrengsbreuken aanbrengt in de bordersequenties. Hoe de enkele streng, de T-streng, vervolgens los komt is nog niet bekend. Na transport van de T-streng naar de kern van de gastheer wordt de T-streng uiteindelijk omgezet naar dubbelstrengs DNA door het in de gastheer aanwezige DNA-replicatieapparaat, maar het is nog onbekend hoe dat precies gebeurt (Liang & Tzfira, 2013). Voor dit proefschrift zijn verschillende DNA sequenties naast de bordersequenties geplaatst zodat er na het vormen van een T-streng een mogelijkheid zou kunnen ontstaan om DNA haarspeldstructuren te vormen doordat het DNA op zichzelf terug zou kunnen vouwen na het vormen van de T-streng. Dit zou *in vivo* kunnen zorgen voor bevordering van de vorming van dubbelstrengs DNA en daarmee van de transiënte expressie van genen op de T-streng voorafgaand aan de integratie in het genoom van de gastheer. Tijdens het onderzoek is een DNA sequentie gevonden naast de linker bordersequentie die zorgt voor een toename in dubbelstrengs DNA vorming en in de transiënte expressie van het construct in *Nicotiana benthamiana*. Deze mutatie heeft daarnaast een inhiberend effect op de integratie van het T-DNA construct wanneer het wordt getransporteerd naar *Saccharomyces cerevisiae*. Dit

was de eerste keer dat een aanpassing van een T-DNA bordersequentie een noemenswaardig effect had op het tot expressie brengen en integreren van het T-DNA.

In **hoofdstuk 3** wordt gekeken naar de overdracht van DNA via het $T4SS_{Ti}$ door *A. tumefaciens* naar *Streptomyces lividans*. Voor het op grotere schaal testen van T-DNA overdracht zou het handig zijn om een testplatform met een hoge doorvoersnelheid op te zetten. *S. lividans* zou hiervoor, gelet op haar specifieke biologie, een goede kandidaat zijn geweest. Een studie door Kelly & Kado (2002) rapporteerde dat DNA-overdracht van *A. tumefaciens* naar *S. lividans* mogelijk was. De in dit hoofdstuk uitgevoerde experimenten waren erop gericht om te bezien of deze overdracht van DNA via het $T4SS_{Ti}$ een geschikt systeem zou zijn om als efficiënt testplatform te kunnen dienen. Bij die experimenten is gebruik gemaakt van zowel (i) de originele methode van Kelly & Kado (2002) als van (ii) de originele methode met aanpassingen in de experimentele condities en het groeimedium. Helaas kon geen gebruik worden gemaakt van de originele stammen en DNA-constructen. In de experimenten is daarom gewerkt met een andere set stammen en constructen. Tijdens de experimenten is geen bewijs gevonden voor de overdracht van DNA door het $T4SS_{Ti}$. Er is daarom vervolgens gekeken of het überhaupt mogelijk is om DNA over te dragen van *A. tumefaciens* naar *S. lividans*. In dat verband is de zogeheten $T4SS_{RP4}$ gemedieerde DNA-overdracht door het RP4 plasmide bestudeerd. Hiermee is voor de eerste keer aangetoond dat $T4SS_{RP4}$ gemedieerde DNA-overdracht tussen *A. tumefaciens* naar *S. lividans* kan plaatsvinden.

In **hoofdstuk 4** is de conjugatieve overdracht van het RP4 plasmide verder onderzocht. Bij de experimenten die in het kader van hoofdstuk 4 zijn verricht, is het translocatiesignaal van het $TraI_{RP4}$ relaxase geïdentificeerd. Tijdens conjugatie wordt een stuk enkelstrengs DNA, waaraan een relaxase covalent is gebonden is, getransporteerd (Pansegrau, Schröder & Lanka, 1993). Het $TraI_{RP4}$ relaxase maakt een enkelstrengsbreuk op een specifieke locatie, 'the origin of transfer' ($oriT_{RP4}$) in RP4, waarbij het covalent verbonden blijft aan het RP4 DNA. Dit DNA-eiwit complex komt vervolgens los van het RP4 plasmide en wordt getransporteerd naar de ontvangende bacterie. Het complex wordt door het CP_{RP4} alleen toegelaten in het $T4SS_{RP4}$ als er een specifiek translocatiesignaal (TS) aanwezig is in het $TraI_{RP4}$-eiwit. In onderhavige studie is aangetoond dat dit signaal zich bevindt in het C-terminale deel van $TraI_{RP4}$, terwijl het relaxase gedeelte zich in het N-terminale deel bevindt. Daarnaast is ook voor het eerst aangetoond dat het $T4SS_{RP4}$ eiwitten voorzien van het $TraI_{RP4}$ TS kan transporteren zonder bijbehorende DNA-overdracht.

In **hoofdstuk 5** is gekeken naar de wisselwerking tussen de Dtr-componenten van zowel het Ti- en als het RP4-overdrachtsysteem. Beide overdrachtssystemen komen wat betreft structuur en functie sterk overeen (Farrand, Hwang & Cook, 1996; Pansegrau & Lanka, 1996; Christie & Gordon, 2014). Specifieke onderdelen, zoals bijvoorbeeld het CP, kunnen soms functioneel worden uitgewisseld (Cabezón, Lanka & Cruz, 1994; Hamilton *et al.*, 2000). In hoofdstuk 5 is gekeken naar het Dtr-systeem en, meer in het bijzonder, naar de modulariteit van de relaxases. Voor dit proefschrift is een hybride-relaxase gemaakt met het relaxase gedeelte van het Ti-relaxase $VirD2_{Ti}$ en het TS van $TraI_{RP4}$. De functionaliteit van deze hybride-relaxase is getest op een artificieel DNA-substraat dat was ontworpen om getransporteerd te worden door het $T4SS_{RP4}$ in combinatie met specifieke Dtr_{Ti}-ondersteuningseiwitten. De overdracht van het artificiële DNA-substraat kon niet worden gedetecteerd, maar de overdracht van het hybride-relaxase zelf door $T4SS_{RP4}$ kon wel worden aangetoond. Met een ander hybride-relaxase, samengesteld uit het relaxase domein van $TraI_{RP4}$ en het TS van $VirD2_{Ti,}$ kon ook

geen DNA-overdracht worden gedetecteerd. De negatieve resultaten worden in dit hoofdstuk bediscussieerd. Daarnaast biedt dit hoofdstuk een veelbelovende blik op de 'toolbox' voor verder onderzoek naar de modulariteit van conjugatieve systemen.

Referenties

Cabezón E, Lanka E & de la Cruz F (1994) Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: trwB and RP4 traG are interchangeable. *J. Bacteriol.* **176:** 4455–4458

Cascales E & Christie PJ (2003) The versatile bacterial type IV secretion systems. *Nat. Rev. Microbiol.* **1:** 137–149

Christie PJ & Gordon JE (2014) The *Agrobacterium* Ti plasmids. *Microbiol. Spectr.* **2:** 295–313

Escobar M & Dandekar A (2003) *Agrobacterium tumefaciens* as an agent of disease. *Trends in Plant Science*. **8:** 380-386

Farrand S, Hwang I & Cook D (1996) The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* **178:** 4233–4247

Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, von Bodman SB, Lanka E, Ream W & Farrand SK (2000) TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. *J. Bacteriol.* **182:** 1541–1548

Harrison E & Brockhurst MA (2012) Plasmid-mediated horizontal gene transfer is a coevolutionary process. *Trends Microbiol.* **20:** 262–267

Kelly B & Kado CI (2002) *Agrobacterium*-mediated T-DNA transfer and integration into the chromosome of *Streptomyces lividans*. *Mol. Plant Pathol.* **3:** 125–134

Liang Z & Tzfira T (2013) *In vivo* formation of double-stranded T-DNA molecules by T-strand priming. *Nat. Commun.* **4:** 1–8

Nishikawa M, Suzuki K & Yoshida K (1990) Structural and functional stability of IncP plasmids during stepwise transmission by trans-kingdom mating: promiscuous conjugation of *Escherichia coli* and *Saccharomyces cerevisiae*. *Japanese J. Genet.* **65:** 323–334

Pansegrau W & Lanka E (1996) Enzymology of DNA transfer by conjugative mechanisms. *Prog. Nucleic Acid Res. Mol. Biol.* **54:** 197–251

Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D, Helinski DR, Schwab H, Stanisich VA & Thomas CM (1994) Complete nucleotide sequence of Birmingham IncP alpha plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239:** 623–663

Pansegrau W, Schröder W & Lanka E (1993) Relaxase (TraI) of IncP alpha plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci.* **90:** 2925–2929

Schröder G & Lanka E (2005) The mating pair formation system of conjugative plasmids - a versatile secretion machinery for transfer of proteins and DNA. *Plasmid* **54:** 1–25

Smillie C, Garcillán-Barcia MP, Francia M V., Rocha EPC & de la Cruz F (2010) Mobility of plasmids. *Microbiol. Mol. Biol. Rev.* **74:** 434–452

Suzuki K, Moriguchi K & Yamamoto S (2015) Horizontal DNA transfer from bacteria to eukaryotes and a lesson from experimental transfers. *Res. Microbiol.* **166:** 753–763

Zatyka M & Thomas CM (1998) Control of genes for conjugative transfer of plasmids and other mobile elements. *FEMS Microbiol. Rev.* **21:** 291–319

Chapter 8

Curriculum vitae

Felix T. Wittleben

Department of Molecular and Developmental Genetics,
Sylvius Laboratory,
Institute of Biology Leiden, Leiden University,
Sylviusweg 72, 2333 BE Leiden, The Netherlands.

Curriculum Vitae

Felix Thomas Wittleben was born on the 20th December 1984 in Schwandorf i.d. Opf. (Germany). He grew up in Schwarzenfeld (Germany) and Regensburg (Germany). From 2001-2002, he attended Hillcrest High School in Greenville (SC, USA). Graduating school from FOS Regensburg in 2004 in Regensburg (Germany), an internship at the Australian Institute of Marine Science (AIMS) in Townsville (Australia) from 2004-2005 followed.

He began to study 'Applied Biology' in September 2005 at the University of Applied Science Bonn-Rhein-Sieg (Bonn, Germany) receiving his BSc diploma in August 2009. Felix wrote his bachelor thesis at the Institute of Virology at University Hospital Bonn (Germany) deciphering the terminal genome sequences of the human Bocavirus (HBoV) under the supervision of Prof. Dr. Oliver E. Schildgen.

In September 2009 he moved to the Netherlands to continue his master studies in the program 'Molecular and Cellular Biosciences' at Leiden University (Leiden, Netherlands), receiving his MSc degree in August 2011. His master thesis was on the subject of identifying cellular interactions of the *Arabidopsis thaliana* transposon *DAYSLEEPER* under the supervision of Dr. Marijn Knip. He started his PhD research in the group of Molecular and Developmental Genetics at the Institute of Biology Leiden (IBL) under the supervision of Dr. Bert J. van der Zaal and Prof. Dr. Paul J.J. Hooykaas. The subject of the PhD thesis was the comparison of horizontal gene transfer machineries encoded in the Ti and RP4 plasmid and the study of trans-kingdom DNA and protein transfer.

From August 2017 to April 2018 he worked as program manager at the Netherlands Centre for Electron Nanoscopy (NeCEN) in Leiden (Netherlands), and since then has worked as scientific project manager and grant advisor at the Leiden Institute of Advanced Computer Science (LIACS) of Leiden University.