



Universiteit  
Leiden  
The Netherlands

## **The effect of the DNA metabolism of the recipient cell on the fate of T-DNA translocated by Agrobacterium**

Rolloos, A.M.

### **Citation**

Rolloos, A. M. (2013, November 14). *The effect of the DNA metabolism of the recipient cell on the fate of T-DNA translocated by Agrobacterium*. Retrieved from <https://hdl.handle.net/1887/22213>

Version: Corrected Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/22213>

**Note:** To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/22213> holds various files of this Leiden University dissertation.

**Author:** Rolloos, Arie Martijn

**Title:** The effect of the DNA metabolism of the recipient cell on the fate of T-DNA translocated by Agrobacterium

**Issue Date:** 2013-11-14

**The effect of the DNA metabolism of the recipient  
cell on the fate of T-DNA translocated by  
*Agrobacterium***

**Arie Martijn Rolloos**

*“Without deviation from the norm, progress is not possible”*

Frank Zappa

Aan mijn familie, vrienden en mijn lieve vrouw en dochter: Xenia en Melina

|             |                                                                                                                             |
|-------------|-----------------------------------------------------------------------------------------------------------------------------|
| Cover:      | Front: Ouroboros, pen drawing by A.M. Rolloos<br>Back: Coffee staining, envisioned by C.V. Henkel, executed by A.M. Rolloos |
| Printed by: | Wöhrmann Print Service, Zutphen, The Netherlands                                                                            |
| ISBN:       | ISBN 978-94-6203-466-2                                                                                                      |

# **The effect of the DNA metabolism of the recipient cell on the fate of T-DNA translocated by *Agrobacterium***

## **Proefschrift**

Ter verkrijging van de graad van Doctor aan de Universiteit Leiden, op gezag  
van de Rector Magnificus Prof. mr. C.J.J.M. Stolker volgens besluit van het  
College voor Promoties te verdedigen op donderdag 14 november 2013 te  
klokke 11.15 uur

door

**Arie Martijn Rolloos**

geboren te Papendrecht, Nederland  
in 1978

## **PROMOTIE COMMISSIE**

|                |                                     |
|----------------|-------------------------------------|
| Promoter:      | Prof. Dr. P.J.J. Hooykaas           |
| Co-promoter:   | Dr. Ir. E.J. van der Zaal           |
| Overige leden: | Dr. A.C. Vergunst                   |
|                | Prof. Dr. C.J. ten Cate             |
|                | Prof. Dr. C.A.M.J.J. van den Hondel |
|                | Prof. Dr. J. Memelink               |

## CONTENTS

|                                                                                                                                                   | Page |
|---------------------------------------------------------------------------------------------------------------------------------------------------|------|
| <b>Chapter 1</b>                                                                                                                                  |      |
| General introduction                                                                                                                              | 7    |
| 1 Introduction                                                                                                                                    | 8    |
| 1.1 <i>Agrobacterium</i> mediated transformation: the preparative phase                                                                           | 9    |
| The <i>Agrobacterium tumefaciens</i> Ti-plasmid                                                                                                   | 9    |
| Adhesion to host cells                                                                                                                            | 9    |
| Induction of virulence genes                                                                                                                      | 10   |
| T-DNA processing                                                                                                                                  | 11   |
| 1.2 <i>Agrobacterium</i> mediated transformation: the intrusive phase                                                                             | 13   |
| Translocation of the T-strand and effector proteins                                                                                               | 13   |
| Nuclear transfer of the T-strand                                                                                                                  | 14   |
| Characteristics of the integration site                                                                                                           | 15   |
| Mechanisms behind the formation of aberrant T-DNA integration                                                                                     | 16   |
| The involvement of host DNA repair mechanisms in T-DNA integration                                                                                | 16   |
| Interactions between <i>Agrobacterium</i> effector proteins and the host factors                                                                  | 17   |
| 1.3 Outline of this thesis                                                                                                                        | 19   |
| 1.4 References                                                                                                                                    | 20   |
| <b>Chapter 2</b>                                                                                                                                  |      |
| Towards T-DNA processing and integration in <i>Saccharomyces cerevisiae</i>                                                                       | 27   |
| <b>Chapter 3</b>                                                                                                                                  |      |
| Involvement of Rad52 in T-DNA circle formation during <i>Agrobacterium tumefaciens</i> mediated transformation of <i>Saccharomyces cerevisiae</i> | 61   |
| <b>Chapter 4</b>                                                                                                                                  |      |
| Enhanced gene targeting by translocated I-SceI during <i>Agrobacterium</i> mediated transformation of yeast                                       | 87   |
| <b>Chapter 5</b>                                                                                                                                  |      |
| General discussion                                                                                                                                | 111  |
| Introduction                                                                                                                                      | 112  |
| Discussion Chapter 2                                                                                                                              | 112  |
| Discussion Chapter 3                                                                                                                              | 115  |
| Discussion Chapter 4                                                                                                                              | 117  |
| Double stranded DNA formation as a limiting factor for transkingdom AMT                                                                           | 119  |
| Future perspectives of the application of genetically modified plants                                                                             | 121  |
| References                                                                                                                                        | 122  |
| <b>Chapter 6</b>                                                                                                                                  |      |
| Dutch summary / Nederlandse samenvatting                                                                                                          | 127  |
| <i>Curriculum vitae</i>                                                                                                                           | 135  |



# Chapter 1

## General introduction

***Agrobacterium tumefaciens* T-DNA processing, transfer and integration in the eukaryotic genome; mechanisms and variables**

## Introduction

*Agrobacterium tumefaciens* is a Gram-negative phytopathogen that is the causal agent of crown gall disease in a broad range of dicotyledonous plants. Crown gall disease is characterized by the formation of tumors at sites where plant tissue has been infected by *Agrobacterium*. This ground-dwelling bacterium has drawn a lot of attention during the past three decades because of its ability to stably integrate a discrete fragment of a large tumor-inducing (Ti) plasmid, into the genome of host cells. The *Agrobacterium* mediated transformation (AMT) of plants has become the technology of choice to introduce or alter genetic traits of plants for plant biotechnologists.

AMT could be put to use to meet some of the needs that are expected to become more prevalent in a future with a growing world population. To give a few examples of potential applications of AMT worth looking forward to: The development of drought resistant plants that can be grown at localities currently unavailable for agriculture, the development of crops offering higher yields of their edible parts or the development of pest resistant plants that could effectively reduce the amount of pesticides that are used now.

A number of research groups have successfully applied AMT to several non-plant organisms ranging from yeast [1, 2], filamentous fungi [3] to sea urchin [4]. *Agrobacterium* apparently is able to adhere sufficiently strong to a broad range of organisms to allow for AMT, provided that the *Agrobacterium* virulence machinery is triggered. AMT of yeast cells has proven to be a very convenient tool to assay the relevance of several host factors for T-DNA integration. Knock-out mutant yeast lines can be easily obtained and transformant-screening after AMT only takes a few days. However, the transformation efficiencies that can be achieved in yeast are relatively low compared to those following AMT of plant cells, especially when regarding the integration of T-DNA without sequence homology to the recipient cells genome. The search for variables that can increase the efficiency of T-DNA integration via the NHEJ pathway during AMT of yeast could provide a deeper insight in the rate limiting steps during AMT.

The process of AMT can be divided into three main parts. Firstly, there is a preparative phase which sets on with the expression of virulence genes, the processing of the transfer-DNA (T-DNA) region located on the Ti plasmid, leading to the formation of a single stranded T-DNA-protein conjugate (the T-strand) and the attachment of *Agrobacterium tumefaciens* to the host cells. Secondly, there is an intrusive phase here the transfer of the T-strand and effector proteins into the host cell via a type IV secretion system (T4SS) takes place and the successive nuclear targeting of the T-strand. The intrusive phase is finalized with the integration of the T-strand in the host cells genome. Thirdly, there is the submissive phase where transcription of the genes located on the integrated T-strand takes place leading to opine synthesis and enhanced cell proliferation. In this introduction the focus will be mostly on the preparative and intrusive phase.

## 1.1 *Agrobacterium* mediated transformation: the preparative phase

### The *Agrobacterium tumefaciens* Ti-plasmid

*Agrobacterium* strains that are able to cause crown gall disease all carry a circa 200 kb Ti plasmid that is required for virulence [5, 6]. On the Ti plasmid, a specific region is present that is transferred and integrated into the genome of host cells: the transfer-DNA (T-DNA) [7, 8]. The T-DNA is flanked by two imperfect direct 25 bp repeats: the T-DNA *left border* (LB) and *right border* (RB) [9]. The genes required for T-DNA transfer are not located on the T-DNA itself but are transcribed from the *vir* region located on the Ti plasmid. The virulence genes located on the *vir* region are involved in T-DNA processing, the T4SS mediated transport of T-strands and also act as accessory proteins in both the *Agrobacterium* and host cells.

The T-DNA contains genes that are involved in the synthesis of opines, a diverse group of amino acid derivatives that *Agrobacterium tumefaciens* utilizes as carbon and nitrogen source [10]. Although these opine synthesis genes are transferred from a bacterial plasmid, they are equipped with eukaryotic expression signals [7]. The *Agrobacterium* Ti plasmids are classified depending on the type of opine synthesis genes they transfer [11]. Among others, there are nopaline and octopine type Ti plasmids [12]. Besides genes for opine biosynthesis, the T-DNA also carries genes involved in the production of the phytohormones auxin and cytokinin which induce the tumor genesis that is characteristic for crown gall disease. The formation of these tumorous outgrowths enable a further increase of the opine production.

### Adhesion to host cells

For pathogenic bacteria, adhesion to host cells is a necessary step for the initiation and maintenance of an infection and is also an important determinant for host specificity. Adherence to the cell surface of the host organism is mediated by a specialized group of proteins, the adhesins. For some human pathogens, the adhesins together with their cognate host receptors are well characterized, for example for the human gastric pathogen *Helicobacter pylori* [13, 14]. The importance of host binding for virulence and host specificity of plant pathogens has also become evident [15-17]. *Agrobacterium* attaches efficiently to both plant cells as well as a variety of abiotic surfaces [18]. Possibly, the aspecific binding to a wide range of surfaces allows *Agrobacterium* to transform so many different cell types.

In Gram-negative bacteria, fimbriae often function as adhesins. However, in many cases it is only a minor subunit at the tip of the fimbriae that is acting as the actual adhesin. Also non-fimbrial adhesins are known to occur in Gram-negative bacteria. *Agrobacterium* strains lacking a Ti plasmid, hence without any T4SS constituents, still confer some capacity to bind tobacco tissue culture cells although strongly reduced [19].

VirB2 is known as the major constituent of the T4SS pilus, while VirB5 is a minor constituent located at its very tip. Both proteins are considered candidates to act as adhesins [17]. VirB2 has been shown to bind to several proteins expressed from an *Arabidopsis thaliana* cDNA library in yeast: BTI1, BTI2, which are currently of unknown function and the membrane associated ATPase AtRAB8. Additionally, *Arabidopsis* with silenced BTI and AtRAB8 expression are less susceptible to transformation by *Agrobacterium* [16].

Although these findings are promising, whether the genes encoded by the Ti plasmid are required for host cell adherence is not clear yet. In a pioneering research a reduced adherence of *Agrobacterium* cells to cultured plant tissue cells was found for *Agrobacterium* cells cured from the Ti plasmid [19] however in another study determinants for adherence were predicted to reside on the chromosomal DNA of *Agrobacterium* [20]. Until now, a direct link between any of the potential *Agrobacterium* adhesins and host cell binding has not yet been established. Performing a reliable binding assay to validate candidate adhesins will be very challenging considering the high rate of aspecific binding that can be expected and the likeliness of redundancy among the *Agrobacterium* proteins that facilitate the attachment to its host cells.

## Induction of virulence genes

*Agrobacterium tumefaciens* carries over 20 virulence genes on the Ti plasmid that are involved in the many steps that ultimately lead to tumor formation. The currently described genes required for T-DNA processing and transfer are all located on the *vir* region of the Ti plasmid, grouped in a varying number of operons. *Agrobacterium tumefaciens* utilizes a two component signal transduction pathway to regulate the expression of the virulence genes. This transduction pathway consists of the response regulator VirG and the transmembrane receptor kinase VirA. VirA is responsive to phenolic compounds like acetosyringone that are released by damaged plant cells (Figure 1A, 1B) [21-23].

There are also a number of sugars like D-(+)-Glucose, D-(+)-Galactose and L-(+)-Arabinose that act synergistically with the phenolic compounds in the induction of *Agrobacterium vir*-genes. The presence of these sugars is sensed by the chromosomally encoded periplasmic protein ChvE that acts in concert with VirA (Figure 1B) [24]. The cytoplasmic domain of VirA is prone to autophosphorylation and also has been shown to phosphorylate the transcriptional activator VirG (Figure 1C) [25, 26]. VirG in turn activates the virulence genes present on the *vir* region (Figure 1D) [27, 28].

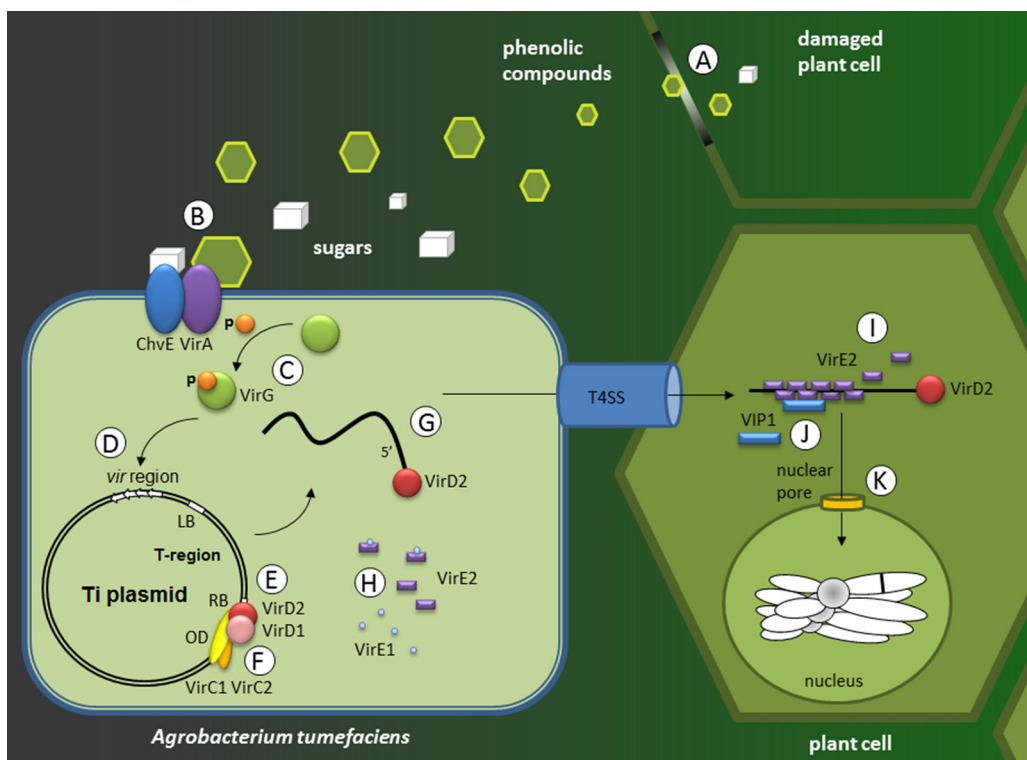


**Figure 1. Schematic overview of *Agrobacterium* mediated transformation (AMT) of plant cells.** (A) Damaged plant cells release phenolic compounds and sugars. (B) Phenolic compounds, like acetosyringone, are recognized by the transmembrane receptor kinase VirA. ChvE senses the presence of sugars. (C) VirA phosphorylates VirG as a response to sugars or phenolic compounds released by plant cells. (D) VirG in turn activates the virulence genes present on the *vir* region of the Ti plasmid. (E) The combined activity of VirD1 and VirD2 induces a nick in the T-DNA RB and LB. (F) VirC1 enhances the assembly of the T-DNA relaxosome by interacting with VirD1, VirD2 and VirC2. (G) A T-strand, is released whereby VirD2 remains covalently attached to the 5'-end. (H) VirE1 competes with the T-strand for binding to VirE2. This enables VirE2 to be translocated separately from the T-strand. (I) In the plant cell the T-strand forms a complex with VirE2 which protects the T-strand from degradation by plant nucleases. (J) VIP1 binds via VirE2 to the T- complex and facilitates nuclear targeting. (K) The T-complex enters the nucleus via the nuclear pores.

## T-DNA processing

T-DNA processing involves *virD2*, *virD1*, *virC2* and *virC1*. VirD2 induces a single strand break (nick) between the 22nd and 23rd base of the border sequences and remains covalently attached via a tyrosine residue to the 5'-end of the T-strand that is eventually released. The transesterification reaction carried out by VirD2 is typical for enzymes belonging to the bacterial relaxase family [29, 30]. Relaxases are involved in the initiation and termination of conjugal transfer of DNA by catalyzing a nicking reaction within their cognate origin of transfer (*oriT*) using supercoiled DNA as substrate and  $Mg^{2+}$  as a cofactor [17]. A typical relaxase has one active site equipped with one or two tyrosyl groups [31].

When the relaxase has formed a complex with the oriT, the free hydroxyl of the active site tyrosine is correctly positioned to perform a nucleophilic attack on the phosphate-deoxyribose backbone of the DNA [18]. The cleaving reaction results in a new phosphodiester bond between the 5'-end of the cleaved DNA strand and the tyrosyl group of the relaxase (Figure 2) [19, 20]. For *Agrobacterium* relaxase VirD2 it has been confirmed by in vitro studies showing that it catalyzes both the site specific nicking reaction and the strand transfer of single stranded DNA (ssDNA) border substrate [22].

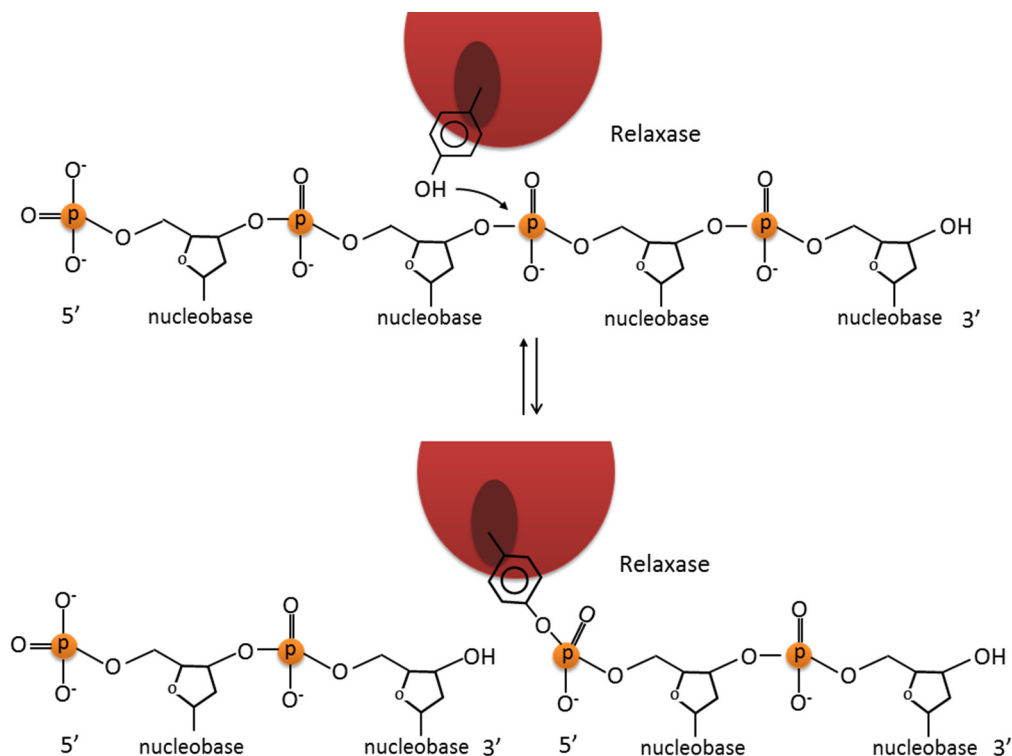


**Figure 2. Schematic of a typical nicking reaction catalyzed by a relaxase on a ssDNA substrate.** The active site tyrosine of the relaxase performs a nucleophilic attack on the phosphate group of the DNA backbone resulting in a phosphodiester bond between the tyrosyl group and the 5'-end of the nicked DNA strand.

*In vitro* experiments showed that the combined activity of VirD2 and VirD1 is required to induce a nick at the RB of double-stranded plasmid DNA (Figure 1E) [32]. These experiments also revealed that purified VirD2 and VirD1 could process a superhelical plasmid DNA substrate but not linearized DNA or plasmid DNA that was brought into a relaxed conformation by treatment with topoisomerase I. Likely, after the first nicking reaction the upper and lower strand of the T-DNA are separated to enable the release of a T-strand. The melting of the dsDNA during T-DNA processing is governed by either displacement synthesis or by helicase activity [33]. Displacement synthesis would involve complementary strand synthesis where a DNA polymerase elongates the 3'-end hydroxyl terminus of the nick, thereby simultaneously separating the T-DNA strands towards the LB sequence.

Displacement synthesis is well described as part of the rolling circle replication of plasmids [34]. Helicases unwind dsDNA using the hydrolysis of nucleoside 5'-triphosphates as an energy source while translocating along the DNA strand. To date, there is no candidate helicase described for *Agrobacterium* that acts on T-DNA after RB nicking. Although many relaxases are known to contain a helicase domain, both VirD2 and VirD1 do not have the A and B motifs of the Walker Box ATP binding site that are conserved among all known members of the helicase families [35].

In the octopine type Ti plasmids, VirC1 binds to *overdrive* (OD) a DNA sequence located upstream of the T-DNA RB [36, 37]. VirC1 enhances the assembly of the T-DNA relaxosome by interacting with VirD1, VirD2 and VirC2 (Figure 1F) and is involved in the polar localization of the T-DNA relaxosome hence priming the T-DNA for secretion by a T4SS (Figure 1G) [38]. Recently, the X-ray crystal structure of the DNA binding domain of VirC2 was elucidated. This domain contained a ribbon-helix-helix fold (RHH fold) that strongly resembles the dimeric RHH folds that are found in bacterial DNA binding proteins like the Arc repressor and CopG. VirC2 however is exceptional within the RHH family because it binds dsDNA with a pseudo-2-fold symmetric RHH fold in a single polypeptide chain instead of forming a symmetric dimer of RHH motifs [39]. It has been suggested that VirC2 facilitates T-DNA processing by VirD1 and VirD2 by destabilization of the border regions [39].

## 1.2 *Agrobacterium* mediated transformation: the intrusive phase

### Translocation of the T-strand and effector proteins

T4SSs are highly complex cell-envelope spanning transport channels that are considered to be ancestrally related to conjugal transfer mechanisms [40]. The T4SS is unique among the prokaryotic secretion systems because it can transfer both DNA and effector proteins into its cognate recipient cells. Effector proteins are proteins produced by prokaryotic pathogens which are transported to their cognate host cells where they interact with host factors in order to modulate their functions thereby facilitating or enhancing the virulence of the pathogen. T4SSs that translocate effector proteins have been the focus of much research because many human pathogens like *Helicobacter pylori* and *Bordetella pertussis* utilize a T4SS to translocate effector proteins to their host cells that have great impact on the disease outcome.

The Ti plasmid of *Agrobacterium* encodes for two structurally similar T4SSs: one for the conjugal transfer of the Ti plasmid itself and an additional T4SS for T-strand transfer to susceptible host cells. The onset of conjugal transfer of Ti plasmids is regulated on both the protein and the genetic level by two environmental signals: on the protein level the inactive monomeric form of the transcriptional activator TraR is dimerized to its active form by interacting with the acyl-homoserine lactone quorum-sensing signal termed *Agrobacterium* autoinducer. When dimeric, TraR activates the *tra* regulon thereby initiating conjugal transfer.

On the genetic level the expression of *traR* is regulated by the presence of a subset of opines that are excreted by the crown gall tumors developed by transformed plant cells [41]. Ti plasmid conjugation thus only occurs in the direct vicinity of plant tumors. The onset of T-strand and effector protein translocation by the second T4SS is differently regulated. Genes encoding the components of the latter T4SS are located within the *vir* region and are acetosyringone-inducible. In the *vir* region, eleven *virB* genes, *virB1-virB11*, together with *virD4*, form the constituents of the T4SS. The genes encoding VirB1-VirB11 are tightly organized in the *virB* operon [42]. VirD4 is encoded within the *virD* operon and is regarded the coupling protein, forming the interface between substrates which are to be transported and the actual pilus itself [43, 44]. The hydrolyzation of ATP by VirD4 and VirB11 ATP-binding subunits triggers VirB10 to undergo a conformational change [45]. This structural change is a prerequisite for DNA substrate passage through the T4SS making the ATPase activity of VirD4 and VirB11 an analogy to the passing of silver coins to Charon, the ferryman of the underworld, in Greek mythology [46].

In addition to the T-strand, with VirD2 covalently attached to the 5'-end, *Agrobacterium* translocates the proteins VirE2, VirE3, VirD5 and VirF to plant cells [47, 48]. Proteins known to be transferred have translocation signals present at or near their C terminus, consisting of clusters of positively charged amino acids which seem to mark these proteins a substrate for the VirB/VirD4 complex [47]. These translocation signals proved to be a portable trait, allowing the *Agrobacterium* mediated translocation of other proteins via the T4SS by making translational fusions with these proteins and *Agrobacterium* translocation signals [47-49].

## Nuclear transfer of the T-strand

In *Arabidopsis thaliana* it has been shown that the importin IMP $\alpha$ -4 is involved in AMT [50]. Importin  $\alpha$  recognizes proteins that are targeted for nuclear import by their NLS and consequently binds to Importin  $\beta$ . Importin  $\beta$  targets the complex to the nuclear pore by binding to nuclear pore proteins (Figure 1K) [51]. After the T-complex is assembled in the plant cytosol, it is generally accepted that VirD2 and VirE2 play a part in the nuclear import of the T-complex. VirD2 has a monopartite nuclear localization signal (NLS) located in its N-terminal part and a bipartite NLS near its C-terminus. Both these NLSs allow nuclear import of reporter genes in plant cells, and also in other eukaryotes like animal and yeast cells [52-56]. However, for the nuclear import of the T-complex only the C-terminal bipartite NLS proved to be indispensable [52, 53]. VirE2 binds to ssDNA without a preference for a specific sequence [57].

It is generally assumed that in the recipient cell VirE2 covers the entire length of the T-strand forming the T-complex although to date no *in vivo* data has been published supporting this. VirE2 associated with the T-strand is thought to offer the enclosed single stranded DNA protection against degradation by plant nucleases (Figure 1I) [58-60]. Despite the affinity of VirE2 for ssDNA it has been shown that VirE2 enters the plant cells independent of the T-strand [61].

This phenomenon was better understood once the crystal structure of the chaperone-like protein VirE1 in complex with VirE2 became available; VirE1 forms a soluble heterodimer with VirE2 in the *Agrobacterium* cells, thus avoiding premature VirE2 binding to the T-strand preventing its translocation (Figure 1H) [62].

Because of this interaction, the nuclear import machinery of the host cell is manipulated to facilitate the nuclear import of the T-complex. VIP1 is a transcription factor that is targeted for nuclear import by the mitogen-activated kinase (MAPK) MPK3 that is part of the defense mechanism by which plants usually fight off pathogens. In that manner, the plant defense pathway is hijacked to facilitate AMT [64]. This strategy for the nuclear targeting of the T-complex does not apply to the AMT of non-plant organisms; VirE2 requires co-expression with VIP1 for optimal nuclear uptake of VirE2 in both yeast cells and mammalian cells, neither of which possess any VIP homologues themselves [65]. Nevertheless, AMT of yeast and mammalian cells is possible without the expression of VirE2 in *Agrobacterium*. It therefore seems that a host factor or another virulence protein can facilitate the nuclear targeting of the T-strand, or alternatively, that the VirD2 moiety of the T-strand already suffices for a basic level of nuclear import. There is experimental evidence suggesting that the VirE3 protein from *Agrobacterium* can mimic the “carrier” function of VIP1, thereby adding to the spectrum of potential host organisms [66].

### Characteristics of the integration site

Several research groups have recovered integrated T-strands from transformed plants to determine if there is a preference for specific genomic loci for T-DNA integration. Several publications described a preference for T-strands to integrate in the transcriptionally active regions of the chromatin [67, 68]. A more recent genome wide analysis of T-DNA integration sites in *Arabidopsis* plants, that were grown under nonselective conditions after AMT, revealed a randomized integration pattern with respect to transcriptionally active or non-active regions of the genome [69]. Nevertheless, when looking at the details, the integration of T-DNA is not always completely random; sequencing of the genomic DNA flanking the integrated T-DNA revealed that there is often a region of micro-homology between the T-strand LB and the integration site [70, 71]. These findings support an early model for single stranded T-DNA integration where the free 3'-end of the incoming T-strand scans the genomic DNA for micro-homologies to anneal to one of the DNA strands of the recipient cell's genome [72]. An additional model for T-DNA integration has been coined more recently that describes the integration process as the integration of a dsDNA molecule, derived from the T-strand, that integrates in double stranded breaks (DSBs) present in the genome of the host cell [73].

## Mechanisms behind the formation of aberrant T-DNA integrations

The integration of T-strand does not always result in the presence of a perfect copy of the original T-DNA region in the recipient cell genome; several categories of aberrant T-strand derived structures were found to be integrated in the recipient cell's genome. These aberrations can offer more insight into some of the peculiarities of AMT and possibly will provide leads for the further improvement of the currently used transformation protocols. Deviant T-strand inserts that are frequently observed are vector backbone fragments. A common cause for the integration of T-DNA containing vector DNA is LB skipping [74-76]. The addition of extra LB repeats [77], the preservation of the natural inner and outer sequence on the binary vector [78] and the usage of a nopaline instead of an octopine type LB sequence have been postulated as ways of increasing the efficiency of left border T-strand termination [78]. Alternatively, the transfer of vector DNA can also result from a combination of VirD2 mediated LB nicking combined with RB skipping [79, 80].

Other aberrant T-strand insertions that are commonly found are T-DNA repeats, present either as inverted or as direct repeats. In plants that were subjected to AMT with two separate *Agrobacterium* strains, more than 50% of the transformed cells contained integrated DNA sequences that can be regarded as fusions of independently transferred T-strand complexes [81]. These data suggest that such T-DNA repeats are forged *in planta* and not in the *Agrobacterium* cells [82]. In the case that a second incoming T-strand successively anneals with its 3'-end to the first T-strand in a similar fashion, it seems plausible that complementary DNA strands are synthesized during repair replication after the T-strands are connected. Another route leading to the formation of complex T-DNA loci was proposed by another group after analyzing several fused T-DNA inserts where the T-strands seemed to be fused together by an illegitimate recombination followed by integration as paired double stranded T-strands [83].

## The involvement of host DNA repair mechanisms in T-DNA integration in plants and yeast

It is tempting to speculate that strand-transferase activity of VirD2 covalently bound to the incoming T-strand plays part in the genomic integration of T-DNA. However, sufficient evidence has been collected that contradicted the direct involvement of VirD2 in T-DNA integration [84-86]. Instead, the DNA repair pathways of the recipient cell involved in the repair of DSBs turned out to be a determining factor in genomic T-DNA integration. DSBs of genomic DNA in eukaryotic organisms are repaired by two conserved major pathways, homologous recombination (HR) and non-homologous end joining (NHEJ). For HR the sister chromatids are used as a template to restore DSBs while during NHEJ the broken ends are directly ligated to each other without a necessity for sequence homology. Although contradicting reports have been published regarding this subject, it is generally assumed that the NHEJ protein Ku80 is involved in

T-DNA integration in plants [87-89]. It is likely that there are several pathways present in plants that facilitate T-DNA integration [90]. T-DNA integration in *Arabidopsis* is also aided by Mre11 that is a part of the MRN complex which among other cellular processes like meiosis and telomere maintenance is part of both the HR and NHEJ DNA repair pathways [91].

For *Saccharomyces cerevisiae*, it was shown that NHEJ components Ku70, Rad50, Mre11, Xrs2, Lig4 and Sir4 are required for the chromosomal integration of non-homologous T-strands [92]. This study also predicted an alternative pathway for T-DNA integration at the telomeres and in an rDNA region in *rad50Δ*, *mre11Δ* and *xrs2Δ* strains associated with drastic chromosomal rearrangements. Strikingly, telomeres and rDNA have in common that, despite not forming hetero chromatin-like structures, these regions are not accessible for RNA polymerase and various recombination proteins [93,94]. It would be interesting to know if the silencing pathway for telomeric DNA and rRNA is involved in this alternative integration mechanism for example by using a *sir2Δ* that is deficient in DNA silencing [95,96]. In an earlier publication, where lithium acetate/single-stranded carrier DNA/PEG transformation [97] was applied instead of AMT to introduce DNA into yeast cells, similar observations are described for the illegitimate genomic integration of dsDNA of a *rad50Δ* yeast strain [98]. Here the deletion of *rad50* led to 47% of the inserts to integrate into repetitive sequences like telomeres, rDNA regions and Ty elements compared to 11% of the inserts analyzed in WT. In addition to this, also chromosomal deletions and translocation events were observed at the integration sites. For the efficient integration of T-DNA homologous to the yeast genome, the HR DNA repair pathway proteins Rad51 and Rad52 are required. *Rad51Δ* and *rad52Δ* yeast strains retained only 9 to 10% of the integration efficiency observed after AMT of WT strains while the deletion of genes specifically involved in NHEJ, *ku70*, *lig4*, *rad50*, *mre11* and *xrs2* had no significant impact on the transformation efficiency [99]. These results taken together showcase a key role for either the NHEJ or HR DNA repair pathway for T-DNA integration in yeast, depending on the absence or presence of regions of homology to the yeast genome on the T-strand.

### Interactions between *Agrobacterium* effector proteins and the host factors

While it is relatively easy to assay T-DNA integrations by determining their final positions in the genome of a recipient cell, it is much more challenging to define those factors within these cells that mediate the integration. In a pioneering study, Mysore and coworkers described a screening where they analyzed an *Arabidopsis* mutant line that is recalcitrant to *Agrobacterium* mediated root transformation but still has the ability to transiently express beta-glucuronidase from the T-strand [100]. This mutant carried a mutation in the core histone protein H2A-1, a member of the 13-gene H2A family in *Arabidopsis*. Interestingly, expression of H2A-1 is induced upon infection with a

transfer competent *Agrobacterium* [101]. These findings make H2A-1 a likely candidate to form the main docking station for the T-complex thereby priming T-DNA integration. However there is probably considerable redundancy in T-complex binding histones in *Arabidopsis* as the overexpression of many of the other core histones leads to an elevation in transformation efficiency [102].

As described above, in *Arabidopsis* the VirE2 protein of the cytosolic T-complex interacts with the transcriptional activator VIP1 which then facilitates nuclear import of the T-complex [63]. VIP1 and VirE2 are good candidates to mediate the association of the T-strand and the host cell's chromatin. Indeed it is reported that VIP1 associates with the chromatin component H2A, probably as a homomultimer [103]. In addition to this, it was shown that the region of VIP1 that enables homomultimerization is required for tumorigenesis but not for the transient transformation of *Arabidopsis* roots [103]. These findings suggest that VIP1 is not only involved in the nuclear targeting of the T-strand but also in the association of the T-complex with the plant chromatin. In line with such a model, it was shown that purified VirE2 is not able to form a complex with nucleosomes isolated from cauliflower florets unless VIP1 acted as a molecular link between VirE2 and the cauliflower nucleosomes [104]. For T-DNA integration to occur it is generally assumed that at a certain point the T-complex must be disassembled by removal of the proteins associated with the T-strand. A recent article reports that synthetic T-complex disassembly can be mediated by Skp1/Cullin/F-box protein VBF pathway present in the recipient plant cells [105]. Earlier reports suggested the involvement of VirF in the disassembly of the T-complex [106]. VirF is dispensable for the AMT of yeast [1] but aids in the AMT of some plant species [107-109].

## 1.3 Outline of this thesis

For almost two decades yeast has been used as a model to study the prerequisites for T-DNA integration and proved to be a very decent working horse. AMT of yeast cells enabled swift transformant screening and offered high convenience when many different mutant backgrounds are required for cocultivation experiments. The AMT of yeast is however not without limitations. AMT of yeast typically yields low transformation efficiencies compared to the AMT of plant cells, especially when NHEJ is required for the integration of the T-strand. Also the variation between experimental replicas is often considerable. **Chapter two** describes the first steps that have been taken to construct an all-in-yeast system enabling T-DNA processing and T-DNA integration in a single yeast cell to address these issues.

A big step forward in plant biotechnology would be the generation of genetically enhanced crops without leaving any genetic marks that are not directly required for the intended plant phenotype. To achieve this, new tools are required that enable the editing of the host cell's genome but without the requirement of T-DNA integration. For this, the expression of transgenes or DNA modifying proteins from circular non-integrative T-DNA derivatives (T-DNA circles) would be an interesting option. **Chapter three** describes a search for the prerequisites of T-DNA circle formation using yeast as a model, focusing mainly on the involvement of the host cell's DNA repair pathways. This chapter provides a model that explains the formation of T-DNA circles and the role of the HR protein rad52 therein.

The natural mechanism behind AMT of plants results in the random integration of T-DNA. More ideally, for plant biotechnological purposes, the T-strand would integrate into a predetermined locus, preferentially via the plant's HR mechanism to ensure highly accurate integration events. In spite of this, in plants T-DNA integration occurs predominantly via the NHEJ DNA repair pathway. **Chapter four** describes the effects of DSB induction on the efficiency of T-DNA integration via the HR pathway in yeast and the role the nucleosome occupancy of the target locus has on DSB induction.

To conclude the research chapters, **Chapter five** will provide an overview of the new insights that can be derived from this thesis and additionally highlight some future prospects.

## 1.4 References

1. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J* 14: 3206-3214.
2. Piers KL, Heath JD, Liang X, Stephens KM, Nester EW (1996) *Agrobacterium tumefaciens*-mediated transformation of yeast. *Proc Natl Acad Sci U S A* 93: 1613-1618.
3. de Groot MJ, Bundock P, Hooykaas PJ, Beijersbergen AG (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat Biotechnol* 16: 839-842.
4. Bulgakov VP, Kiselev KV, Yakovlev KV, Zhuravlev YN, Gontcharov AA, et al. (2006) *Agrobacterium*-mediated transformation of sea urchin embryos. *Biotechnol J* 1: 454-461.
5. Van Larebeke N, Engler G, Holsters M, Van den Elsacker S, Zaenen I, et al. (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* 252: 169-170.
6. Watson B, Currier TC, Gordon MP, Chilton MD, Nester EW (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol* 123: 255-264.
7. Thomashow MF, Nutter R, Postle K, Chilton MD, Blattner FR, et al. (1980) Recombination between higher plant DNA and the Ti plasmid of *Agrobacterium tumefaciens*. *Proc Natl Acad Sci U S A* 77: 6448-6452.
8. Chilton MD, Saiki RK, Yadav N, Gordon MP, Quetier F (1980) T-DNA from *Agrobacterium* Ti plasmid is in the nuclear DNA fraction of crown gall tumor cells. *Proc Natl Acad Sci U S A* 77: 4060-4064.
9. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A* 79: 6322-6326.
10. Zambryski P, Tempe J, Schell J (1989) Transfer and function of T-DNA genes from *agrobacterium* Ti and Ri plasmids in plants. *Cell* 56: 193-201.
11. Guyon P, Chilton MD, Petit A, Tempe J (1980) Agropine in "null-type" crown gall tumors: Evidence for generality of the opine concept. *Proc Natl Acad Sci U S A* 77: 2693-2697.
12. Moore LW, Chilton WS, Canfield ML (1997) Diversity of opines and opine-catabolizing bacteria isolated from naturally occurring crown gall tumors. *Appl Environ Microbiol* 63: 201-207.
13. Ilver D, Arnqvist A, Ogren J, Frick IM, Kersulyte D, et al. (1998) *Helicobacter pylori* adhesin binding fucosylated histo-blood group antigens revealed by retagging. *Science* 279: 373-377.
14. Mahdavi J, Sonden B, Hurtig M, Olfat FO, Forsberg L, et al. (2002) *Helicobacter pylori* SabA adhesin in persistent infection and chronic inflammation. *Science* 297: 573-578.
15. Matthysse AG, Jaeckel P, Jeter C (2008) attG and attC mutations of *Agrobacterium tumefaciens* are dominant negative mutations that block attachment and virulence. *Can J Microbiol* 54: 241-247.
16. Hwang HH, Gelvin SB (2004) Plant proteins that interact with VirB2, the *Agrobacterium tumefaciens* pilin protein, mediate plant transformation. *Plant Cell* 16: 3148-3167.

17. Backert S, Fronzes R, Waksman G (2008) VirB2 and VirB5 proteins: specialized adhesins in bacterial type-IV secretion systems? *Trends Microbiol* 16: 409-413.
18. Tomlinson AD, Fuqua C (2009) Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. *Curr Opin Microbiol* 12: 708-714.
19. Matthyse AG, Wyman PM, Holmes KV (1978) Plasmid-dependent attachment of *Agrobacterium tumefaciens* to plant tissue culture cells. *Infect Immun* 22: 516-522.
20. Whatley MH, Bodwin JS, Lippincott BB, Lippincott JA (1976) Role of *Agrobacterium* cell envelope lipopolysaccharide in infection site attachment. *Infect Immun* 13: 1080-1083.
21. Winans SC (1991) An *Agrobacterium* two-component regulatory system for the detection of chemicals released from plant wounds. *Mol Microbiol* 5: 2345-2350.
22. Stachel SE, Zambryski PC (1986) virA and virG control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell* 46: 325-333.
23. Cangelosi GA, Ankenbauer RG, Nester EW (1990) Sugars induce the *Agrobacterium* virulence genes through a periplasmic binding protein and a transmembrane signal protein. *Proc Natl Acad Sci U S A* 87: 6708-6712.
24. Ankenbauer RG, Nester EW (1990) Sugar-mediated induction of *Agrobacterium tumefaciens* virulence genes: structural specificity and activities of monosaccharides. *J Bacteriol* 172: 6442-6446.
25. Jin S, Roitsch T, Ankenbauer RG, Gordon MP, Nester EW (1990) The VirA protein of *Agrobacterium tumefaciens* is autophosphorylated and is essential for vir gene regulation. *J Bacteriol* 172: 525-530.
26. Huang Y, Morel P, Powell B, Kado CI (1990) VirA, a coregulator of Ti-specified virulence genes, is phosphorylated in vitro. *J Bacteriol* 172: 1142-1144.
27. Jin SG, Roitsch T, Christie PJ, Nester EW (1990) The regulatory VirG protein specifically binds to a cis-acting regulatory sequence involved in transcriptional activation of *Agrobacterium tumefaciens* virulence genes. *J Bacteriol* 172: 531-537.
28. Pazour GJ, Das A (1990) virG, an *Agrobacterium tumefaciens* transcriptional activator, initiates translation at a UUG codon and is a sequence-specific DNA-binding protein. *J Bacteriol* 172: 1241-1249.
29. Durrenberger 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.
30. Jasper F, Koncz C, Schell J, Steinbiss HH (1994) *Agrobacterium* T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proc Natl Acad Sci U S A* 91: 694-698.
31. Garcillan-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.
32. 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.
33. Byrd DR, Matson SW (1997) Nicking by transesterification: the reaction catalysed by a relaxase. *Mol Microbiol* 25: 1011-1022.

34. Khan SA (2005) Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* 53: 126-136.
35. Lohman TM, Bjornson KP (1996) Mechanisms of helicase-catalyzed DNA unwinding. *Annu Rev Biochem* 65: 169-214.
36. Peralta EG, Hellmiss R, Ream W (1986) Overdrive, a T-DNA transmission enhancer on the *A. tumefaciens* tumour-inducing plasmid. *Embo J* 5: 1137-1142.
37. Toro N, Datta A, Carmi OA, Young C, Prusti RK, et al. (1989) The *Agrobacterium tumefaciens* virC1 gene product binds to overdrive, a T-DNA transfer enhancer. *J Bacteriol* 171: 6845-6849.
38. 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.
39. Lu J, den Dulk-Ras A, Hooykaas PJ, Glover JN (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.
40. 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.
41. Oger P, Farrand SK (2002) Two opines control conjugal transfer of an *Agrobacterium* plasmid by regulating expression of separate copies of the quorum-sensing activator gene *traR*. *J Bacteriol* 184: 1121-1131.
42. Ward JE, Akiyoshi DE, Regier D, Datta A, Gordon MP, et al. (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. *J Biol Chem* 263: 5804-5814.
43. 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.
44. Kumar RB, Das A (2002) Polar location and functional domains of the *Agrobacterium tumefaciens* DNA transfer protein VirD4. *Mol Microbiol* 43: 1523-1532.
45. Cascales E, Christie PJ (2004) *Agrobacterium* VirB10, an ATP energy sensor required for type IV secretion. *Proc Natl Acad Sci U S A* 101: 17228-17233.
46. Banta LM, Kerr JE, Cascales E, Giuliano ME, Bailey ME, et al. An *Agrobacterium* VirB10 mutation conferring a type IV secretion system gating defect. *J Bacteriol* 193: 2566-2574.
47. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, et al. (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.
48. Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, et al. (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290: 979-982.
49. Schrammeijer B, den Dulk-Ras A, Vergunst AC, Jurado Jacome E, Hooykaas PJ (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.
50. Bhattacharjee S, Lee IY, Oltmanns H, Cao H, Veena, et al. (2008) IMPa-4, an Arabidopsis importin alpha isoform, is preferentially involved in *agrobacterium*-mediated plant transformation. *Plant Cell* 20: 2661-2680.

51. Bednenko J, Cingolani G, Gerace L (2003) Importin beta contains a COOH-terminal nucleoporin binding region important for nuclear transport. *J Cell Biol* 162: 391-401.
52. Howard EA, Zupan JR, Citovsky V, Zambryski PC (1992) The VirD2 protein of *A. tumefaciens* contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. *Cell* 68: 109-118.
53. Relic B, Andjelkovic 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.
54. Ziemienowicz A, Gorlich D, Lanka E, Hohn B, Rossi L (1999) Import of DNA into mammalian nuclei by proteins originating from a plant pathogenic bacterium. *Proc Natl Acad Sci U S A* 96: 3729-3733.
55. Tinland B, Koukolikova-Nicola Z, Hall MN, Hohn B (1992) The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. *Proc Natl Acad Sci U S A* 89: 7442-7446.
56. Herrera-Estrella A, Van Montagu M, Wang K (1990) A bacterial peptide acting as a plant nuclear targeting signal: the amino-terminal portion of *Agrobacterium* VirD2 protein directs a beta-galactosidase fusion protein into tobacco nuclei. *Proc Natl Acad Sci U S A* 87: 9534-9537.
57. Abu-Arish A, Frenkiel-Krispin D, Fricke T, Tzfira T, Citovsky V, et al. (2004) Three-dimensional reconstruction of *Agrobacterium* VirE2 protein with single-stranded DNA. *J Biol Chem* 279: 25359-25363.
58. Das A (1988) *Agrobacterium tumefaciens* virE operon encodes a single-stranded DNA-binding protein. *Proc Natl Acad Sci U S A* 85: 2909-2913.
59. Citovsky V, Wong ML, Zambryski P (1989) Cooperative interaction of *Agrobacterium* VirE2 protein with single-stranded DNA: implications for the T-DNA transfer process. *Proc Natl Acad Sci U S A* 86: 1193-1197.
60. Christie PJ, Ward JE, Winans SC, Nester EW (1988) The *Agrobacterium tumefaciens* virE2 gene product is a single-stranded-DNA-binding protein that associates with T-DNA. *J Bacteriol* 170: 2659-2667.
61. Lundquist RC, Close TJ, Kado CI (1984) Genetic complementation of *Agrobacterium tumefaciens* Ti plasmid mutants in the virulence region. *Mol Gen Genet* 193: 1-7.
62. Dym O, Albeck S, Unger T, Jacobovitch J, Branzburg A, et al. (2008) Crystal structure of the *Agrobacterium* virulence complex VirE1-VirE2 reveals a flexible protein that can accommodate different partners. *Proc Natl Acad Sci U S A* 105: 11170-11175.
63. Tzfira T, Rhee Y, Chen MH, Kunik T, Citovsky V (2000) Nucleic acid transport in plant-microbe interactions: the molecules that walk through the walls. *Annu Rev Microbiol* 54: 187-219.
64. Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. *Science* 318: 453-456.
65. Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an Arabidopsis protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *Embo J* 20: 3596-3607.
66. Lacroix B, Vaidya M, Tzfira T, Citovsky V (2005) The VirE3 protein of *Agrobacterium* mimics a host cell function required for plant genetic transformation. *Embo J* 24: 428-437.

67. Koncz C, Martini N, Mayerhofer R, Koncz-Kalman Z, Korber H, et al. (1989) High-frequency T-DNA-mediated gene tagging in plants. *Proc Natl Acad Sci U S A* 86: 8467-8471.
68. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, et al. (2003) Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301: 653-657.
69. Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J* 51: 779-791.
70. Thomas CM, Jones JD (2007) Molecular analysis of *Agrobacterium* T-DNA integration in tomato reveals a role for left border sequence homology in most integration events. *Mol Genet Genomics* 278: 411-420.
71. Tzfira T, Li J, Lacroix B, Citovsky V (2004) *Agrobacterium* T-DNA integration: molecules and models. *Trends Genet* 20: 375-383.
72. Tinland B (1996) The integration of T-DNA into plant genomes. *Trends Plant Sci* 1: 178-184.
73. Chilton MD, 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.
74. Wenck A, Czako M, Kanevski I, Marton L (1997) Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation. *Plant Mol Biol* 34: 913-922.
75. Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J* 11: 945-957.
76. Abdal-Aziz SA, Pliego-Alfaro F, Quesada MA, Mercado JA (2006) Evidence of frequent integration of non-T-DNA vector backbone sequences in transgenic strawberry plant. *J Biosci Bioeng* 101: 508-510.
77. Kuraya Y OS, Fukuda M, Hiei Y, Murai N, Hamada K, Ueki J, Imaseki H, Komari T (2004) Suppression of transfer of non-T-DNA 'vector backbone' sequences by multiple left border repeats for transformation of higher plants mediated by *Agrobacterium tumefaciens*. *Mol Breed* 14: 309-320.
78. Podevin N, De Buck S, De Wilde C, Depicker A (2006) Insights into recognition of the T-DNA border repeats as termination sites for T-strand synthesis by *Agrobacterium tumefaciens*. *Transgenic Res* 15: 557-571.
79. Ramanathan V, Veluthambi K (1995) Transfer of non-T-DNA portions of the *Agrobacterium tumefaciens* Ti plasmid pTiA6 from the left terminus of TL-DNA. *Plant Mol Biol* 28: 1149-1154.
80. van der Graaff E, den Dulk-Ras A, Hooykaas PJ (1996) Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants. *Plant Mol Biol* 31: 677-681.
81. De Neve M, De Buck S, Jacobs A, Van Montagu M, Depicker A (1997) T-DNA integration patterns in co-transformed plant cells suggest that T-DNA repeats originate from co-integration of separate T-DNAs. *Plant J* 11: 15-29.
82. Krizkova L, Hroudka M (1998) Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. *Plant J* 16: 673-680.

83. De Buck S, Jacobs A, Van Montagu M, Depicker A (1999) The DNA sequences of T-DNA junctions suggest that complex T-DNA loci are formed by a recombination process resembling T-DNA integration. *Plant J* 20: 295-304.
84. Tinland B, Schoumacher F, Gloeckler V, Bravo-Angel AM, Hohn B (1995) The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *Embo J* 14: 3585-3595.
85. Ziemienowicz A, Tinland B, Bryant J, Gloeckler V, Hohn B (2000) Plant enzymes but not *Agrobacterium* VirD2 mediate T-DNA ligation in vitro. *Mol Cell Biol* 20: 6317-6322.
86. Bravo-Angel AM, Hohn B, Tinland B (1998) The omega sequence of VirD2 is important but not essential for efficient transfer of T-DNA by *Agrobacterium tumefaciens*. *Mol Plant Microbe Interact* 11: 57-63.
87. Friesner J, Britt AB (2003) Ku80- and DNA ligase IV-deficient plants are sensitive to ionizing radiation and defective in T-DNA integration. *Plant J* 34: 427-440.
88. Gallego ME, Bleuward JY, Daoudal-Cotterell S, Jallut N, White CI (2003) Ku80 plays a role in non-homologous recombination but is not required for T-DNA integration in *Arabidopsis*. *Plant J* 35: 557-565.
89. Li J, Vaidya M, White C, Vainstein A, Citovsky V, et al. (2005) Involvement of KU80 in T-DNA integration in plant cells. *Proc Natl Acad Sci U S A* 102: 19231-19236.
90. Qi Jia AdD-R, Sylvia de Pater, Paul J.J. Hooykaas (2011) Ku80 and At5Parp are involved in distinct NHEJ pathways. 93-109 p.
91. Qi Y, Zhang Y, Zhang F, Baller JA, Cleland SC, et al. Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. *Genome Res* 23: 547-554.
92. van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J* 20: 6550-6558.
93. Loo S, Rine J (1995) Silencing and heritable domains of gene expression. *Annu Rev Cell Dev Biol* 11: 519-548.
94. Gottschling DE (1992) Telomere-proximal DNA in *Saccharomyces cerevisiae* is refractory to methyltransferase activity in vivo. *Proc Natl Acad Sci U S A* 89: 4062-4065.
95. Li C, Mueller JE, Bryk M (2006) Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. *Mol Biol Cell* 17: 3848-3859.
96. Sandmeier JJ, Celic I, Boeke JD, Smith JS (2002) Telomeric and rDNA silencing in *Saccharomyces cerevisiae* are dependent on a nuclear NAD(+) salvage pathway. *Genetics* 160: 877-889.
97. Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11: 355-360.
98. Chan CY, Zhu J, Schiestl RH Effect of rad50 mutation on illegitimate recombination in *Saccharomyces cerevisiae*. *Mol Genet Genomics* 285: 471-484.
99. van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31: 826-832.
100. Mysore KS, Nam J, Gelvin SB (2000) An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *Proc Natl Acad Sci U S A* 97: 948-953.

101. Yi H, Mysore KS, Gelvin SB (2002) Expression of the Arabidopsis histone H2A-1 gene correlates with susceptibility to Agrobacterium transformation. *Plant J* 32: 285-298.
102. Gelvin SB, Kim SI (2007) Effect of chromatin upon Agrobacterium T-DNA integration and transgene expression. *Biochim Biophys Acta* 1769: 410-421.
103. Li J, Krichevsky A, Vaidya M, Tzfira T, Citovsky V (2005) Uncoupling of the functions of the Arabidopsis VIP1 protein in transient and stable plant genetic transformation by Agrobacterium. *Proc Natl Acad Sci U S A* 102: 5733-5738.
104. Lacroix B, Loyter A, Citovsky V (2008) Association of the Agrobacterium T-DNA-protein complex with plant nucleosomes. *Proc Natl Acad Sci U S A* 105: 15429-15434.
105. Zaltsman A, Lacroix B, Gafni Y, Citovsky V 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.
106. Magori S, Citovsky V Hijacking of the Host SCF Ubiquitin Ligase Machinery by Plant Pathogens. *Front Plant Sci* 2: 87.
107. Hooykaas PJ, Hofker M, den Dulk-Ras H, Schilperoort RA (1984) A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid, and an Ri plasmid by complementation analysis of Agrobacterium tumefaciens mutants. *Plasmid* 11: 195-205.
108. Schrammeijer B, Risseuw E, Pansegrau W, Regensburg-Tuink TJ, Crosby WL, et al. (2001) Interaction of the virulence protein VirF of Agrobacterium tumefaciens with plant homologs of the yeast Skp1 protein. *Curr Biol* 11: 258-262.
109. Regensburg-Tuink AJ, Hooykaas PJ (1993) Transgenic N. glauca plants expressing bacterial virulence gene virF are converted into hosts for nopaline strains of A. tumefaciens. *Nature* 363: 69-71.

## Chapter 2

### **Towards T-DNA processing and integration in *Saccharomyces cerevisiae***

Martijn Rolloos , Paul J.J. Hooykaas and Bert J. van der Zaal

## Abstract

*Agrobacterium tumefaciens* has the unique ability to transfer genetic material to a wide range of different eukaryotes, ranging from plants to yeast and filamentous fungi. Once translocated to a recipient cell, transfer DNA (T-DNA) lacking its own means to replicate as an episomal piece of DNA, can only be stably maintained after integration into the genome of the recipient. The yeast strain *Saccharomyces cerevisiae* has been developed as a model for the study of T-DNA transfer and integration, but unfortunately such studies are limited by a low transformation frequency. Since many aspects regarding T-DNA integration have not yet been resolved, we wanted to overcome this by developing a system where T-DNA processing would occur in yeast, mediated by expression of *Agrobacterium*-derived proteins. Therefore we introduced a set of *Agrobacterium tumefaciens* virulence genes together with a plasmid carrying an artificial T-DNA (pArT) in this host. To screen for yeast cells which were likely to have experienced T-DNA processing and its subsequent integration at a novel genomic position, we used a promoter-trap assay where a promoterless G418 selection marker was positioned near one of the expected flanks of the processed T-DNA. To maintain control over the onset of the T-DNA processing, the virulence genes were equipped with galactose-inducible promoters. A fraction of the yeast cells did indeed acquire G418 resistance after galactose induction. However, the acquired resistance proved to be transient and could therefore not be attributed to genomic integration of processed T-DNA. Optimization of the experimental setup resulted in a strong reduction of background G418 resistance. Analysis of the resulting yeast strains with acquired G418 resistance uncovered that all acquired resistance could be ascribed to modified versions of pArT.

## Introduction

*Agrobacterium tumefaciens* is a soil-borne Gram-negative phytopathogen that is the causal agent of crown gall disease in a broad range of dicotyledonous plants. Tumor inducing *Agrobacterium tumefaciens* species carry large tumor inducing plasmids (Ti-plasmids) that contain a region of transferable DNA termed T-DNA from which a T-strand is mobilized to enter the host cell and successively integrates as processed T-DNA into the host cells genome. Wild type T-DNA comprises genes involved in the production of phytohormones that are involved in the tumorigenesis that is symptomatic for crown gall disease. In addition to this genes required for the synthesis of opines that *Agrobacterium tumefaciens* utilizes as a source for carbon and nitrogen are present. The part of the Ti-plasmids designated for transfer is defined by two imperfect direct repeats of 25 bp; the *left border* (LB) and *right border* (RB) sequence [1].

The *Agrobacterium* virulence genes required for T-DNA processing and transfer, are located at the *vir* region of the Ti-plasmid and are grouped in a varying number of operons. It has been shown with an *in vitro* study that the relaxase VirD2 catalyzes

a cleaving reaction at the RB sequence on a single stranded DNA (ssDNA) substrate but not on a double stranded DNA (dsDNA) substrate [2]. Another *in vitro* study demonstrated that for VirD2 mediated processing of dsDNA both VirD2 and VirD1 are required [3]. After the nicking reaction has taken place at both the RB and LB sequence, a T-strand with VirD2 covalently attached to the 5' end is released. The T-strand is targeted towards a type IV secretion system (T4SS) that transfers the T-strand and a number of effector proteins to the host cell [4-7]. In the octopine type Ti plasmids, VirC1 binds to *overdrive* (OD) a recognition sequence located upstream the T-DNA RB [8,9]. VirC1 enhances the assembly of the T-DNA relaxosome by interacting with VirD1, VirD2 and VirC2 and is involved in the polar localization of the T-DNA relaxosome facilitating the secretion of the T-DNA by the T4SS [10]. There are indications that VirC2 facilitates T-DNA processing by the destabilization of the border regions, thereby adding leverage to the nicking reaction catalyzed by VirD2 and VirD1 [11].

During AMT, VirE2 enters the plant cells independent of the T-strand and associates with the T-strand forming a T-complex in the recipient plant cell [12]. VirE2 is thought to in this way protect the T-strand from degradation and is involved in the nuclear entry of the T-complex [13-18]. Once the T-complex has formed in the recipient cell, several factors present in the recipient cell are of great importance to accomplish the genomic integration of the T-DNA. The *Arabidopsis thaliana* transcriptional activator VIRE2-INTERACTING PROTEIN 1 (VIP1) is known to facilitate the nuclear import of the T-complex by associating with the T-complex via VirE2 [19]. In addition to this, it has been shown that VIP1 associates with the chromatin component H2A [20]. Considering these two findings it is likely that VIP1 is involved in both the nuclear targeting of the T-complex and the localization of the T-complex to the plant chromatin. The association of the T-complex to the chromatin of the recipient cell was confirmed by the finding that purified VirE2 can only form a complex with nucleosomes isolated from cauliflower florets if VIP1 is also present forming a molecular link between VirE2 and mono nucleosomes [21].

A number of research groups managed to find conditions to broaden the host range of *Agrobacterium* to eukaryotic organisms outside the plant kingdom. The *Agrobacterium* mediated transformation (AMT) of *Saccharomyces cerevisiae* has especially proven its value. Yeast cells have a short generation time allowing for fast transformant screening and mutant backgrounds can easily be obtained. Pioneering studies regarding the effects of the recipient host genotype on T-DNA integration after AMT were performed in this organism. In contrast to plants, T-DNA preferentially integrates at genomic locations sharing sequence homology with the T-DNA by a process involving homologous recombination (HR) via the yeast Rad52 protein [22,23]. When homology is lacking or when the HR pathway is inhibited, integration occurs with low frequency by nonhomologous recombination often co-occurring with small deletions of the target DNA, truncated T-DNA borders and the presence of filler DNA as is typically found with plants after AMT [24]. These integration events are mediated by the enzymes of the NHEJ machinery as defects in this system were shown to prevent NHEJ-mediated

T-DNA integration [22]. Although the AMT of yeast is in principle an attractive model system, establishing the role of particular gene functions within the recipient cell during the process of T-DNA integration is prone to experimental bias due to known and unknown experimental variables during the AMT procedure. Some examples of such variables are the adhesion of *Agrobacterium* cells to the recipient cells and the subsequent translocation of T-strands via the T4SS that takes place during the cocultivation period. The AMT of yeast is characterized by a rather long period of cocultivation, in most protocols between 3 and 5 days [25]. During such an extended period of simultaneous growth of organisms with short generation times, small initial differences in the growth dynamics between experimental replicas can be hugely amplified, adding up to considerable experimental variation when finally assessing the fraction of transformed yeast cells in even the most carefully laid out replicas. In addition to this, differences in the relative input of yeast and *Agrobacterium* cells between experimental replicas have a significant impact on the transformation efficiency (unpublished data).

Considering the issues with the AMT of yeast mentioned above, it is clear that a radically different approach is needed to characterize recipient cell factors that influence T-DNA integration into the genome of the recipient cell in an unbiased manner. Ideally, adhesion and cocultivation should be avoided completely by seeking methods for administering T-strands or T-complexes to the cellular nucleus or generating them there. Since border sequences and virulence (Vir) proteins involved in T-DNA processing and T-complex formation have been characterized and studied for about three decades, we reasoned that it should by now be possible to develop an “all-in-yeast system” by expressing a minimal set of Vir proteins mediating T-DNA processing and the consequent generation of T-strands. In agreement with this idea, *in vivo* T-strand production had already been described in *E. coli* cells; when expression vectors carrying the ORFs of VirD2 and VirD1 were combined with a plasmid based T-DNA substrate, processing of the T-DNA substrate by VirD2 and VirD1 activity resulted in the production of T-strands [26]. A more recent report describes the *in vivo* production and integration of T-DNA in plants cells. Here plant cells were transformed using particle bombardment with VirD1 and VirD2 expressing plasmids together with a plasmid carrying a T-DNA substrate and processed integrated T-DNA was recovered afterwards. In a related publication, chemical transformation of maize protoplasts with a similar combination of plasmids also resulted in the recovery of processed T-DNA [27,28]. In another report, an adenoviral transgene cassette was constructed harboring a T-DNA substrate that was delivered to the nuclei of mammalian cells. Co-delivery of VirD2 and VirD1 ORFs resulted in elevated genomic integration levels of the T-DNA substrate [29]. However, in this case actual processing of the T-DNA substrate by VirD2 and VirD1 remained uncertain.

These promising steps towards controlled T-strand formation in other organisms prompted us to pursue the development of an all-in-yeast system for unbiased characterization of yeast gene functions required for T-DNA integration in the yeast genome. In this study we describe these efforts using galactose-inducible expression cassettes for the production of Vir proteins in yeast combined with a plasmid vector harboring an artificial T-DNA

substrate with a promoterless KanMX resistance marker between border sequences. It was expected that conditions could be found where galactose-induced kanamycin (G418) resistance in yeast cells was indicative of T-DNA integration at a genomic position that could lead to KanMX expression, via a promoter trap mechanism or via a translational fusion to a coding sequences. As described, the systems developed readily suffered from background G418 resistance independent of the expression of *Agrobacterium* virulence proteins. Only the most stringently controlled system led to the detection of T-DNA integration events that could have resulted from T-DNA processing and ectopic integration.

## Results

### Virulence gene expression cassettes

For the development of an all-in-yeast system for the processing of T-DNA and subsequent generation of T-strands, Vir proteins from *Agrobacterium* known to be involved in these processes had to be expressed in yeast cells, preferably in a controllable manner. To maintain control over the onset of T-DNA processing, an expression cassette was constructed that could be loaded with up to three ORFs encoding Vir proteins or other proteins of interest (Figure 1). The promoter sequences used were derived from genes encoding proteins required for galactose metabolism. These promoters are sufficiently divergent to avoid recombination within a single expression cassette and contain well characterized upstream activating sequences (UAS) mediating induction of gene expression if galactose is present but repressing transcription in the presence of glucose [30,31,32,33]. The promoter sequence of the *GAL7* gene (*UAS7*) and the dual promoter driving *GAL1* and *GAL10* expression (*UAS10,1*) were utilized for the construction of two expression cassettes that, when combined, allow the induced expression of up to six proteins. The cognate terminator sequences were used for transcriptional termination. The expression cassettes can be either placed on two separate plasmids or can be integrated in a genomic locus as will be indicated below. The organization of two fully loaded expression cassettes (I and II) is given in Figure 1. The ORFs present in the expression cassettes are mentioned for each experiment described below.

The minimal set of virulence proteins that is required for the processing of T-DNA border sequences consists of VirD2 and VirD1 [3,34]. In addition to these ORFs, also ORFs encoding VirC2 and VirC1 were used for the experiments as these proteins have been reported to enhance border processing by interacting with the OD sequence which is present upstream octopine typed RB sequences [9,10,26,35]. Other ORFs present on the expression cassettes encode VirE2, that, among other functions, is supposedly involved in the protection of the T-strands against nucleases [17,18,36], and the Arabidopsis VIP1 that is reportedly required for the nuclear targeting of VirE2 and the association of the T-complex to the host cell's chromatin [20,37,38]. No sequences encoding tags or nuclear localization signals (NLS) were added to the ORFs to avoid the possibility that such

### Virulence gene expression cassette I



### Virulence gene expression cassette II

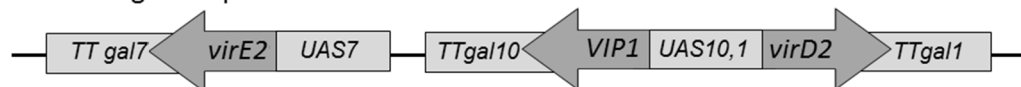


**Figure 1. Organization of the *Agrobacterium* virulence gene expression cassettes I and II.** The *Agrobacterium* virulence ORFs and the *Arabidopsis* VIP1 are regulated by promoter sequences derived from the yeast galactose pathway: the promoter sequence of the *GAL7* gene (*UAS7*) and the dual promoter driving *GAL1* and *GAL10* expression (*UAS10,1*) were utilized for the construction of two expression cassettes that can mediate the expression of up to six ORFs. These promoter sequences induce gene expression if galactose is used as a carbon source in the growth medium but repress transcription whenever glucose is added.

structures could affect protein functionality. Virulence cassettes I and II harboring the ORFs chosen for each experiment, were introduced into yeast cells using two *ARS/CEN* shuttle vectors equipped with non-reverting autotrophy markers respectively: pRS316 and pRS314 [39] for expression cassette II and I, respectively. At a later stage, to avoid the presence of more than one plasmid in yeast cells, expression cassettes I and II were combined in a single integrative vector derived from pINT [40].

### Visualizing expression and nuclear targeting of VirD2 and VirD1 in yeast

To confirm proper inducible expression from different positions within the expression cassette, a GFP-encoding ORF was inserted in frame within either the *virD1* or the *virD2* ORF allowing for visualization of fusion protein expression as well as assessing their cellular localization in yeast. With VirD1 and VirD2 being key virulence proteins, their expression and localization needed to be ascertained. This resulted into the expression vectors pRS314:*virE2::virD2::GFP::virD2* and pRS316:*virD1::GFP::virD1::virC2::virC1*. The galactose induced expression of GFP fused to VirD2 and GFP fused to VirD1 would also provide evidence that the *UAS7* and *UAS1* promoter sequences were active and properly induced. Activity of the *UAS10* promoter, which forms a dual bidirectional promoter with *UAS1*, was confirmed with the expression of VirC2 fused to GFP (data not shown). Yeast cells expressing VirD1 fused to GFP encoded within expression cassette I on pRS316, also exhibited GFP fluorescence but mainly in the cytosol, as is shown in Figure 2 (panels 2A-2C). This is in line with prior expectations since the VirD1 protein does not contain a predicted NLS. Some nuclear localization appeared to be present as well, possibly due to the small size of the GFP fusion protein, allowing some entry via the nuclear pores. In addition to that, a number of bright foci were visible near the nucleus in most of the observed cells. Interestingly, the randomized localization pattern of the VirD1 GFP fusion protein shifted towards a more exclusive nuclear localization in yeast

cells when a pRS314 plasmid was present harboring expression cassette II with an ORF encoding VirD2 (Figure 2, panels 3A-3C). When the above described yeast cells were grown on medium containing glucose instead of galactose no obvious fluorescence was present while using the same microscope settings (data not shown). Nuclear localization of the VirD2 and VirD1 fusion proteins was confirmed with a DAPI staining of the nuclei of yeast cells harboring the fusion genes described above (Figure 3).

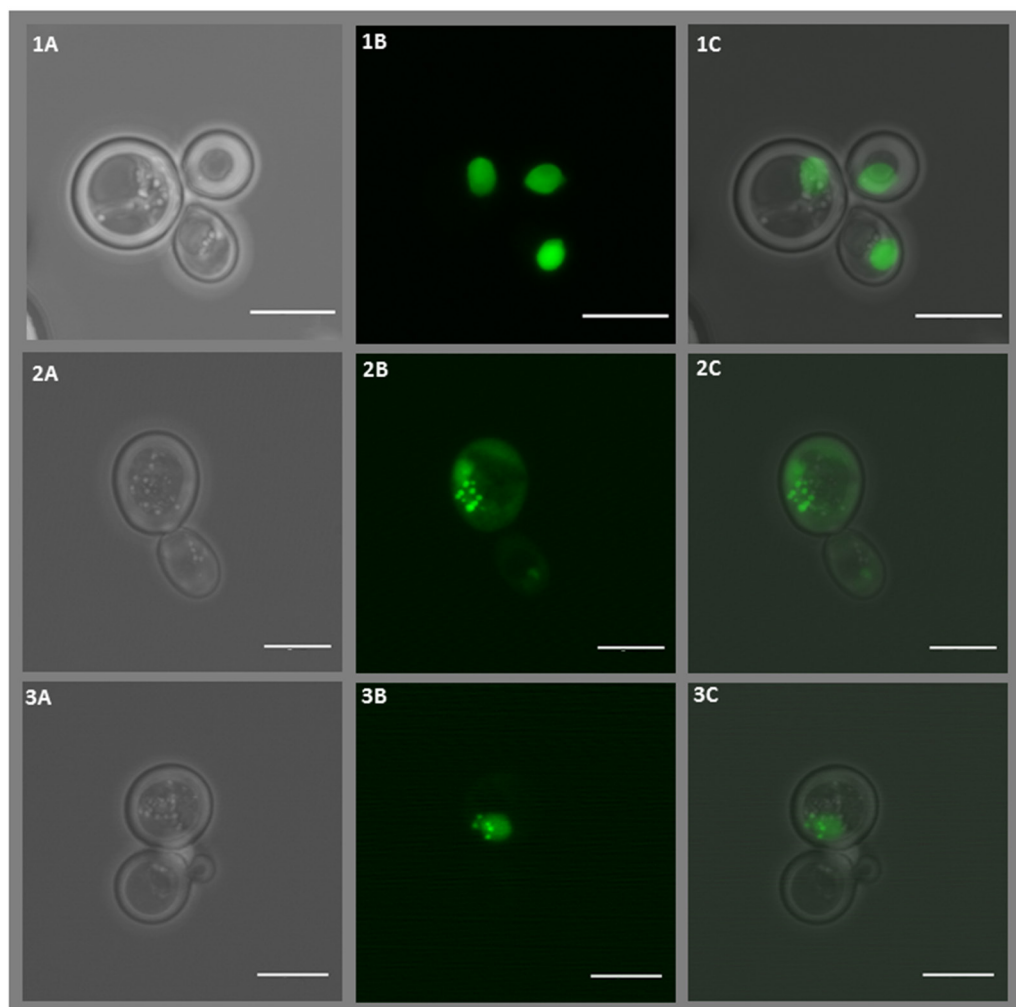


**Figure 2. Cellular localization of VirD2 and VirD1 fused to GFP. (1A) Bright-field image of cells expressing VirD2 fused to GFP. (1B) GFP signal of cells expressing VirD2 fused to GFP (1C) Superimposition of a bright-field image and the GFP signal of VirD2 fused to GFP. (2A) Bright-field image of cells expressing VirD1 fused to GFP. (2B) GFP signal of cells expressing VirD1 fused to GFP (2C) Superimposition of a bright-field image and the GFP signal of VirD1 fused to GFP. (3A) Bright-field image of cells expressing VirD1 fused to GFP and VirD2. (3B) GFP signal of cells expressing VirD1 fused to GFP and VirD2. (3C) Superimposition of a bright-field image and the GFP signal of cells expressing both VirD1 fused to GFP and VirD2. Each white scale bar represents a 5  $\mu$ m distance.**

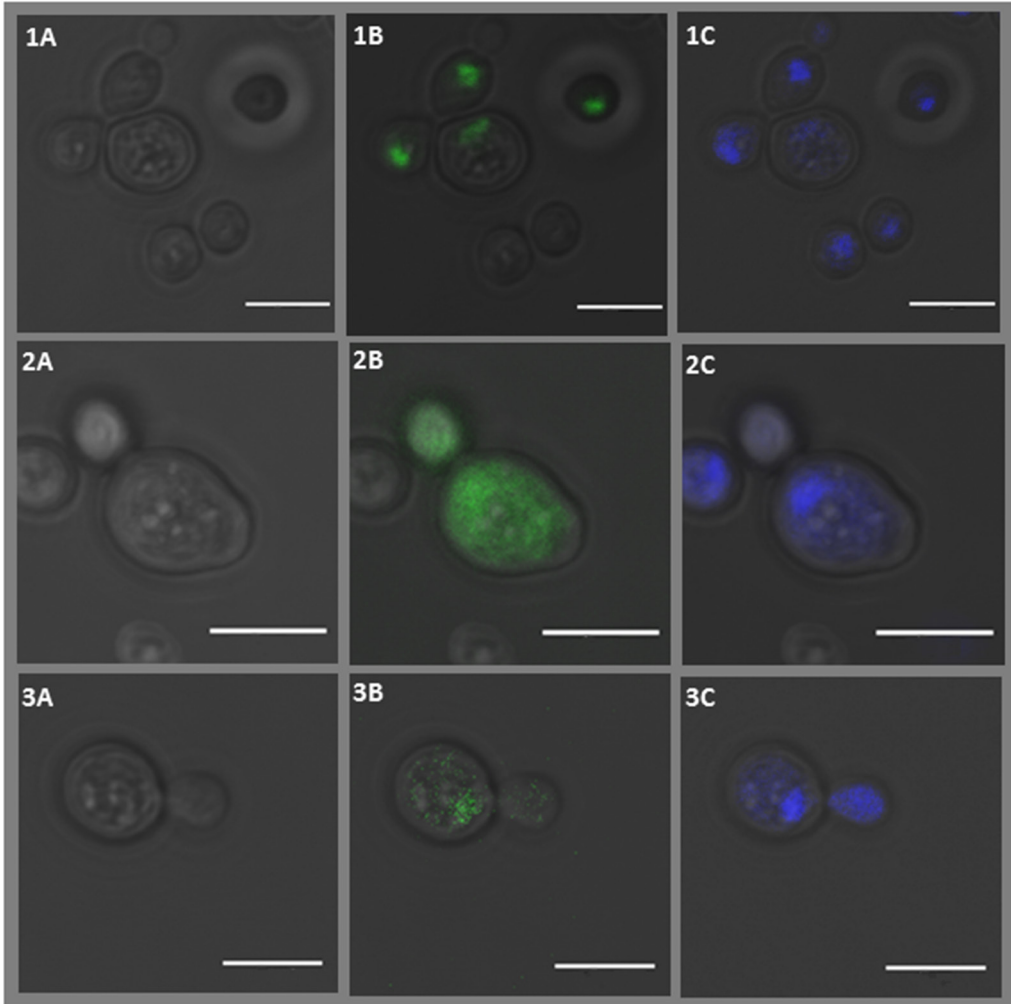


**Figure 3. VirD2 facilitated nuclear localization of VirD1 GFP fusion proteins.** (1A) Bright-field image of cells expressing VirD2 fused to GFP. (1B) Superimposition of a bright-field image and the GFP signal of VirD2 fused to GFP (1C) Superimposition of a bright-field image and the DAPI signal of cells expressing VirD2 fused to GFP (2A) Bright-field image of cells expressing VirD1 fused to GFP. (2B) Superimposition of a bright-field image and the GFP signal of VirD1 fused to GFP (2C) Superimposition of a bright-field image and the DAPI signal of cells expressing VirD1 fused to GFP. (3A) Bright-field image of cells expressing VirD1 fused to GFP and VirD2. (3B) GFP signal of cells expressing VirD1 fused to GFP and VirD2. (3C) Superimposition of a bright-field image and the DAPI signal of cells expressing both VirD1 fused to GFP and VirD2. Each white scale bar represents a 5  $\mu\text{m}$  distance.

### Substrate plasmids for T-strand formation

To provide the all-in-yeast system with a substrate for border processing and T-strand formation, an artificial T-DNA was cloned into the yeast *ARS/CEN* vector pRS313 resulting in the plasmid pArT1 (Figure 4A). In order to be consistent with the source of

the *vir* genes used in the expression cassettes, elements for the artificial T-DNA were also obtained from the same octopine Ti-plasmid pTi-15955 (NCBI reference sequence: NC\_002377.1) [43]. Octopine Ti-plasmids harbor two adjacent T-regions, T-left (TL) and T-right (TR) [44]. The border sequences that were used for the construction of pArT1 were minimal length TL border sequences, just consisting of the 25 nucleotide border repeats preceded by the immediate 40 upstream nucleotides. In that manner, the TL RB sequence also contained the OD sequence upstream of the RB border repeat, the binding site for VirC1 [8,9]. In between the LB and ODRB sequences the pArT1 plasmid harbors a promoterless KanMX marker gene. In pArT1, the actual KanMX coding sequence starts with its own ATG initiation codon, but is preceded by an ORF that starts upstream of the RB sequence. The KanMX marker gene used for the construction of pArT1 is known to tolerate N-terminal protein fusions [45]. When the system would be operational, VirD2 and VirD1 mediated processing of the T-DNA section of pArT1 should lead to production of a T-strand carrying the KanMX marker gene (Figure 4B). The consecutive integration of the KanMX marker gene should lead to G418 resistance when integrating directly downstream of an active yeast promoter or when integration takes place in a yeast ORF leads to the generation of a translational fusion protein (Figure 4C).

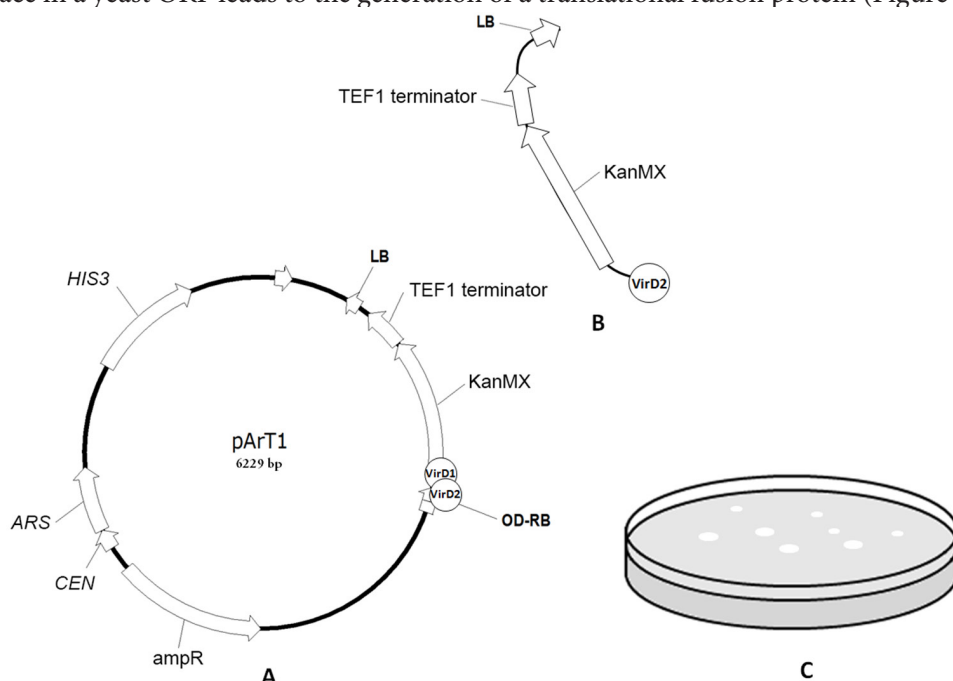


**Figure 4. Strategy for the detection of T-DNA integration events in yeast.** (A) The yeast shuttle vector pRS313 with an ampicillin resistance gene (*ampR*) for selection in *E. coli* and a yeast *HIS3* auxotrophy marker and chromosomal *ARS/CEN* replication origin has been equipped with an *Agrobacterium leftborder* (LB) and *overdrive -rightborder* (OD-RB) sequence. In between the border sequences a promoterless KanMX marker gene is present. (B) The combined catalytic activity of VirD1 and VirD2 results into LB and RB processing and the mobilization of a T-strand with VirD2 covalently attached to the 5'-end. (C) The integration of the mobilized T-DNA just downstream of a yeast promoter or within a genomic yeast ORF restores the expression of the KanMX marker gene leading to G418 resistance.

The envisioned strategy requires that yeast cells with an unmodified pArT1 vector are sensitive to G418. To assess the sensitivity, 50  $\mu$ l of a yeast cell suspension with an OD<sup>600</sup> of  $10^{-3}$  was plated on YPD medium with G418 concentrations ranging from 2  $\mu$ g/ml to 150  $\mu$ g/ml (Figure 5). As a positive control for G418 resistance, a plasmid designated pArT1prom was used. This vector is identical to pArT1, except that it harbors a TEF1 promoter upstream the KanMX ORF. According to the assay, all yeast cells harboring pArT1 were killed when the G418 concentration exceeded 50  $\mu$ g/ml. The yeast cells harboring pArT1prom still divided and formed readily developing colonies even when challenged with a G418 concentration of 150  $\mu$ g/ml. For all the experiments described below, G418 was applied at a concentration of 200  $\mu$ g/ml, assuming that this concentration would suffice for a reliable selection of yeast cells with Vir protein mediated T-DNA translocation events.

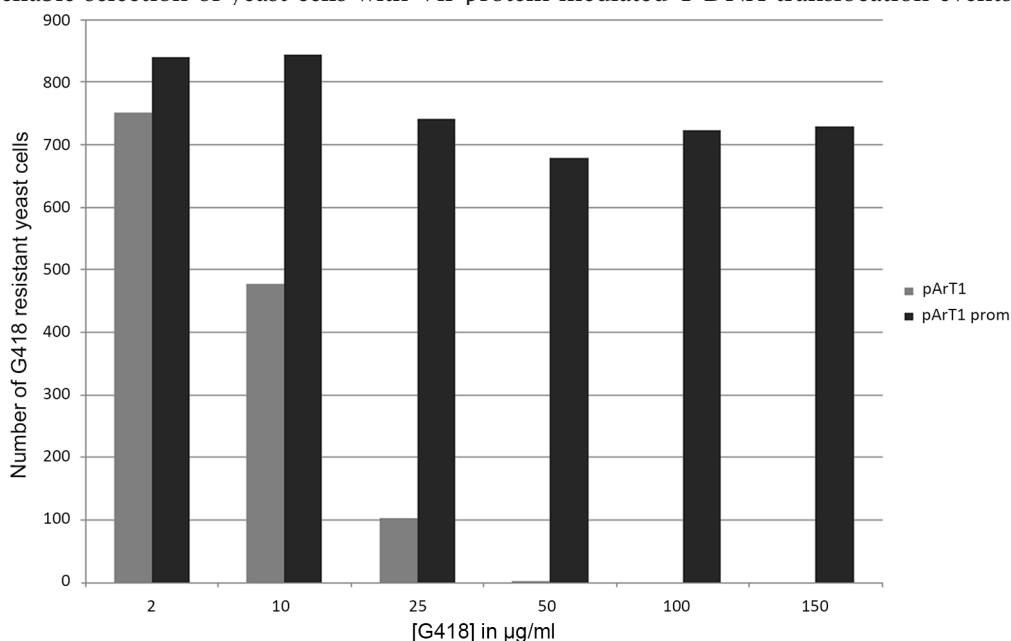


**Figure 5. G418 resistance conferred by pArT1 and pArT1 prom in the yeast strain YPH250.** The yeast strain YPH250 harboring pArT1 and pArT1 prom were assayed on G418 sensitivity. For this, 50  $\mu$ l of a yeast cell suspension with an OD<sup>600 nm</sup> adjusted to  $10^{-3}$  was plated on plates with YPD medium mixed with a dilution series of G418.

### Validation of the compatibility of the border sequences of pArT1 to VirD2 mediated border processing

To confirm that the minimal LB and ODRB sequences used for the construction of pArT1 are compatible with *Agrobacterium* virulence gene mediated border processing, we wanted to demonstrate that these borders were functional during AMT of yeast. Therefore, a plasmid with these borders (pArT1 stop, which will be described below) was equipped with the broad host range replication origin pVS1 [46,47] so that it could be maintained in *Agrobacterium* as a donor of T-strands as a binary plasmid.

To be able to select transformed yeast cells, the promoterless KanMX marker gene was replaced with a restored KanMX marker gene equipped with a TEF1 promoter and terminator sequence. The yeast strain YPH250 was subsequently cocultivated with *Agrobacterium* strain LBA1100 carrying the disarmed octopine Ti-plasmid pTiB6 [48] and the newly developed pArT1 derived binary plasmid, which resulted in high levels of AMT only when the *Agrobacterium* cells were induced with acetosyringone (Figure 6). These data demonstrated that the minimal border sequences used for the pArT1 vector in principle form adequate substrates for Vir protein mediated processing, by comparison about as efficient as the more extended octopine type border sequences which are described in Chapter 3 of this thesis.

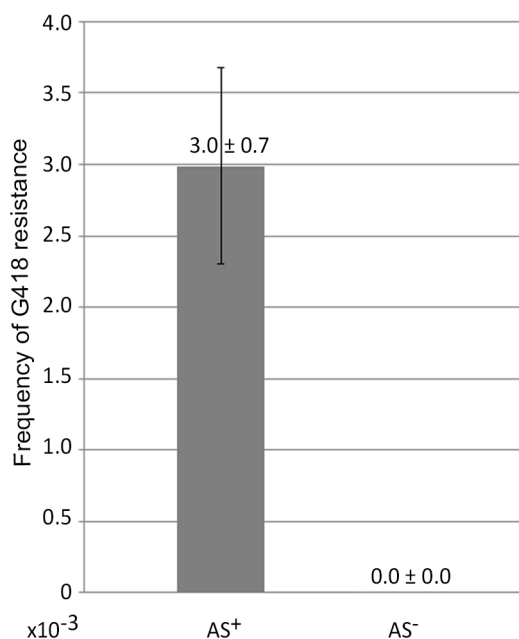


**Figure 6. *Agrobacterium* virulence protein mediated processing of the pArT1 minimal border sequences.**

Cocultivation of the yeast strain YPH250 with *Agrobacterium* strain LBA1100 harboring a pArT1 derived vector with an active KanMX resistance marker and a pVS1 origin of replication allowing for plasmid replication in *Agrobacterium*. These cocultivation experiments were performed both with and without acetosyringone induction indicated here as AS<sup>+</sup> and AS<sup>-</sup>. The averages of five independent experiments are shown. Error bars represent the SEM. The transformation frequencies used to determine these averages were calculated by dividing the total number of G418 resistant colonies by an estimate of the total number of yeast cells

### All-in-yeast T-DNA integration assay in liquid medium

For the first series of all-in-yeast T-DNA integration assays, pArT1 was used to serve as substrate for T-DNA processing. This ARS/*CEN* type vector was introduced in yeast strain YPH250 together with the related yeast shuttle vectors pRS314 and pRS316 from which several combinations of virulence genes were expressed. To assess whether expression of Vir proteins could mediate integration of processed T-DNA to a new position in the yeast genome, all yeast strains were provided with the T-DNA donor pArT1 as well as pRS314 and pRS316, with or without differently loaded Vir protein expression cassettes. These yeast strains were grown overnight in a non-selective rich medium containing either 2% glucose or 2% galactose. After overnight growth, samples from the different yeast cell cultures were plated on G418 containing selective medium to assess whether processed T-DNA had integrated at other positions in the yeast genome. An overview of the yeast strains used and the observed frequency of G418 resistance is given in Figure 7.

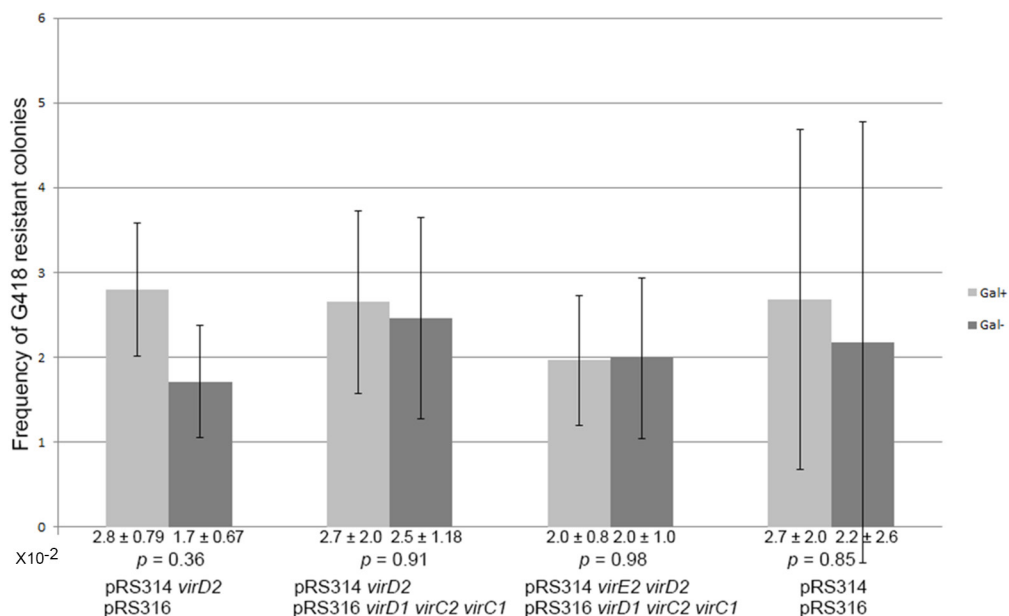


**Figure 7. All-in-yeast promoter trap-assay to detect integration of processed T-DNA.**

All yeast strains were provided with pArT1 combined with both pRS314 and pRS316, with or without differently loaded Vir protein expression cassettes. As negative control pRS314 and pRS316, both without expression cassettes, were combined with pArT1. The expression of virulence proteins was induced by the presence of galactose and repressed in the presence of glucose indicated here as Gal+ and Gal- respectively. The percentage of transformed yeast cells was calculated by dividing the number of G418 resistant colonies by an estimate of the total number of yeast cells. The averages of three independent experiments are shown. Error bars represent the SEM. A two-tailed student's T-test revealed no significant difference between samples grown on galactose or glucose for each of the vector combinations.

Strikingly, the percentage of cells that acquired G418 resistance was about 2% for all plasmid combinations present and did also not differ significantly between growth conditions inducing virulence gene expression or growth conditions repressing virulence gene expression. This observation was confirmed by a two-tailed student's T-test assuming unequal variance (heteroscedastic T-test). In addition to this, pArT1 combined with the empty vector controls showed a frequency of G418 resistance comparable to the yeast cells that could express one or more virulence genes. The observed background levels necessitated an optimization of the signal-to-noise ratio of the all-in-yeast system for T-DNA processing and integration.

To investigate whether proper induction or repression of at least the VirD2 protein had occurred, a sample of yeast cells that were grown overnight for the T-DNA integration assay, depicted in Figure 7, was used for Western blot analysis with a VirD2 antibody (Figure 5). The protein samples B, C and D represent those vector combinations that would allow for the production of VirD2. Only if the yeast cells were cultured on growth medium with galactose, a 55 kDa band likely to represent VirD2 appeared on the Western blot.

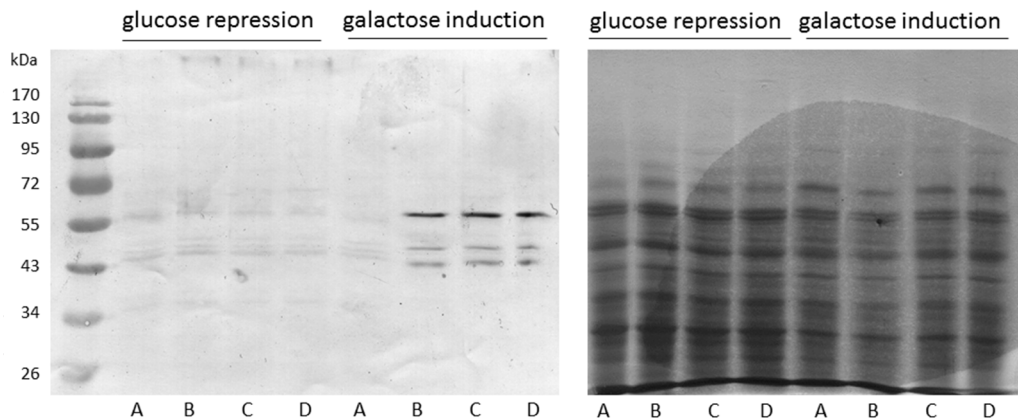


**Figure 8. Expression levels of VirD2 in samples derived from an all-in-yeast T-DNA integration assay.** Yeast cells with varying vector content here indicated as A-D, were cultured overnight on a non-selective growth medium containing either glucose or galactose as carbon source. Protein samples were isolated from yeast strains harboring the following plasmids: (A) pArT1, pRS314 and pRS316, (B) pArT1, pRS314:*virD2* and pRS316, (C) pArT1, pRS314:*virD2* and pRS316:*virD1 virC2 virC1*, (D) pArT1, pRS314:*virE2 virD2* and pRS316:*virD1 virC2 virC1*. Protein samples B, C and D were isolated from yeast cells containing a galactose-inducible VirD2 expression vector. Protein samples A represent the empty vector control. The left panel shows a western blot after immunostaining with a VirD2 antibody. The right panel depicts the corresponding Coomassie brilliant blue staining showing no obvious fluctuations in total protein content between the collected protein samples.

### Adaptions of the all-in-yeast system to reduce background G418 resistance

Yeast cells harboring pArT1 were already fully sensitive to 50 µg/ml G418 (Figure 5). As mentioned above, for our selection procedure the G418 concentration was even raised to 200 µg/ml. The relatively high frequency of G418 resistance observed in all yeast strains were pArT1 was combined with pRS314 and pRS316 expressing several combinations of virulence genes (Figure 7) thus required further explanation. Since the acquired resistance to G418 did not correlate with the expression of the virulence protein VirD2, which is crucial for T-DNA processing, insertion of processed T-DNA could not be the cause for the G418 resistance. Moreover, the observed G418 resistance of yeast cells was quickly lost when such G418 resistant cells were further grown on rich non-selective medium. With Southern blot analysis, no evidence for integration events was found, but only plasmid-derived signals. All available data indicated that there was a plasmid-based source for the acquired G418 resistance; it seemed as if the DNA sequence upstream the KanMX selection marker acquired significant promoter activity in a rather large fraction of cells, thereby obscuring any acquired G418 resistance that might had resulted from T-DNA integration events.

To counter the phenomenon of virulence gene independent acquired G418 resistance from pArT1, a stop codon-rich sequence (stopper) was developed, containing stop codons in all possible reading frames, and cloned directly upstream of the RBOD sequence. This adjusted pArT1 vector was renamed to pArT1stop.

The stopper sequence was introduced directly upstream the ODRB sequence to inhibit any expression of the KanMX marker gene present on pArT1stop without interfering with the expression of the KanMX marker gene after genomic integration as a T-strand (Fig 9).

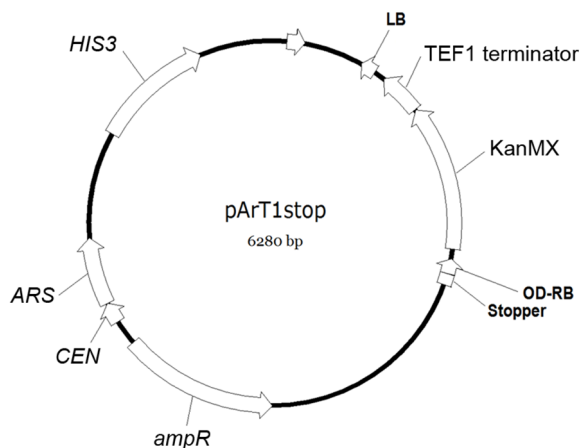


**Figure 9. Vector map of pArT1stop.** A stop codon-rich DNA fragment termed “stopper sequence” was introduced in pArT1 directly upstream the ODRB sequence to counter the expression of the KanMX marker gene. Since border processing takes place downstream of the stopper sequence, any integrated T-strands derived from pArT1stop are unaffected by the stopper sequence.

In an experiment where the pArT1stop vector was combined with pRS314 and pRS316 as empty vector controls strongly reduced background levels of G418 resistance were observed ( $0.8 \times 10^{-5}$  with a SEM. of  $3.7 \times 10^{-5}$  for  $n=3$ ). In the same experiment pArT1stop combined with pRS314 expressing VirD2, VirD1, VirC2 and VirC1 showed much higher levels of G418 resistance of  $25 \times 10^{-3}$  with a SEM. of  $5.7 \times 10^{-3}$  for  $n = 3$  however a Southern blot analysis revealed only the presence of vector DNA and no indications that any integration events had occurred (data not shown).

Analysis of the yeast strain used that gave the latter results revealed that the sequence upstream the ODRB of pArT1stop had reverted to a sequence identical to pArT1. This loss of the stopper sequence might be explained by homologous recombination between the extensive region of homology upstream the stopper sequence and the six base pairs of sequence homology downstream the stopper sequence between pArT1stop and the other members of the pRS vector series that were present in that cell.

Although it has been reported that up to four ARS/CEN containing plasmids of the pRS vector series can be maintained in the yeast strain YPH250 [49], this was in fact merely based upon the ability of the transformed yeast cells to maintain the autotrophic characteristics mediated by the introduced vectors. Considering the stochastic behavior of the KanMX marker gene, one could consider the possibility that reshuffling of plasmid sequences via host dependent recombination processes can occur with concomitant

changes in expression of plasmid-borne traits. Hence, to assay if the background resistance previously observed with pArT1 was exacerbated by the addition of one or two plasmids with high sequence homology, the percentage of G418 resistant yeast cells was compared between yeast cells containing pArT1, pArT1 together with pRS314, and pArT1 together with pRS314 and pRS316. As can be seen in Figure 10, under the conditions used the presence of pArT1 led to about 0.1% G418 resistance. However, around 5% of the plated cells formed G418 resistant colonies if one or more plasmids were introduced in addition to pArT1. Since none of the used vectors contained *Agrobacterium* virulence genes, the frequency of the yeast cells that acquired G418 resistance reflect the frequency of false positives that can be expected during a promoter trap-assay for the detection of T-DNA integration events.

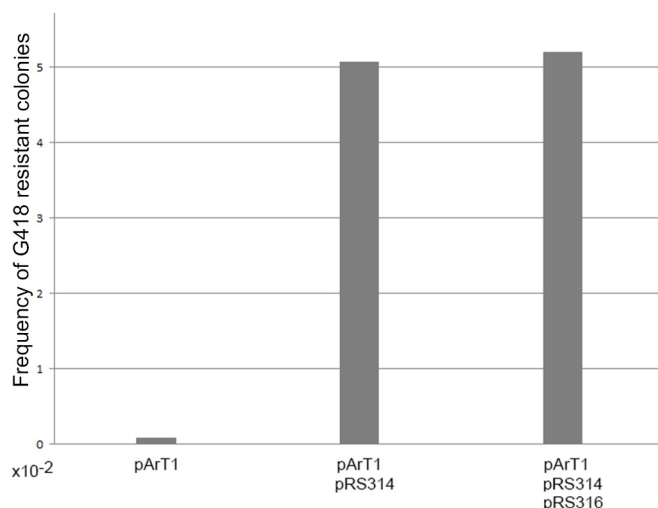


**Figure 10. The influence of additional plasmids on the frequency of acquired G418 resistance.** The yeast cells were screened for G418 resistance after 24 hours of growth on a liquid YPD medium containing 2% glucose. For this, 100  $\mu$ l of a yeast cell suspension with an OD<sub>600 nm</sub> adjusted to 10<sup>-3</sup> was plated on G418 selective plates with YPD medium. The percentage of transformed yeast cells was calculated by dividing the number of G418 resistant colonies by an estimate of the total number of yeast cells.

Because of these observations, it seemed prudent to stop with the development of a flexible plasmid-based system for the expression of *Agrobacterium* virulence genes and instead continue with the expression cassettes integrated in the yeast genome. The *Agrobacterium* virulence cassettes were integrated in the *pdcc6* locus of the yeast genome, leaving the vector delivering the T-DNA to be the only plasmid required. In addition to this, theoretical problems regarding the translation of VirD1 and VirC2 in yeast were avoided by adjusting some rare arginine encoding codons to meet the optimal yeast codon usage. To integrate the virulence genes, the yeast integrative pINT vector [40] was adjusted to harbor several combinations of expression cassettes for the *Agrobacterium* virulence proteins VirD2, VirD1, VirC2, VirC1, VirE2, as well as VIP1 derived from *Arabidopsis*.

To be used as a negative control, a pINT vector containing no virulence gene expression cassettes was integrated into the yeast genome.

### All-in-yeast T-DNA integration assay on solid medium

2 The new yeast strains containing integrated virulence gene expression cassettes were all supplemented with pArT1stop. To accumulate more integration events, the yeast cells were induced for a period of seven days on solid medium while limiting their growth-rate by reducing the temperature to 21 °C and using an induction medium (IM) that is low on carbon sources. This approach is based on the protocol that was successfully used for the AMT of yeast cells [25]. The IM that was used contained galactose instead of glucose to allow the expression of virulence genes. No antibiotics were added to the IM to avoid any potential experimental biases, also no autotrophy selection for maintaining pArT1stop was applied to avoid missing integration events that co-occurred with plasmid loss. During the induction, the yeast cells were maintained on a nitrocellulose filter. To screen for transformants, the yeast cells were washed from these nitrocellulose filters and assayed on G418 resistance.

As shown in Figure 11, the level of acquired G418 resistance in all pArT1stop containing strains was very low under the adapted experimental conditions. For the negative controls, the yeast strain without any Vir protein expression (empty cassette) and the strain expressing VirD1, VirC2, VirC1 and VirE2, but not the crucial relaxase protein VirD2, G418 resistance was only found at a frequency of  $0.07 \times 10^{-6}$  and  $0.06 \times 10^{-6}$ , respectively. For the strain expressing VirD2 and the strain expressing VirD2, VirD1, VirC2 and VirC1, the frequencies were  $5.2 \times 10^{-6}$  and  $4.4 \times 10^{-6}$ , respectively. Although not high in absolute terms, these frequencies were at least 60-fold higher than for the negative controls, statistically significant ( $p < 0.05$ ) for the VirD2 producing strain and nearly so ( $p < 0.08$ ) for the VirD2, VirD1, VirC2 and VirC1 producing strain.

A puzzling aspect of the results depicted in Figure 11, is that the addition of the *Agrobacterium* virulence gene *virD1* to *virD2* did not seem to have a positive impact on the generation of potential integration events, when comparing the first and the second bar of Figure 11. As further discussed below, VirD2 is an enzyme able to cleave ssDNA rather than dsDNA. Expression of VirD2 will then be sufficient for T-strand generation when ssDNA is available, as it might be during stages of plasmid replication. Such opportunistic VirD2-mediated processing of ssDNA might outnumber the processing events where VirD1 in addition to VirD2, is required for the processing of double stranded border sequences [3,26]. If this hypothesis is valid, one would expect that the expression of VirE2, the protein forming a complex with ssDNA, would compete with VirD2 in binding to single stranded RB sequences, hence reducing border processing. Indeed the frequencies of G418 resistance found in the yeast strains carrying an additional *virE2* ORF are low compared to the corresponding frequencies without this ORF (Figure 11).

As for the presence of the *Arabidopsis* VIP1, supposedly involved in the nuclear targeting of VirE2 [37,38], we did not observe an elevation in the acquired G418 resistance when compared with any of the virulence gene combinations harboring the *virE2* ORF without the VIP1 encoding ORF.

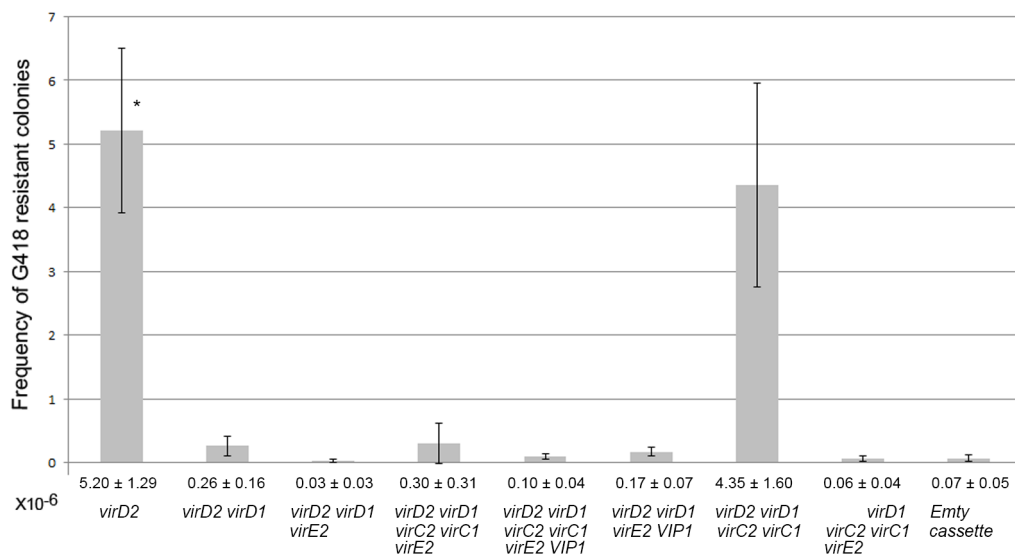


**Figure 11. The frequency of acquired G418 resistance as observed after 7 days of growth on solid medium with galactose.** The yeast strain YPH250 contained pArT1stop in addition to combinations of expression cassettes for VirD2, VirD1, VirC2, VirC1, VirE2 and VIP1, or empty cassettes, integrated in the genome. The averages of four independent experiments are shown. The percentage of transformed yeast cells was calculated by dividing the number of G418 resistant colonies by an estimate of the total number of yeast cells. Error bars represent the SEM. The p-values indicated in this figure resulted from a two-tailed student's T-test comparing the different yeast backgrounds with the empty vector control. An asterisk indicates significant differences with the empty vector control with  $\alpha = 0.05$ .

## Validation of integration events

To analyze potential integration events resulting from the all-in-yeast T-DNA integration assay on solid medium, genomic DNA was isolated from 35 of the G418 resistant colonies. As a negative control for probe hybridization gDNA was isolated from a WT YPH250 strain. As a positive control for probe hybridization and G418 resistance YPH250 transformed with pArT1 *TEF1:KanMX* was used. Southern blot analysis of *Nco*I-digested DNA was performed using a probe against the *KanMX* marker gene. As can be seen in Figure 12, signals of about 6 kb are present in nearly all DNA samples. In lane 2, this signal agrees with a predicted 6.6 kb *Nco*I-linearized pArT1 *TEF1:KanMX* vector, slightly larger than the circa 6 kb signals in most other lanes which agree with a predicted size of 6.3 kb for linearized pArTstop vector. Apart from these vector-derived signals, lane 7, 13, 17 and 28 exhibited fragments of smaller size. These signals have some of the characteristics of randomly integrated processed T-DNA. Firstly, all of these DNA

fragments exceed the 1.25 kb which is the expected size of a fully processed T-DNA. Secondly, the fragments are of varying molecular weight, which is in accordance with the random integration pattern typical for T-DNA integration. Lane 15 and 32 lacked the vector signal that is present in the other lanes, each showing a different signal at higher molecular weight. The increased fragment size could be an indication that T-DNA had integrated far removed from a genomic *NcoI* site. Alternatively, considering the simultaneous disappearance of the normal vector signal, it could be that the complete vector backbone of pArT1 stop had integrated in the yeast genome. However, all the above mentioned DNA fragments that deviating from the expected size of a linearized pArTstop vector could represent drastically modified pArT1stop vectors harboring DNA inserts from an unknown source. DNA isolated from G418 resistant strains harboring pArT1stop but without an additional gDNA fragment that hybridizes to the probe are likely to represent false-positives that have reverted to a state of active expression of the KanMX marker gene. The most likely route leading to the expression of the KanMX marker gene would be the loss or disruption of the previously introduced stopper sequence.

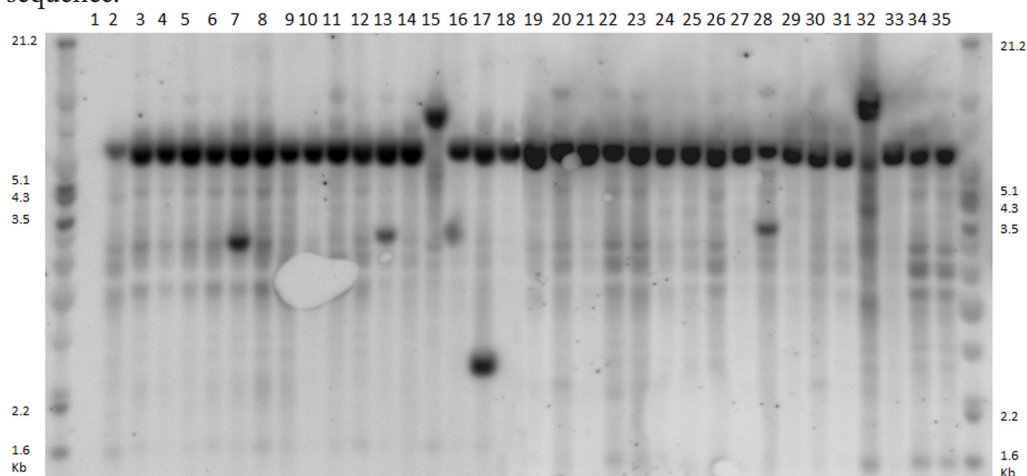


**Figure 12. Southern blot of yeast DNA samples analysed for potential T-DNA integration events.** Genomic DNA derived from G418 resistant colonies was digested with *NcoI* prior to the Southern blotting. The Southern blot was hybridized with a probe against the KanMX marker gene. After linearization by *NcoI* the expected size of the pArT1stop vector backbone is 6.3 Kb. A T-strand resulting from RB and LB processing of the T-DNA is predicted to measure 1255 bp. (1) YPH250 WT, (2) YPH250 (pArT1 prom KanMX), (3-14) YPH250 (pArT1 stop) with *virD2*, (15-18) YPH250 (pArT1 stop) with *virD2 virD1*, (19-30) YPH250 (pArT1 stop) with *virD2 virD1 virC2 virC1*, (31-32) YPH250 (pArT1 stop) with *virE2 virD2 virD1 virC2 virC1*, (33-34) YPH250 (pArT1 stop) with *virE2 VIP1 virD2 virD1 virC2 virC1* (35) YPH250 (pArT1 stop) with *virD1 virC2 virC1*.

Plasmid DNA was isolated from all the yeast strains that potentially harboured T-DNA inserts. The transfer of these plasmids to a WT YPH250 strain always resulted in varying levels of G418 resistance. After these yeast strains were cured from their plasmid content by repetitively transferring the yeast cells to fresh nonselective plates until their His autotrophy was lost, the G418 resistance was also abolished.

These results taken together indicate that modified pArTstop vectors were a likely source for the acquired G418 resistance. Since this observation still does not exclude the presence of genomic T-DNA inserts, a second Southernblot was performed on the yeast strains that had lost their plasmid content. Without exception the loss of the plasmid content accompanied loss of the additional DNA fragments shown in Figure 12 (data not shown). As a positive control the concomitant yeast cells still harbouring pArt1stop were taken along.

## Discussion

In order to get more insight into the role of recipient cell factors during the process of T-DNA integration in the genome of eukaryotic cells, we started the development a system that would skip all preceding steps associated with the cocultivation of *Agrobacterium* cells and the eukaryotic cells subjected to AMT. With the AMT of yeast cells as a model, we therefore attempted to establish an all-in-yeast system, where T-DNA processing, T-strand formation, and T-DNA integration would all occur within yeast cells. This necessitated the construction of a vector harboring an artificial T-DNA substrate, pArT1, as well as inducible expression cassettes harboring *Agrobacterium* virulence gene ORFs and the ORF of VIP1 derived from *Arabidopsis*. All these genes are known to function in T-DNA processing or the integration of processed T-DNA. Ideally, a fully plasmid-based all-in-yeast system for the processing and integration of T-DNA would be the most flexible experimental tool kit to assay the effect of host factors on T-DNA integration; already available mutant yeast backgrounds could simply be equipped with the relevant plasmids to determine the contribution of particular yeast genes during T-DNA integration.

For the construction of a plasmid based all-in-yeast system, two galactose inducible virulence gene expression cassettes were prepared that were loaded with up to six ORFs: five ORFs encoding the *Agrobacterium* Vir proteins VirD2, VirD1, VirC2, VirC1 and VirE2 and one ORF encoding the *Arabidopsis* transcriptional activator VIP1. To assay if VirD2 and VirD1, which are pivotal to T-DNA processing, were properly expressed by the inducible virulence gene expression cassettes and to investigate whether these proteins were able to reach the yeast nucleus, GFP fusions with virD2 and virD1 were constructed. With confocal microscopy it was shown that VirD2 was targeted to the nucleus. VirD1, which has no predicted NLS in contrast to VirD2, was only targeted to the nucleus when co-expressed with VirD2. The second requirement for the construction of a well-defined all-in-yeast T-DNA processing and integration system was the construction of a T-DNA like substrate for VirD2 catalyzed border processing. For this, the small *ARS/CEN* yeast shuttle vector pRS313 [39] was equipped with an artificial T-DNA consisting out of minimal octopine type ODRB and LB sequences flanking a promoterless KanMX marker gene. Following VirD2 mediated border processing, the promoterless KanMX marker gene present on the released T-strand, would only become active after being integrated as a T-DNA into a transcriptionally active region of the yeast genome. The borders used proved to be highly active when tested for their ability to support AMT.

The first all-in-yeast assay for the integration of processed T-DNA derived from pArT1 resulted into acquired G418 resistance for 1-3% of the yeast cells, independent from the expression of virulence genes. Further analysis of the resulting G418 resistant colonies pointed out that the source of the G418 resistance was plasmid-based and thus could not be the result of T-DNA integration events. The high frequency of G418 resistance was not expected as plating of about 700 yeast cells harboring pArT1 did not yield any resistant colonies at G418 concentrations of 50 µg/ml and higher (Figure 5), while 200 µg/ml was used for the integration assays. Remarkably, comparing the G418 resistance of yeast cells only transformed with pArT1 to that of cells combining pArT1 with one or two additional *ARS/CEN* type vectors of the pRS series, pRS314 and pRS316, showed a steep increase in the frequency of G418 resistance when more plasmids were present (Figure 9). Since pRS314 and pRS316 were not equipped with *Agrobacterium* virulence genes, the integration of processed T-DNA could not be the source of the acquired G418 resistance. When plasmids were isolated from an liquid overnight yeast culture, after transforming *E. coli* cells with the yeast isolates to amplify the amount of plasmid to reach sufficient levels for restriction analysis [50], we frequently found structural changes when more than one plasmids was present. These aberrant plasmids were not recovered when the plasmids were isolated from a yeast strain with the homologous repair gene *rad52* deleted (data not shown) showing that indeed these highly similar plasmids are the subject of intensive homologous recombination. Although not the focus of this research, it would be interesting to know if the act of homologous repair itself could lead to stochastic expression from an ORF independent of a promoter sequence being present. For further experimentation, although up to four *ARS/CEN* type plasmids were claimed to be stably maintained in yeast [39,49], we decided to refrain from using more than a single plasmid of this type.

Trying to reduce the background G418 resistance, the virulence gene expression cassettes were integrated into the yeast genome to avoid any detrimental effects resulting from the maintenance of multiple plasmids. In addition to that, pArT1 was equipped with a stop codon rich sequence upstream the ODRB to construct pArT1stop. As can be seen in Figure 11, the presence of pArT1 alone did result in the presence of about 0.1% of G418 resistant cells per 1000. Whatever the reason, we expected that the presence of the stopper sequence should prevent such an apparently stochastic expression of the KanMX marker gene. Using this revised system, while also allowing for a 7 day period of T-DNA processing and T-DNA integration, we obtained indications that the presence of Vir proteins affected the development of G418 resistance in yeast cells. First of all, expression of VirD2 only in a yeast strain harboring pArT1stop resulted into a significantly higher level of acquired G418 resistance when compared to an empty vector control (Figure 11). Although not fully significant, also the VirD2, VirD1, VirC2 and VirC1 expressing strain exhibited an elevated frequency of G418 resistance. Most strikingly, expression of VirD1 in addition to VirD2 did not at all enhance the frequency of obtained G418 resistance, although it has been shown that VirD1 is required for VirD2-mediated T-DNA processing on a dsDNA template [3] in its absence, VirD2 can only process ssDNA [2].

To validate if the G418 resistant yeast colonies that resulted from this experiment were representing genomic integration events of processed T-DNA, genomic DNA derived from a selection of the resistant colonies was analyzed by Southern blotting. Only in DNA derived from 4 out of 32 colonies that in theory could have processed T-DNA, T-DNA might have integrated in the yeast genome while the pArT1stop plasmid was still maintained. In two additional G418 resistant colonies, a high molecular weight signal was detected while the original pArT1stop signal was no longer present.

The genomic DNA samples of these six yeast strains both cured from their plasmid content and with the pArT1stop vector still present, were subjected to further analysis. As it turned out, the observed G418 resistance was linked to the presence of the pArT1stop plasmid or a derivative thereof. As would be expected for a plasmid based property, the G418 resistance proved to be a portable trait when transferring the plasmids isolated from the G418 resistant strains to a fresh WT yeast strain. A second Southern blot using these yeast strains, both with and without their plasmid content, elucidated that the observed extra DNA fragments were after all not of genomic origin and could therefore not constitute T-DNA integrants. A plausible explanation for the role of VirD2 in the generation of G418 resistance could be that the binding of VirD2 to single stranded plasmid DNA during cell proliferation is causing faulty DNA replication. This could lead to modifications of the pArT1stop vector thereby reinstating KanMX marker gene transcription. The results depicted in Figure 11 indicating that VirD1, known to aid VirD2 in the processing of double stranded border sequences [3,26], does not have a positive impact on the acquisition of G418 resistance would be in line with this explanation.

Why this seemingly straight forward approach for *in vivo* T-DNA processing and integration did not result in actual integration events and why the expression of VirD2 apparently aided in the acquisition of plasmid based G418 resistance, are difficult questions to address that urge to look closely into several aspects of cellular biology. First of all, it is important to consider that the screening method used for the detection of T-DNA integration events is based on the promoter trap-assay. A limitation of using a promoter trap-assay for this purpose is that only a fraction of the integration events is detected. Theoretically, only one of every six random integration events within a genomic ORF will be in the correct reading frame thereby allowing for the expression of the KanMX marker gene. In addition to this, any integration events outside the transcriptionally active regions of the yeast chromosome are not likely to result in a strong activation of the KanMX marker gene. Therefore, the integration frequency needs to be reasonably high, at the very least higher than the frequency of false positives we have encountered. There are still several leads that could help to further improve the all-in-yeast approach to detect T-DNA integration events that are worth looking into. Pioneering *in vitro* experiments revealed that purified VirD2 and VirD1 can process superhelical plasmid DNA substrate but not linearized DNA or plasmid DNA that was brought into a relaxed conformation by treatment with topoisomerase I [3]. Experiments where border processing was successfully induced, using a double stranded border sequences as a substrate, utilized superhelical plasmid DNA isolated from *E. coli*.

The earliest report concerning the *in vivo* T-strand production describes experiments where expression vectors carrying the ORFs of VirD2 and VirD1 were combined with a plasmid based, T-DNA substrate in *E. coli* [26]. In a more recent report plant cells were transformed by particle bombardment with VirD1 and VirD2 expressing plasmids in combination with a plasmid carrying a T-DNA substrate which led to the integration of newly processed T-strands. Later maize protoplasts were chemically transformed with a similar combination of plasmids also resulting in T-DNA processing and integration [27,28]. It is plausible that for efficient *in vivo* T-DNA processing superhelical plasmid DNA is required as has been shown for *in vitro* T-DNA processing. Differences in plasmid topology between budding yeast and Gram-negative bacteria like *Agrobacterium* and *E. coli* could very well be the Achilles heel of the all-in-yeast approach to induce *in vivo* T-DNA processing and integration. Indeed DNA isolated from bacterial plasmids has been shown to be more negatively supercoiled than plasmids isolated from budding yeast equipped with an *ARS* type of replication origin [51].

The observations described here are in line with the idea that the ds border sequences of pArT1stop are recalcitrant to border processing when in yeast but not in *Agrobacterium* or another prokaryote. It would be interesting to test if the expression of a prokaryotic topoisomerase in yeast in concert with the expression of the *Agrobacterium* virulence proteins could enhance negative supercoiling of pArT1stop and thus enhance T-DNA processing in yeast. Additionally, it could be worthwhile to repeat this experiment with pArT1 integrated into the yeast genome, with the centromere sequence removed to avoid erroneous chromosome segregation during cell division, to later screen for the VirD2 mediated formation of T-DNA circles as are described in Chapter 3 of this thesis.

## Materials and methods

### Assembly of the *Agrobacterium* virulence gene expression cassettes I and II

#### *PCR amplification and nucleotide sequencing*

For the amplification of the *Agrobacterium tumefaciens* virulence genes used for the assembly of the yeast expression vectors, genomic DNA from the *Agrobacterium tumefaciens* octopine Ti plasmid; pTi-15955, NCBI reference sequence: NC\_002377.1 [43] was used as a template. Primers used for the amplification of *Agrobacterium tumefaciens* virulence genes are listed in Table 1. The PCR products were first cloned into pSKN-SgraI [52] which was then renamed to pOPSL n as is listed in table 1 column A, “n” indicating the number assigned to each separate amplicon. These pOPSL n vectors were used for nucleotide sequencing of the inserts using M13 primers. For the successive cloning steps, fragments were excised from the pOPSL vectors as depicted in Table 1.

For the amplification of the upstream activating sequences (UAS) and transcriptional terminators (TT) from the yeast galactose pathway, genomic DNA from *Saccharomyces cerevisiae* strain YPH250 was used (*MAT $\alpha$* , *ura3-52*, *lys2-801*, *ade2-101*, *trp1- $\Delta$ 1*, *his3- $\Delta$ 200*, *leu2- $\Delta$ 1* [39]). Primers used for the amplification of *Saccharomyces cerevisiae* DNA are listed in Table 2. In table 1 and 2 restriction sites are underlined. Column A shows the name of the vector and the restriction sites that were used to clone the digested amplicon into pSKN-SgraI. Column C depicts an number assigned to each (excised) fragment.

**Table 1 Amplicons used for the construction of the virulence gene cassettes I and II.** *Agrobacterium tumefaciens* virulence genes and the *Arabidopsis* transcriptional activator VIPI.

| Name primers                  | Sequence 5'-3' Target cleavage site                                                                       | A                                      | B                                           | C  |
|-------------------------------|-----------------------------------------------------------------------------------------------------------|----------------------------------------|---------------------------------------------|----|
| VirD1 BamHI F<br>VirD1 NheI R | ATTAGGATCCAAAAATGTCAAAACACACCAGAGTC<br>ATATGCTAGCTACAAGGCACCTTTCAATAGC                                    | pOPSL 2<br><i>Bam</i> HI, <i>Nhe</i> I | <i>virD1</i><br><i>Bam</i> HI, <i>Nhe</i> I | 2  |
| VirC2 Sall F<br>VirC2 XbaI R  | TAAAGTCGACCTTCACCAATTCCTCGATGG<br>ATTATCTAGAAAATGGCAATTCGCAAGCCCG                                         | pOPSL 4<br><i>Xba</i> I, <i>Sal</i> I  | <i>virC2</i><br><i>Xba</i> I, <i>Sal</i> I  | 4  |
| VirC1 Sall F<br>VirC1 BglII R | ATATGTCGACTTCAAGCCTCCAAGATTTTGC<br>TAAGAGATCTAAAAATGCAACTTTTGACGTTTGTCTC                                  | pOPSL 6<br><i>Sal</i> I, <i>Bgl</i> II | <i>virC1</i><br><i>Sal</i> I, <i>Bgl</i> II | 6  |
| VirD2 BglII F<br>VirD2 XhoI R | TAATAGATCTAAAAATGCCCGATCGCGCTCAAG<br>ATTACTCGAGCCTTCCTGTCTCTAGGTCC                                        | pOPSL 8<br><i>Bgl</i> II, <i>Xho</i> I | <i>virD2</i><br><i>Bgl</i> II, <i>Xho</i> I | 8  |
| VirE2 BamHI F<br>VirE2 NheI R | ATTAGGATCCAAAAATGGATCTTCTGGCAATGAG<br>TCATGCTAGCCAGAAGGAACGTCAAAGC                                        | pOPSL 9<br><i>Bam</i> HI, <i>Nhe</i> I | <i>virE2</i><br><i>Bam</i> HI, <i>Nhe</i> I | 9  |
| VIPI SpeI F<br>VIPI XhoI R    | ATTA <del>ACTAGT</del> AAAAATGGAAGGAGGAGGAAGAGG<br>TAAT <del>CTCGAGT</del> AGGTGACACGA <del>ACTCAGC</del> | pOPSL 11<br><i>Spe</i> I, <i>Xho</i> I | <i>VIPI</i><br><i>Spe</i> I, <i>Xho</i> I   | 11 |

**Table 2 Amplicons used for the construction of the virulence gene cassettes I and II.** *Saccharomyces cerevisiae* regulatory elements.

| Name primers                                                    | Sequence 5'-3' Target cleavage site                                          | A                                      | B                                                      | C |
|-----------------------------------------------------------------|------------------------------------------------------------------------------|----------------------------------------|--------------------------------------------------------|---|
| TTGAL7 <i>Xba</i> I F<br>TTGAL7 <i>Nhe</i> I R                  | CAGCTCTAGAAGCATAAGAAGTTGTGGCG<br>ATGCGCTAGCGAAAGTGAATATTCATTCATATC           | pOPSL 1<br><i>Nhe</i> I                | <b>TTgal7</b><br>Not excised                           | 1 |
| UASGAL7TTGAL10 <i>Bam</i> HI F<br>UASGAL7TTGAL10 <i>Xho</i> I R | GAGTAGGATCCAATATTCAACTGTTTTTTTATCATG<br>ATCGCTCGAGTTGCCAGCTTACTACTCCTTC      | pOPSL 3<br><i>Bam</i> HI, <i>Xho</i> I | <b>UASgal7, TTgal10</b><br><i>Bam</i> HI, <i>Xho</i> I | 3 |
| UASGAL10-GAL1 <i>Spe</i> I F<br>UASGAL10-GAL1 <i>Bam</i> HI R   | TGACACTAGITTTGAATTTCAAAATCTTACTTTTTTTTGG<br>GCATGGATCCCTTGACGTTAAAGTATAGAGGT | pOPSL 5<br><i>Spe</i> I, <i>Bam</i> HI | <b>UASgal10,1</b><br><i>Spe</i> I, <i>Bam</i> HI       | 5 |
| TTGAL1 <i>Sal</i> I F<br>TTGAL1 <i>Xho</i> I R                  | GATGGTCGACTTCTTTTTTTTACTTTGTTCAGAAC<br>TATACTCGAGACATATGATAAGACGTCAACTGC     | pOPSL 7<br><i>Sal</i> I                | <b>TTgal1</b><br>Not excised                           | 7 |

To create a series of expression cassettes harboring several combinations of *Agrobacterium* virulence genes and *VIPI* derived from *Arabidopsis*, the amplicons mentioned in Table 1 were successively cloned in pSKN-SgraI. The resulting *Agrobacterium* expression cassettes were initially cloned into the yeast shuttle vectors pRS314 and pRS316 and were later combined into an adjusted version of the integrative vector pINT [40]. The construction of the *Agrobacterium* expression cassette I (Figure 1) was initiated as follows: the cloning vector pOPSL1 was digested with *Bam*HI and *Xho*I and fragment 2 was inserted into this vector. The resulting plasmid was digested with *Nhe*I and *Bam*HI. Fragment 3 was inserted into this vector. Independent of these cloning steps, pOPSL1 was digested with *Bgl*II and *Sal*I and similarly digested fragment 6 was inserted into this vector. The resulting plasmid was digested with *Bgl*II and *Spe*I and similarly digested fragment 5 was inserted subsequently. These steps resulted in the two vectors containing the galactose-inducible expression cassettes; pSKN-SgraI with *TTgal7:virD1:UASgal7:TTgal10* and pSKN-SgraI with *UAS10,1:virC1:TTgal1* to which will be referred as vector A and vector B, respectively. Vector A and B were combined to form a single vector by digesting both vectors with *Sca*I and *Xho*I. Digestion of vector A resulted in a 3269 fragment containing the *virD1* ORF and the pBR322 replication origin of pSKN-SgraI and can thus be considered the vector backbone. Digestion of vector B resulted in the required 2887 bp insert containing the *virC1* ORF. Ligating these two fragments reconstituted the Ampicillin maker gene of pSKN-SgraI. This cloning step resulted in plasmid “pSKN-SgraI with *TTgal7:virD1:UASgal7:TTgal10:UAS10,1:virC1:TTgal1*”. Finally, this plasmid was digested with *Xho*I and *Spe*I and fragment 4 was inserted, resulting in “pSKN-SgraI with *TTgal7:virD1:UASgal7: virC2:TTgal10:UAS10,1:virC1:TTgal1*” which was renamed as *Agrobacterium* virulence cassette I (Figure 1).

To generate *Agrobacterium* virulence gene expression cassette II (Figure 1) pOPSL1 was digested with *Bam*HI and *Xho*I and fragment 2 was inserted into this vector. The resulting plasmid was digested with *Nhe*I and *Bam*HI and similarly digested fragment 9 was inserted. Independent of these cloning steps, pOPSL7 was digested with *Bgl*II and *Sal*I and similarly digested fragment 8 was inserted. The resulting plasmid was digested with *Bgl*II and *Spe*I, followed by insertion of similarly digested fragment 5. These steps resulted in two different vectors; “pSKN-SgraI with *TTgal7:virE2:UASgal7:TTgal10*” and “pSKN-SgraI with *UAS10,1:virD2:TTgal1*” to which will be referred to as vector C and vector D respectively. Vector C and D were combined to form a single vector by digesting both vectors with *Sca*I and *Xho*I and ligating the two fragments carrying the *Agrobacterium tumefaciens* virulence genes together. Digestion of vector C resulted in a 4439 bp fragment containing the *virE2* ORF and the pBR322 replication origin of pSKN-SgraI and can thus be considered the vector backbone. Digestion of vector D resulted in the required 3476 bp insert containing the *virD2* ORF. Ligating these two fragments reconstituted the Ampicillin maker gene of pSKN-SgraI. This cloning step step resulted into the plasmid “pSKN-SgraI with *TTgal7:virE2:UASgal7:TTgal10:UAS10,1:virD2:TTgal1*”. Finally, this plasmid was digested with *Xho*I and *Spe*I and fragment 11 was inserted, resulting in

“pSKN-SgrAI with *TTgal7:virE2:UASgal7:TTgal10:VIPI:UAS10,1:virD2:TTgal1*” which was renamed as *Agrobacterium* virulence cassette II. By skipping cloning steps, also a vector A without *virD1*, a vector B without *virC1*, a vector C without *virE2*, and a vector D without *virD2* were constructed enabling the construction of several versions of *Agrobacterium* expression cassettes where ORFs of interest were omitted while maintaining the inducible promoter and terminator sequences.

The *Agrobacterium* virulence gene expression cassettes were cloned as *NotI* fragments into the *ARS/CEN* shuttle vectors pRS314 and pRS316 [26]. Virulence gene cassette I was cloned into pRS316 and virulence gene cassette II was cloned into pRS314 resulting in pRS316 expressing *virD1*, *virC2* and *virC1* and pRS314 expressing *virE2* and *virD2*. In second instance, to avoid the presence of *ARS/CEN* plasmids besides pArT1 or its derivative pArT1stop, the expression cassettes were transferred to an adjusted pINT vector [40] that by means of homologous recombination via a double cross-over, can be integrated into the *PDC6* locus of yeast [53]. To make pINT compatible with the experimental set up, where G418 was used for the screening of integration events, a small adjustment was required. For this, the *APT1* marker gene was excised by digesting pINT with *HindIII*. The resulting 1401 bp fragment that carried a part of the *pdcc6* locus was ligated back to the 5714 bp fragment. The resulting plasmid was cut with *SpeI* and *BamHI* and ligated with a *SpeI* *BglII* restriction fragment of pYM25 carrying *hygromycin-B-phosphotransferase* conferring resistance to hygromycin. Several versions of *Agrobacterium* virulence gene expression cassettes I and II were cloned in this modified pINT vector. The resulting integrative vectors were linearized using *BssHII* to enhance the integration efficiency into the yeast genome after using a LiAc/ssDNA/PEG yeast transformation protocol [54].

### *Construction of the virulence gene::GFP expressing vectors*

With PCR, using pSKN-GFPH6 as a template, restriction sites were added to the GFP ORF corresponding to restriction sites present within the *virD2* and *virD1* ORFs and unique to the virulence gene expression plasmids. The resulting amplicons were digested with the corresponding restriction enzymes and cloned into the similarly digested virulence gene expression vectors pRS314 *virE2:virD2* and pRS316 *virD1:virC2:virC1*. To clone a GFP ORF in *virD1*, GFP was amplified using the primers VirD1-GFP FW (5' GTCAGGTAGTAAAGGAGAAGAAGCTTTTCACTGG 3') and VirD1-GFP RV (5' GTGCCTTTGTATAGTTCATCCATGCCATG 3'). The resulting amplicon could directly be cloned as a blunt fragment to pRS316 with *virD1* *virC2* and *virC1* digested with *AjiI*. To clone the GFP ORF in *virD2*, GFP was amplified using the primers: VirD2-GFP FW (5' ATTACGGACCGTGGTAGTAAAGGAGAAGAAGCTTTT 3') and VirD2-GFP RV (5' ATTACGGACCGATTTGTATAGTTCATCCATGCCATG 3'). The *CpoI* restriction sites added to the original sequence are underlined. The resulting amplicon was digested with *CpoI* and cloned into pRS314 with *virE2* and *virD2* also digested with *CpoI*. The resulting vectors pRS314:*virE2::virD2::GFP::virD2* and pRS316:*virD1::GFP::virD1:virC2:virC1* were transferred to the yeast strain YPH250 by LiAc/ssDNA/PEG transformation.

## Confocal Microscopy

To capture the GFP signal, a band pass filter was used of 505-530 nm. The laser intensity was set to 4.5% with 1 airy unit and 500 master gain. For the DAPI staining, slightly modified settings were used to increase the signal. For the post-processing of the images, the following software was used: Image J [55] was used to restack a selection of confocal images derived from z-stacks. Adobe Photoshop was used to superimpose the GFP signal on the bright-field images.

## Construction of the pArT vectors

### *Cloning strategy*

For the construction of pArT1, the yeast shuttle vector pRS313 [39] was used as backbone. This plasmid is equipped with an autonomously replicating sequence (*ARS*) and a centromere. For selection, a histidine autotrophy marker is available that is non-reverting in the yeast strain YPH250 [39]. A minimal octopine ODRB and LB were cloned in pRS313 to serve as substrate for VirD2 catalyzed border processing. The ODRB and LB sequences are identical to the *TLrb* and *TLlb* of the octopine Ti plasmid pTi-15955, NCBI reference sequence: NC\_002377.1 [43]. For the construction of the border sequences the following oligos were annealed (recognition sites added to the original sequence are underlined, the RB and LB sequences are depicted in lowercase).

LB-FW:

5'- GCTAGCGTCTCGcggcaggatatattcaattgtaaatGGCTTCATGTCCGGGAAATCTACATG  
GATCAGCAATGAGTATGCAGCT-3',

LB-RV:

5'-ATCCATGTAGATTTC<sup>CGG</sup>GACATGAAGCCatttacaattgaatatatcctgccgCGAC  
GCTAGCGAGCT-3',

RB-FW:

5'-TCGACTGACtggcaggatatataccgttgtaattTGAGCTCGTGTGAATAAGTCGCTGTGT  
ATGTTTGT<sup>TTG</sup>ATTGTGCTAGC-3',

RB-RV:

5'-TCGAGGCTAGCACAATCAAACAAACATACACAGCGACTTATTCACAG  
AGCTCAaattacaacggtatatatcctgccGTCAG-3'

The KanMX marker gene used for vector construction originates from the *E. coli* transposon Tn903. To obtain this gene, the binary vector pSDM8000 [22] was digested with *Nco*I and *Sal*I. This DNA fragment containing the *KanMX* gene with *TEF1* terminator was cloned into pSKN-SgraI. A *Sac*I site was removed by blunting *Sac*I overhangs with T4 DNA polymerase.

Successively, the blunted vector ends were religated. The removal of this *SacI* site was required to facilitate future cloning steps. The resulting vector was digested with *SalI* and *BamHI* to excise the *KanMX* ORF with the TEF1 terminator, but without a promoter sequence. This fragment was cloned into the yeast shuttle vector pRS313 [39] which was similarly digested. It is important to note that in the pArT1 vector, no in frame stop codons are present from the VirD2 nicking site up until the start codon of the *KanMX* ORF.

#### *The addition of the stopper sequence to pArT1;*

To avoid unwanted expression of the *KanMX* marker gene from pArT1, a stop codon-rich “stopper” sequence was introduced in pArT1, containing stop codons in six different reading frames. The adjusted pArT1 vector was renamed to pArT1stop. The stopper sequence was constructed by annealing the two following oligos:

Stopper *XbaI* FW;

5'-TCGACTAGTCTAGACTAGTGATCATTAAGTGATCATTAAGTGATCATTAAGC-3',

Stopper *XbaI* RV;

5'-TCGAGTTAATGATCACTTAATGATCACTTAATGATCACTAGTCTAGACTAG-3'.

(*XbaI* sites underlined)

The stopper sequence was introduced into a unique *XhoI* site with the 3'-end of the FW oligo directly upstream the OD sequence (Figure 10). The *XbaI* restriction site, present in the stopper sequence, was used to determine the presence of the insert in the vector backbone. The orientation of the stopper sequence was determined by means of sequencing. A clone with the *XbaI* site closed to the RBOD sequence was selected for further experimentations although the stopper sequence would have sufficed as a translational stop in both orientations.

### **T-DNA integration assays**

#### *T-DNA integration assay on liquid medium*

For this assay the following vector combinations were transferred to the yeast strain YPH250 harboring pArT1 combined with pRS314 *virD2* and pRS316, pArT1 combined with pRS314 *virD2* and pRS316 *virD1:virC2:virC1*, and pArT1 combined with pRS314 *virE2:virD2* and pRS316 *virD1:virC2:virC1*. As a negative control for the activity of the *Agrobacterium* virulence proteins, pArT1 was combined with pRS314 and pRS316.

The resulting yeast strains were precultured at 30 °C by shaking at, 200 rpm for 18-24 hrs on a selective minimal synthetic defined base (SD) liquid medium. Amino acids were added fitting the auxotrophic requirements of YPH250, omitting Ura and Trp for plasmid maintenance. The cell suspensions were centrifuged for two minutes at 2500 rpm and pellets were washed with 50 ml of sterile H<sub>2</sub>O. The yeast cells were transferred to 50 ml liquid SD medium containing 2% glucose or 2% galactose with the OD<sup>600</sup> adjusted to 0.25. If required, 2% raffinose can be added to the galactose containing medium to boost yeast culture growth to make a better comparison to growth on glucose containing medium, while not interfering with galactose-inducible gene expression. Alternatively, as was done to obtain the results in figure 7 prior to the T-DNA integration assay on liquid medium, the required yeast strains can be precultured on liquid medium with glycerol as carbon source instead of glucose to adapt to growth on alternative carbon sources. Amino acids were added fitting the auxotrophic requirements of YPH250, omitting Ura and Trp for plasmid maintenance. The yeast cells were incubated for the chosen time period at 30 °C, 200 rpm. The cell suspensions were centrifuged for two minutes at 2500 rpm and the pellets were washed with 50 ml sterile H<sub>2</sub>O. To screen for any integration events, a dilution series of the incubated cells was plated on YPD medium containing 200 µg/ml G418 (Duchefa G0175, Haarlem, The Netherlands) and on YPD medium without selection to estimate the total number of viable yeast cells.

#### *T-DNA integration assay on solid medium*

The yeast strains were precultured for 18-24 hrs in 5 ml minimal synthetic defined base (SD) liquid medium with either 2% glucose or 2% galactose and raffinose at 30°C, 200 rpm. To these cell suspensions, 45 ml YPD was added or YPD with 2% galactose instead of glucose. The yeast cells were grown for 5 hrs at 30 °C, 200 rpm. The cell suspensions were centrifuged for two minutes at 2500 rpm. The pellets were washed with 10 ml of a sterile 0.9% NaCl solution and the OD<sup>600</sup> of the yeast cells was adjusted to 0.1 in this solution. Nitrocellulose filters were placed on IM plates as described in [25]. The induction plates contained 0-2% galactose instead of glucose. Amino acids complementing auxotrophic requirements of the yeast strain YPH250 were added. No antibiotics were added to the IM. For each treatment, 100 µl of cell suspension was administered to the nitrocellulose filters. The induction plates were incubated for 7x24 hrs at 21 °C. To screen for integration events, the yeast cells were washed from the filters with the 0.9% NaCl solution. A dilution series of the collected yeast cells was plated on YPD medium containing 200 µg/ml G418, and on YPD medium without G418 to estimate the total number of viable yeast cells.

## Southern blotting

Isolated genomic DNA was digested with *Nco*I and transferred to a positively charged nylon membrane. A DIG labeled probe for the KanMX gene was produced by PCR with pSDM8000 plasmid DNA [40] as a template, using the primers KanMX FW (5' GGGTAAGGAAAAGACTCACG 3') and KanMX RV (5' GCCGTTTCTGTAATGAAGGAG 3') and the "PCR DIG labeling mix" (Roche, cat. no. 11585550910). For the remainder of the Southern blotting procedure, the DIG application manual of Roche was followed ([www.roche-applied-science.com](http://www.roche-applied-science.com)).

## Acknowledgements

The authors would like to thank Philippe Sakalis for his great help with the confocal imaging, Roel de Haan for his participation in the analysis of yeast strains harboring potential integration events and Erich Lanka of the Max-Planck-Institut für Molekulare Genetik for kindly providing us with the VirD2 antibody.

## References

1. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A* 79: 6322-6326.
2. Jasper F, Koncz C, Schell J, Steinbiss HH (1994) Agrobacterium T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proc Natl Acad Sci U S A* 91: 694-698.
3. 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.
4. Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, et al. (2000) VirB/D4-dependent protein translocation from Agrobacterium into plant cells. *Science* 290: 979-982.
5. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, et al. (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.
6. Ward JE, Akiyoshi DE, Regier D, Datta A, Gordon MP, et al. (1988) Characterization of the virB operon from an Agrobacterium tumefaciens Ti plasmid. *J Biol Chem* 263: 5804-5814.
7. Durrenberger F, Cramer 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.
8. Peralta EG, Hellmiss R, Ream W (1986) Overdrive, a T-DNA transmission enhancer on the A. tumefaciens tumour-inducing plasmid. *Embo J* 5: 1137-1142.
9. Toro N, Datta A, Carmi OA, Young C, Prusti RK, et al. (1989) The Agrobacterium tumefaciens virC1 gene product binds to overdrive, a T-DNA transfer enhancer. *J Bacteriol* 171: 6845-6849.
10. 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.
11. Lu J, den Dulk-Ras A, Hooykaas PJ, Glover JN (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.
12. Lundquist RC, Close TJ, Kado CI (1984) Genetic complementation of Agrobacterium tumefaciens Ti plasmid mutants in the virulence region. *Mol Gen Genet* 193: 1-7.
13. Citovsky V, Guralnick B, Simon MN, Wall JS (1997) The molecular structure of agrobacterium VirE2-single stranded DNA complexes involved in nuclear import. *J Mol Biol* 271: 718-727.
14. Citovsky V, Zupan J, Warnick D, Zambryski P (1992) Nuclear localization of Agrobacterium VirE2 protein in plant cells. *Science* 256: 1802-1805.
15. Grange W, Duckely M, Husale S, Jacob S, Engel A, et al. (2008) VirE2: a unique ssDNA-compacting molecular machine. *PLoS Biol* 6: e44.
16. Das A (1988) Agrobacterium tumefaciens virE operon encodes a single-stranded DNA-binding protein. *Proc Natl Acad Sci U S A* 85: 2909-2913.

17. Citovsky V, Wong ML, Zambryski P (1989) Cooperative interaction of *Agrobacterium* VirE2 protein with single-stranded DNA: implications for the T-DNA transfer process. *Proc Natl Acad Sci U S A* 86: 1193-1197.
18. Christie PJ, Ward JE, Winans SC, Nester EW (1988) The *Agrobacterium tumefaciens* virE2 gene product is a single-stranded-DNA-binding protein that associates with T-DNA. *J Bacteriol* 170: 2659-2667.
19. Tzfira T, Rhee Y, Chen MH, Kunik T, Citovsky V (2000) Nucleic acid transport in plant-microbe interactions: the molecules that walk through the walls. *Annu Rev Microbiol* 54: 187-219.
20. Li J, Krichevsky A, Vaidya M, Tzfira T, Citovsky V (2005) Uncoupling of the functions of the Arabidopsis VIP1 protein in transient and stable plant genetic transformation by *Agrobacterium*. *Proc Natl Acad Sci U S A* 102: 5733-5738.
21. Lacroix B, Loyer A, Citovsky V (2008) Association of the *Agrobacterium* T-DNA-protein complex with plant nucleosomes. *Proc Natl Acad Sci U S A* 105: 15429-15434.
22. van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J* 20: 6550-6558.
23. van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31: 826-832.
24. Bundock P, Hooykaas PJ (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.
25. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *Embo J* 14: 3206-3214.
26. De Vos G, Zambryski P (1989) Expression of *Agrobacterium* nopaline-specific VirD1, VirD2, and VirC1 proteins and their requirement for T-strand production in *E. coli*. *Mol Plant Microbe Interact* 2: 43-52.
27. Hansen G, Chilton MD (1996) "Agrolistic" transformation of plant cells: integration of T-strands generated in planta. *Proc Natl Acad Sci U S A* 93: 14978-14983.
28. Hansen G, Shillito RD, Chilton MD (1997) T-strand integration in maize protoplasts after codelivery of a T-DNA substrate and virulence genes. *Proc Natl Acad Sci U S A* 94: 11726-11730.
29. Rademaker R (2007) Evaluation of an *A.tum*-derived system for integration of adenovirally delivered transgene cassettes in mammalian cells. Thesis Chapter 6: 84-92.
30. Bram RJ, Lue NE, Kornberg RD (1986) A GAL family of upstream activating sequences in yeast: roles in both induction and repression of transcription. *Embo J* 5: 603-608.
31. Bryant GO, Prabhu V, Floer M, Wang X, Spagna D, et al. (2008) Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biol* 6: 2928-2939.
32. Selleck SB, Majors JE (1987) In vivo DNA-binding properties of a yeast transcription activator protein. *Mol Cell Biol* 7: 3260-3267.
33. Sellick CA, Campbell RN, Reece RJ (2008) Galactose metabolism in yeast-structure and regulation of the leloir pathway enzymes and the genes encoding them. *Int Rev Cell Mol Biol* 269: 111-150.

34. Jayaswal RK, Veluthambi K, Gelvin SB, Slightom JL (1987) Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules in *Escherichia coli* by a virD-encoded border-specific endonuclease from *Agrobacterium tumefaciens*. *J Bacteriol* 169: 5035-5045.
35. Michiels CB, Ram AF, Hooykaas PJ, van den Hondel CA (2004) *Agrobacterium*-mediated transformation of *Aspergillus awamori* in the absence of full-length VirD2, VirC2, or VirE2 leads to insertion of aberrant T-DNA structures. *J Bacteriol* 186: 2038-2045.
36. Ward DV, Zambryski PC (2001) The six functions of *Agrobacterium* VirE2. *Proc Natl Acad Sci U S A* 98: 385-386.
37. Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. *Science* 318: 453-456.
38. Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *Embo J* 20: 3596-3607.
39. 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.
40. Meijer AH, Ouwerkerk PB, Hoge JH (1998) Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. *Yeast* 14: 1407-1415.
41. Howard EA, Zupan JR, Citovsky V, Zambryski PC (1992) The VirD2 protein of *A. tumefaciens* contains a C-terminal bipartite nuclear localization signal: implications for nuclear uptake of DNA in plant cells. *Cell* 68: 109-118.
42. Tinland B, Koukolikova-Nicola Z, Hall MN, Hohn B (1992) The T-DNA-linked VirD2 protein contains two distinct functional nuclear localization signals. *Proc Natl Acad Sci U S A* 89: 7442-7446.
43. Schrammeijer B, Beijersbergen A, Idler KB, Melchers LS, Thompson DV, et al. (2000) Sequence analysis of the vir-region from *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *J Exp Bot* 51: 1167-1169.
44. Thomashow MF, Nutter R, Montoya AL, Gordon MP, Nester EW (1980) Integration and organization of Ti plasmid sequences in crown gall tumors. *Cell* 19: 729-739.
45. Machowski EE, McAdam RA, Derbyshire KM, Mizrahi V (2000) Construction and application of mycobacterial reporter transposons. *Gene* 253: 67-75.
46. Lee LY, Gelvin SB (2008) T-DNA binary vectors and systems. *Plant Physiol* 146: 325-332.
47. Heeb S, Itoh Y, Nishijyo T, Schnider U, Keel C, et al. (2000) Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. *Mol Plant Microbe Interact* 13: 232-237.
48. Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative Transfer by the Virulence System of *Agrobacterium tumefaciens*. *Science* 256: 1324-1327.
49. Mumberg D, Muller R, Funk M (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156: 119-122.
50. Singh MV, Weil PA (2002) A method for plasmid purification directly from yeast. *Anal Biochem* 307: 13-17.
51. Mayan-Santos MD, Martinez-Robles ML, Hernandez P, Krimer D, Schvartzman JB (2007) DNA is more negatively supercoiled in bacterial plasmids than in minichromosomes isolated from budding yeast. *Electrophoresis* 28: 3845-3853.

52. Neuteboom LW, Lindhout BI, Saman IL, Hooykaas PJ, van der Zaal BJ (2006) Effects of different zinc finger transcription factors on genomic targets. *Biochem Biophys Res Commun* 339: 263-270.
53. Hohmann S (1991) PDC6, a weakly expressed pyruvate decarboxylase gene from yeast, is activated when fused spontaneously under the control of the PDC1 promoter. *Curr Genet* 20: 373-378.
54. Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11: 355-360.
55. Schneider CA, Rasband WS, Eliceiri KW NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671-675.



## Chapter 3

### **Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens* mediated transformation of *Saccharomyces cerevisiae*.**

3

pending for publication in Molecular Microbiology

Martijn Rolloos, Marius H.C. Dohmen, Paul J.J. Hooykaas and Bert J. van der Zaal

## Abstract

*Agrobacterium tumefaciens* cells carrying a tumor inducing plasmid (Ti-plasmid) can transfer a defined region of transfer DNA (T-DNA) as a single stranded intermediate, designated the T-strand, to plant cells as well as to other eukaryotic cells, like yeasts and other fungi. This process is known as *Agrobacterium* mediated transformation (AMT) that results in the incorporation of the T-DNA in the genomic DNA of the recipient cells. All available evidence indicates that T-strand transfer closely resembles conjugal DNA transfer as found between Gram-negative bacteria. However, where conjugal plasmid DNA transfer starts via relaxase-mediated processing of a single origin of transfer (*oriT*) and by default results in the presence of a circular copy of the plasmid in the recipient cells, the T-DNA is flanked by two imperfect direct border repeats which are both substrates for the Ti-plasmid encoded relaxase VirD2. Yeast was used as a model system to investigate the requirements of the recipient cell for the formation of T-DNA circles after AMT. It was found that, despite the absence of self-homology on the T-DNA, the homologous repair proteins Rad52 and Rad51 are involved in T-DNA circle formation. A model is presented involving the formation of T-DNA concatemers derived from T-strands by a process of strand-transfer catalyzed by VirD2. These concatemers are then resolved into T-DNA circles by homologous recombination in the recipient cell.

## Introduction

*Agrobacterium tumefaciens* causes crown gall disease in dicotyledonous plants by the transfer of a part from a 200 kb tumor inducing plasmid (Ti plasmid) to the plant genome. This transferrable region harbors genes that stimulate the production of phytohormones leading to the formation of the tumors that characterize crown gall disease [1, 2]. Within the Ti-plasmid, the T-DNA is flanked by two 25 bp imperfect direct repeats; the *left border* (LB) and the *right border* (RB). Outside of this transferrable region, the Ti plasmid contains a *vir* region containing the virulence genes that among other functions are required for the transfer of the T-DNA [3]. The combined activity of the virulence proteins VirD2 and VirD1 facilitate a nicking reaction of the parental Ti-plasmid at the RB and LB sequences [4]. VirD2 acts as a relaxase by inducing a nick between the twenty-second and twenty-third base of the RB sequence when orientated in the 5'-3' direction [5, 6]. After both border sequences are processed, a single stranded section of the T-DNA with VirD2 covalently attached to the 5' end is separated from the Ti-plasmid. With an *in vitro* experiment it has been shown that the nicking reaction catalyzed by VirD2 is reversible, meaning that VirD2 covalently bound to a T-strand, can act as a strand transferase when a suitable ssDNA substrate is provided thus reconstituting an intact border sequence [7]. In the models used at the present date, the nucleoprotein particle that results from the relaxase activity of VirD2 on the Ti plasmid, termed the T-strand, is targeted towards a type IV secretion system (T4SS) that facilitates the translocation of the T-strand to the recipient host cell [8].

In eukaryotes, the T-strands are transported into the nucleus and can get incorporated within the genomic DNA of the recipient cell as a processed T-DNA [9]. This process in its entirety, leading to genetically modified recipient cells, is known as *Agrobacterium* mediated transformation (AMT).

From early on, it was noted that both the processing and transfer of T-strands by *Agrobacterium* strongly resembles that of conjugal transfer of plasmid DNA in between prokaryotes, especially when comparing the pTi Virulence gene system with the conjugal system IncP [10, 11]. It thus became generally acknowledged that there must be a common ancestry for both DNA transfer mechanisms. Indeed, the RB and LB sequences present on the *Agrobacterium* Ti plasmid still share 6 to 7 nucleotides with the IncP-like *oriT* structure and the cleavage sites are placed at very similar positions [12]. Also the DNA sequence encoding VirD2 contains motives with strong homology to that of the IncP relaxase TraI [13]. The genetic organization of the *virB* operon that encompasses all the components of the T4SS of *Agrobacterium* and the pilus cluster genes of the IncP conjugal system show strong structural and functional homology [14, 15]. The homology between the molecular machinery behind conjugal transfer and VirB/VirD4 mediated DNA transfer is underlined by the discovery that conjugative transfer of plasmid DNA can be mediated by the genes that are part of the Ti plasmid *vir* region [16]. In fact, the *Agrobacterium* T4SS became the paradigm for transfer via conjugal pili. A difference between AMT and conjugal transfer of DNA concerns the fate of the transferred DNA. In the case of conjugation the transferred DNA is usually reformed into plasmids whereas in the case of AMT the transferred DNA integrates in the genome of the recipient cell as a linear molecule. Nevertheless, data has been collected suggesting that T-DNA may also circularize in recipient cells. Such plasmids were found after AMT of yeast if the T-DNA was provided with a 2 micron replication origin [17]. Recently, following AMT circularized T-DNA structures have also been recovered from plants, the natural host of *Agrobacterium* [18].

With AMT emerging as a means of genetic engineering of eukaryotic genomes, investigations regarding the fate of transferred DNA molecules within recipient cells have almost exclusively been inspired by the idea that T-DNA integration represents the default pathway of the process. Therefore, aspects dealing with a possible generation of plasmid-like molecules or T-DNA circles have been largely neglected, certainly in eukaryotic systems which lack portable elements functioning as origins for replicon maintenance. In yeast cells, however, such aspects can be addressed. Previous work using *Agrobacterium* cells harboring a binary vector equipped with a yeast 2 micron replication origin and an autotrophy selection marker in between T-DNA RB and LB sequences showed that plasmid DNA can be found after AMT of yeast cells. However, only two out of 48 plasmids represented the minimal yeast plasmids that were as expected the result from accurate border processing and circle formation by precise border fusion thereby forming a single RB-LB border fusion during circularization in the recipient cell. The remaining plasmids appeared to have been transferred completely, thus with skipping of LB processing [17].

Skipping of the LB would lead to transfer of a T-strand that could subsequently be reformed into the original plasmid in the recipient cell.

As mentioned above, the most widely used models for T-strand transfer assume that the VirD2 protein guides the strand through the T4SS, while being covalently bound to the 5'-end of a processed RB sequence [19, 20]. In fact, the models for relaxase-mediated conjugative plasmid transfer are derived from observations concerning AMT. However, it should be realized that the proposed mechanisms for both T-DNA transfer and conjugal plasmid transfer are not at all underpinned by molecular evidence regarding the entities that enter recipient cells. These could be single stranded molecules corresponding to single plasmid molecules, each with a relaxase at the 5' end, or alternatively, a single relaxase might guide long strands of repeating single stranded units through the T4SS when a rolling-circle replication-like mechanism feeds the translocation process. In the latter case, it is not clear how such a chain of concatemers is resolved into single units. Apart from that, while a double stranded plasmid molecule must be the final result of the conjugation process; there seem to be different means of triggering the synthesis of a complementary strand. Some plasmids might enable the transfer of primase activity to recipient cells like the RP4 [21, 22] and the IncQ plasmid RSF1010 [23], while others just seem to depend upon the fortuitous presence of such activities in the recipient. Since no virulence protein has been recovered from *Agrobacterium* displaying homology to any known primase, an assessment of the contribution of recipient cell functions will thus be essential.

In the present study, we developed a novel series of binary vectors with a varying number of border sequences. For deliberately mimicking conjugal complete vector transfer, a single RB sequence was assumed to function similar to a plasmid *oriT*. Once transformed yeast cells were obtained, T-strand derived plasmids were analyzed to determine the frequency of complete vector transfer and to confirm *Agrobacterium* mediated border processing. We provide evidence that the homologous repair (HR) pathway is the default pathway for the formation of circular T-DNA.

## Results

The construction of plasmid vectors allowing for quantification of partial and complete vector transfer

To learn more about T-DNA circle formation in recipient yeast cells, a new series of plasmid vectors was developed; the pOphis series (Figure 1). These vectors contained elements allowing for their replication in *E. coli* as well as in *Agrobacterium* but also for stable replication in yeast cells. The presence of one or more T-DNA border sequences allowed for their transfer to recipient cells by means of AMT. In order to be able to follow not only transfer of the T-DNA, but also vector backbone transfer, a selection marker was incorporated between the T-DNA border sequences but also outside of the borders

in the vector part. In order to mitigate undesired genomic integration of segments of the pOphis vectors into the yeast genome, homology to the yeast genome was omitted as much as possible.

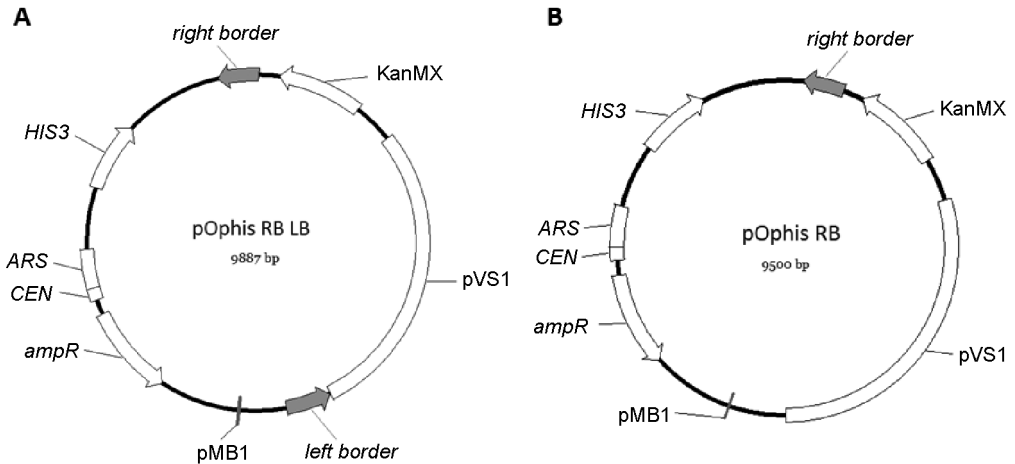


**Figure 1. Vector maps of the pOphis binary vectors.** (A) pOphis RB LB with the *rightborder* (RB) and *leftborder* (LB) sequence. Here both border sequences can be processed leading to the partial transfer of pOphis, while LB (or RB) skipping can lead to complete vector transfer (B) pOphis RB with only the RB sequence, expected to act as a single oriT-like sequence for complete vector transfer. See text for further details regarding plasmid properties.

Construction of the pOphis vectors started using the yeast shuttle vector pRS313 [24], which was altered to function as a binary vector by adding RB and LB sequences originally flanking the TL T-DNA of the well characterized octopine type Ti-plasmid pTi15955 [25]. The 25 bp RB possessed 322 bp of additional upstream DNA and the 25 bp LB was provided with 103 bp upstream TL DNA and 268 bp of downstream DNA, to avoid any loss of function that might occur when using minimal border sequences. Hence, pOphis vectors that contain the RB sequence also contain the OD sequence upstream of the RB core sequence. Each pOphis vector was further equipped with a replication origin derived from pVS1 [26, 27], allowing for replication in *Agrobacterium*. The high copy pMB1 (ColE1/pBR322) origin of replication from pRS313 was maintained allowing for replication in *E. coli*, as was the ARS/CEN sequence allowing for replication in yeast.

For yeast transformant screening, the pOphis vectors were equipped with a KanMX resistance marker gene driven by the *TEF1* promoter and the *HIS3* gene conferring histidine auxotrophy. The border sequences were positioned in such a way that the selection markers on the plasmid can be used to determine if complete or incomplete vector transfer has taken place. For example: if both RB and LB are processed, this will lead to the transfer of fragment A. If LB skipping takes place, as it is known to occur frequently [28], this is likely to result in the transfer of the complete vector AB (Figure 2).

In our experimental set-up the theoretical fragment B, consisting of the pVS1 replication origin and the KanMX marker gene, is not expected to be recovered from transformed yeast cells since it does not contain a yeast replication origin. The translocation of only fragment B would occur if LB processing takes place prior to RB processing as has been reported previously [29].

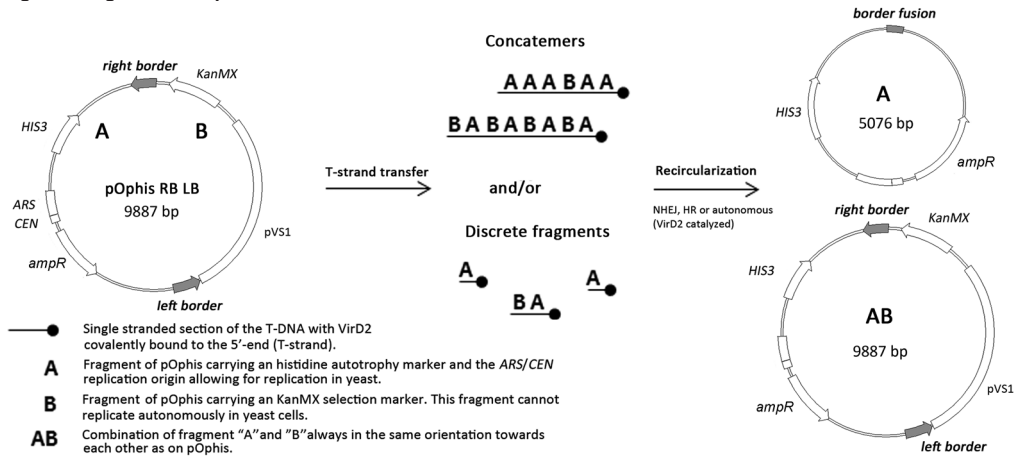


**Figure 2. Experimental set-up, illustrated for pOphis RB LB and its circular derivatives following the AMT of yeast cells.** T-strands produced in *Agrobacterium* will enter the yeast cell as separate units or, alternatively, as concatemers formed either by border fusion or by a rolling circle replication-like mechanism. After the transfer of a T-strand, complementary strand synthesis can convert the T-strand to a legitimate substrate for DNA repair pathways (see text). Alternatively, VirD2 can be involved in the formation of T-DNA circles from T-strands by catalyzing a strand-transfer reaction. Yeast cells with T-strand derived plasmids are first selected for histidine autotrophy. A successive selection step with G418 determines if complete or partial vector transfer has occurred. Border processing at both the RB and LB sequence will result in the transfer of fragments "A" or "B" depending on which of the border sequences is processed first. LB skipping or RB skipping would result in complete vector transfer (fragment "AB"). Note that fragment B cannot be recovered from transformed yeast cells since it does not contain a yeast replication origin. Consequences for other pOphis vectors, with single or additional borders are described in the text.

### T-DNA circle formation in recipient yeast cells is promoted by the homologous recombination pathway

In the first experiments, the plasmids pOphis RB LB and pOphis RB (Figure 1) were used as substrates for AMT. Compared to pOphis RB LB, its derivative pOphis RB contained only the RB sequence as substrate for VirD2 mediated border processing. By removing the LB sequence, the single RB should act in analogy to the single conjugative oriT present in mobilizable plasmids, thus exclusively mediating complete vector transfer (AB; see Figure 2). Because of the presence of ARS/CEN sequences, the T-strands that entered the recipient yeast cells could be stably maintained after double strand formation and circularization. Which factors are responsible for this, is currently unknown.

In order to find out whether any of the key DNA repair processes in yeast, homologous recombination (HR) and nonhomologous end-joining (NHEJ) are involved, transfer was studied not only to WT yeast but also to mutants defective in HR (*rad51Δ*, *rad52Δ*) or NHEJ (*ku70Δ*), respectively. When a single border sequence would trigger a rolling-circle replication-like mechanism in *Agrobacterium*, the yeast Rad52 dependent HR pathway may be required to resolve the entering T-strand concatemers into monomeric, circular units once double stranded structures have been formed. Hence, one might expect relatively low transformation efficiencies in *rad52Δ* or *rad51Δ* mutant strains with such single border vectors. At the same time, a role of Ku70, essential for DNA repair by non-homologous end joining (NHEJ), might play a role in 5' end to 3' end joining of smaller units, for instance for forming circles from double stranded "A" fragments. For transformant screening, the yeast cells were plated on a synthetic defined (SD) yeast medium without histidine, allowing for the growth of yeast cells containing partial as well as complete vector molecules (A and AB, respectively; see Figure 2).

As can be seen in Figure 3, removal of the LB sequence from pOphis RB LB led to an 83% drop in the average transformation efficiency. This suggests that in the absence of the LB sequence the formation of stable T-DNA circles becomes inefficient. Also the transformation of the *rad52Δ* and *ku70Δ* yeast mutant strains with pOphis RB instead of pOphis RB LB resulted into a similar reduction in AMT efficiency of 85% compared to pOphis RB LB. The observed AMT efficiency was strongly dependent on the yeast genotype (Figure 3). Cocultivations with YPH250 *rad52Δ* resulted in about 80% reduction of the average transformation efficiency compared to the WT YPH250 strain. Cocultivation of YPH250 *rad51Δ* strain similarly led to a 75% reduction of the average transformation efficiency compared to WT. This indicates an important role of the HR pathway in AMT of this pOphis vector. T-DNA transfer to the *ku70Δ* yeast strain occurred with a frequency that was not significantly different from that to the WT suggesting that NHEJ does not play an important role in T-DNA circularization.

The importance of the HR pathway for T-DNA circle formation was also observed for translocated pOphis RB, leading to an 85% reduction of AMT efficiency when using the *rad52Δ* strain. With pOphis RB transferred to the *ku70Δ* strain the AMT efficiency was not significantly reduced compared to WT. Multimers of the pOphis vectors present in yeast cells would be unstable since they would harbor multiple *CEN* encoding sequences. The obtained data suggests that such multimers are formed during AMT, with both pOphis RB LB and pOphis RB, and that the HR pathway of the recipient yeast cells is involved in the resolution of such concatemeric T-DNA into stable monomeric T-DNA circles utilizing the self-homology of T-DNA concatemers as a basis for HR.

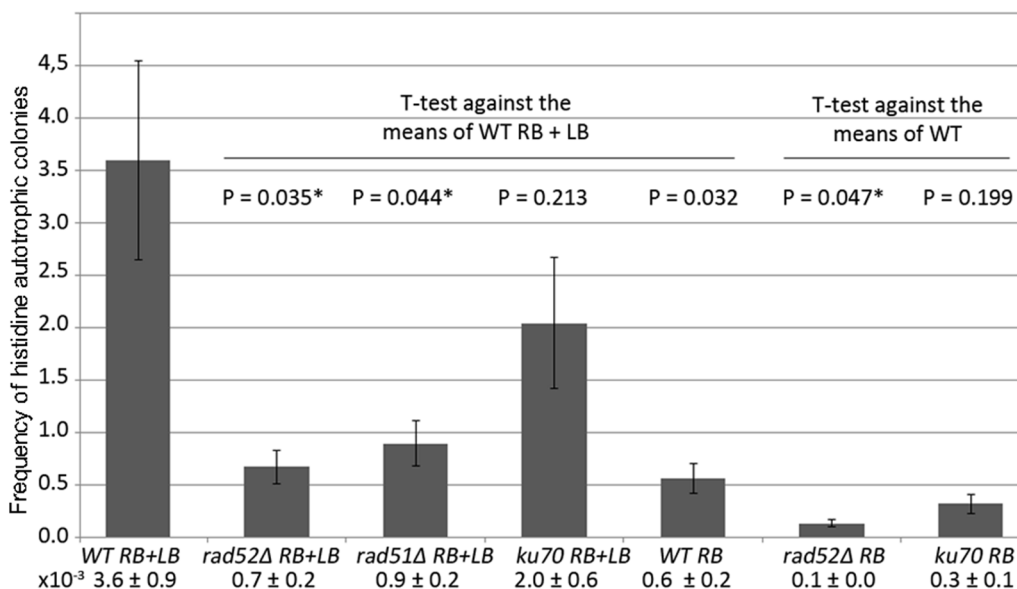


**Figure 3. The influence of the yeast DNA repair pathways on T-DNA circularization.** Yeast strain YPH250 and its derivatives, *rad52Δ*, *rad51Δ* and *ku70Δ* were cocultivated with LBA1100 harboring pOphis RB LB (RB + LB) or LBA1100 harboring pOphis RB (RB). As a negative control the cocultivations were also performed without inducing virulence gene transcription with acetosyringone. The averages of five independent experiments are shown. No transformants were found without induction. The averages of five independent experiments are shown, with the exception of *rad52Δ* RB which shows the average of four independent replications. The AMT frequencies were calculated by dividing the total number of histidine autotrophic colonies by an estimate of the total number of yeast cells. Error bars represent the SEM. The p-values were calculated performing a two-tailed student's T-test. An asterisk is used to indicate statistical significance with  $\alpha$  set to 0.05.

### Stability of pOphis RB LB

Within the pOphis RB LB plasmid, only the 25 bp core sequences of the LB and RB share sequence homology. Therefore, removal of fragment B by an intermolecular HR event cannot be excluded at this point. The frequency by which such an event could occur could be determined by cloning an *URA3* gene downstream to the LB sequence to function as a positive selection marker for the loss of fragment B. YPH250 transformed with pOphis RB LB *URA3* was thereafter cultivated on plates, exactly mimicking the cocultivation procedure. The average frequency of Ura auxotrophy-loss was found to be  $3.8 \times 10^{-6}$  with  $n = 20$  with a SEM of 1.27.

### Varying the border repeat constitution of the vectors

To get more insight into the negative effect of the removal of the LB from pOphis RB LB on the transformation efficiency, the pOphis series was further extended. An additional RB was cloned directly upstream the LB sequence of pOphis RB LB to obtain

pOphis RB RBLB (Figure 4A) which is expected to be less prone to LB skipping. To determine if it has to be a LB sequence complementing the RB sequence to reach transformation efficiencies comparable to pOphis RB LB, pOphis RB RB was constructed (Figure 4B). In addition to this, pOphis LB LB was constructed that lacks the OD sequence that is associated with the RB sequence present in the aforementioned pOphis vectors. A convenient negative control was put into use in the form of pOphis Borderless (Figure 4C), this plasmid does not contain any VirD2 recognition sequences, and thus cannot be processed by VirD2 and because of that, is not expected to undergo T4SS mediated transfer.

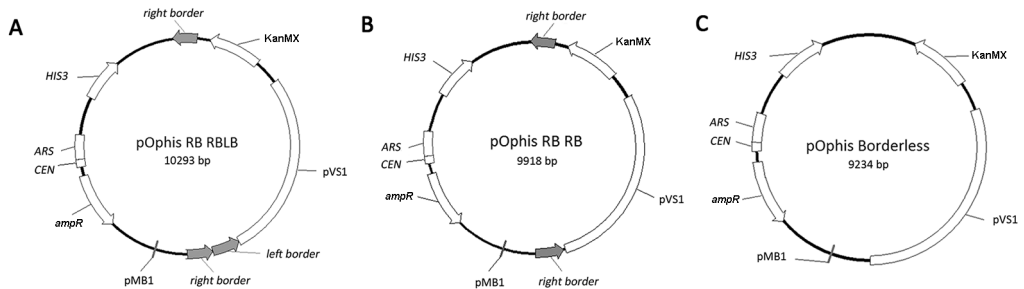


**Figure 4. Vector maps of the extended pOphis binary vector series.** (A) pOphis with RB and RBLB in a tandem configuration. This binary vector is comparable with pOphis RB LB but should be less prone to LB skipping. (B) pOphis with two RB sequences. The additional RB sequence replaces the LB sequence as is present on pOphis RB LB (C) Borderless pOphis, used as a negative control during the AMT of yeast.

A comparison of the AMT frequencies resulting from the usage of different border repeat constitutions indicate that enhancement of border processing at the site of the LB results in a higher AMT frequency. Figure 5 shows that transforming WT yeast with pOphis RB RBLB led to an approximate 2-fold increase in transformation efficiency compared to pOphis RB RB or pOphis RB LB. No significant difference was found between the transformation efficiencies that could be reached using pOphis RB LB and pOphis RB RB (Figure 5). AMT of WT yeast with pOphis LB LB, resulted in a very low transformation efficiency ( $1.1 \times 10^{-8}$ ,  $n = 4$ , SEM =  $0.2 \times 10^{-8}$ ), about 75 fold less than AMT of WT yeast with pOphis RB LB. No transformed yeast cells were observed when using pOphis Borderless.

A very similar relative decline in AMT efficiency was observed for each individual plasmid when using the *rad52Δ* yeast strain as a recipient instead of the WT strain. As mentioned above, AMT of the *rad52Δ* yeast strain with pOphis RB LB and pOphis RB showed a reduction in transformation frequency of about 80%. Similarly, AMT with pOphis RB RBLB showed a reduction of about 70% (Figure 5). The yeast HR pathway is thus needed for T-DNA circularization in yeast cells irrespective of whether the transferred T-DNA consisted predominantly of complete vector molecules (fragment AB, Figure 2) or T-DNA segments (fragment A, Figure 2). When using the *ku70Δ* yeast strain as a plasmid recipient, the differences in transformation efficiency compared to WT were not significant (Figures 3 and 5).

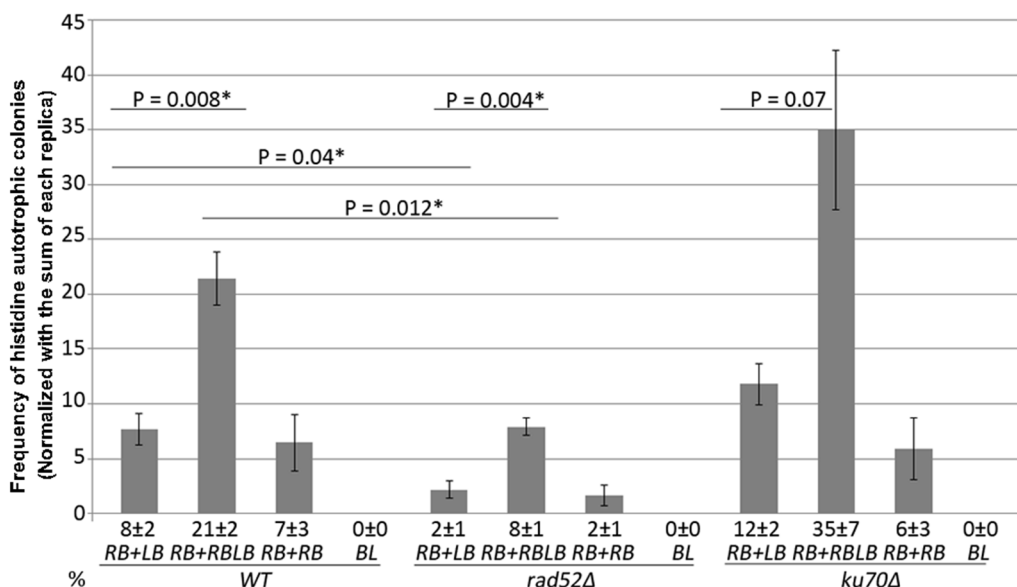


**Figure 5. Effect of border repeat constitution upon AMT efficiency in DNA repair mutant yeast strains.** YPH250 WT, *rad52Δ* and *ku70Δ* were cocultivated LBA1100 harboring pOphis RB LB (RB + LB), pOphis RB RBLB (RB+RBLB) and pOphis RB RB (RB + RB). LBA1100 harboring pOphis Borderless (Borderless) was used as a negative control for T4SS mediated T-DNA transfer. The averages of four independent experiments are shown. The transformation efficiency of each treatment was normalized with the sum of the transformation efficiencies of all treatments of the corresponding replica. Thereafter, the resulting relative transformation efficiencies of all four replications were averaged. The normalization step to obtain relative transformation efficiencies was introduced to remove the effect of varying overall transformation efficiencies between experimental replicas while maintaining the differences in the relative transformation efficiency between treatments. The data presented here represents a total of about 40000 histidine autotrophic yeast mutants. Error bars represent the SEM. The p-values were calculated performing a two-tailed student's T-test. An asterisk indicates statistical significance with a set to 0.05.

### Analysis of the T-DNA circles derived from transformed yeast cells

As already explained above, also in the legends to Figure 1, the vector backbone part, fragment B of the plasmids is functionally dispensable for maintenance in yeast cells and can also not be maintained as an independently replicating unit. Therefore, it should be possible to assess whether T-strands predominantly represented the complete binary vector (plasmid AB) or only a portion thereof (plasmid A).

Random samples of fifty transformed colonies per treatment were assayed for G418 resistance for each replica (Figure 6). G418 resistance is indicative for the presence of part B and *vice versa*. WT and *ku70Δ* strains receiving pOphis RB by AMT invariably combined histidine autotrophy with G418 resistance corroborating transfer of the complete single border plasmid. When using the *rad52Δ* yeast strain as a recipient this was also mostly the case however, 6% of the analyzed cells harbored incomplete pOphis RB molecules, lacking the KanMX gene.

For the pOphis plasmids with two or more borders, pOphis RB LB, pOphis RB RB, pOphis RB RBLB, significant fractions of the transformed yeast colonies were G418 sensitive, indicative for transfer of the T-DNA segment instead of the whole vector during AMT. For pOphis RB LB, a modest 30% of the T-DNA circles in WT yeast were sensitive for G418, indicating a high frequency of border skipping. A similar result was obtained with pOphis RB RB (Figure 6) indicating that the RB could act with a similar efficiency as a LB in delimiting T-DNA transfer and the avoidance of complete vector transfer. For pOphis RB RBLB, about 74% of the transformants were sensitive to G418. This indicates that the presence of an extra RB border repeat next to the LB strongly increases border processing thereby significantly reducing complete vector transfer. As mentioned above (Figure 5), the pOphis RB RBLB plasmid also exhibited the highest AMT frequency. With the *rad52Δ* and *ku70Δ* mutants as recipients also the usage of pOphis RB RBLB resulted in the highest AMT frequencies compared to the other pOphis vectors. However, the percentage of G418 sensitivity was significantly different in the mutant yeast strains transformed with pOphis RB LB: higher in *rad52Δ* (about 56%) and lower than in *ku70Δ* (about 13%) (Figure 6).

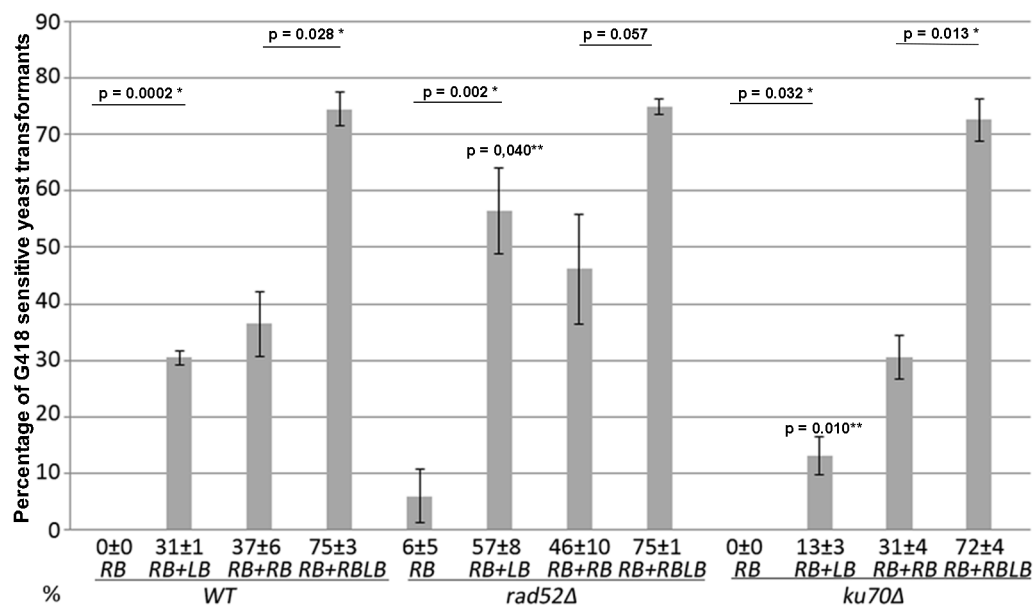


**Figure 6. Percentage of His<sup>+</sup> yeast transformants conferring G418 sensitivity.** WT, *rad52Δ* and *ku70Δ* YPH250 strains were cocultivated with LBA1100 harboring pOphis RB (RB), pOphis RB LB (RB+LB), pOphis RB RB (RB + RB) and pOphis RB RBLB (RB+RBLB). Each bar shows the average ratio of partial vector transfer of four independent cocultivations from each of which 50 histidine autotrophic colonies were assayed for G418 sensitivity. To calculate the ratio of partial vector transfer for each replica, the number of yeast cells sensitive to G418 was divided by 50. Error bars represent the SEM. The p-values were calculated performing a two-tailed student's T-test. A single asterisk indicates statistical significance with  $\alpha$  set to 0.05 between the bars connected with a horizontal stripe. Two asterisks indicate statistical significance with  $\alpha$  set to 0.05 when comparing the marked bar with the WT yeast strain transformed with *Agrobacterium* harboring pOphis RB LB.

From a number of yeast colonies transformed with pOphis RB LB the plasmids representing T-DNA circles were rescued and prepared for nucleotide sequencing of the RB region. From WT, *rad51Δ* and *rad52Δ* yeast strains four colonies were analyzed; two from yeast colonies sensitive to G418 and another two from yeast colonies resistant to G418. The transformed yeast strains resistant to G418, which were hence expected to contain the complete pOphis RB LB, showed perfect RB sequences, identical to the original plasmid (Figure 7, lower panel). In five out of six cases, the plasmids isolated from G418 sensitive transformed yeast strain showed RB-LB fusions as expected to result from RB and LB processing and fusion (Figure 7, upper panel). In one of the rescued T-DNA circles isolated from the *rad51Δ* yeast strain, the border sequences were lost due to a deletion of 402 bp plasmid DNA; the sequence shown was followed by a remainder of fragment A, thus with fragment B lacking completely.

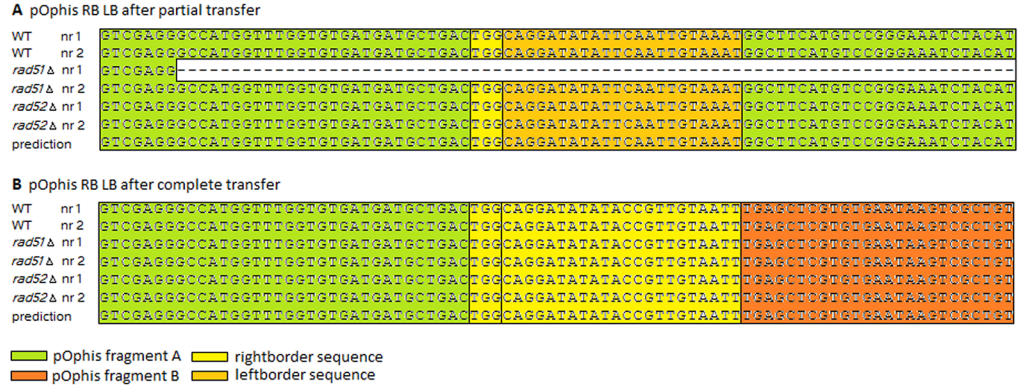


**Figure 7. Alignment of border sequences present on T-DNA circles rescued from different yeast backgrounds.** Sequence data of T-DNA circles rescued from WT, *rad51Δ* and *rad52Δ* yeast strains after AMT with LBA1100 harboring pOphis RB LB. **(A)** represents sequences derived from apparently partially transferred pOphis RB LB, rendering the transformed yeast cells autotrophic for histidine but sensitive to G418. **(B)** represents the RB area from apparently completely transferred pOphis RB LB, rendering the transformed yeast cells autotrophic for histidine and resistant to G418. For each yeast strain, two rescued plasmids were analyzed.

Additional border regions of T-DNA circles rescued from WT YPH250 were sequenced, representing randomly picked representatives of pOphis RB RB, pOphis RB RBLB and pOphis LB LB (Figure 8). In all cases, the sequences were as expected from the phenotype of the yeast cells harbouring the plasmids. None of them exhibited border sequences with filler DNA or with DNA deletions. Interestingly, the sequences derived from plasmids rescued after AMT with pOphis RB RBLB seemed to reflect that after an initial nicking reaction at one of the RB sequences, further processing can occur at either the LB sequence or at the secondary RB sequence, directly upstream of the LB sequence. The results of both scenarios are represented in Figure 8.

|               |           |     |                                          |                 |
|---------------|-----------|-----|------------------------------------------|-----------------|
| pOphis RBRB   | G418 res  | nr1 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRB   | G418 res  | nr2 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRB   | G418 res  | nr3 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRB   | G418 sens | nr1 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRB   | G418 sens | nr2 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRB   | G418 sens | nr3 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 res  | nr1 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 res  | nr2 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 res  | nr3 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 sens | nr1 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 sens | nr2 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis RBRBLB | G418 sens | nr3 | TGTGATGATGCTGACTGGCAGGATATATACCGTTGTAATT | TGAGCTCGTGTGAAT |
| pOphis LBLB   | G418 res  | nr1 | CCAAAGCGGCAGCGGCGCAGGATATATTCGAATTGTAATT | GGCTTCATGTCCGGG |
| pOphis LBLB   | G418 res  | nr2 | CCAAAGCGGCAGCGGCGCAGGATATATTCGAATTGTAATT | GGCTTCATGTCCGGG |
| pOphis LBLB   | G418 sens | nr1 | CCAAAGCGGCAGCGGCGCAGGATATATTCGAATTGTAATT | GGCTTCATGTCCGGG |
| pOphis LBLB   | G418 sens | nr2 | CCAAAGCGGCAGCGGCGCAGGATATATTCGAATTGTAATT | GGCTTCATGTCCGGG |
| pOphis LBLB   | G418 sens | nr3 | CCAAAGCGGCAGCGGCGCAGGATATATTCGAATTGTAATT | GGCTTCATGTCCGGG |

pOphis fragment A   
  rightborder sequence  
 pOphis fragment B   
  leftborder sequence

**Figure 8. Alignment of border sequences present on T-DNA circles derived from different pOphis vectors.** T-DNA circles rescued from WT YPH250 after AMT with LBA1100 harboring pOphis RB RB, pOphis RB RBLB and pOphis LB LB. Prior to the isolation of the T-DNA circles, the sensitivity to G418 of the transformed yeast strains was assayed. For each type of pOphis vector, five or six rescued plasmids were analyzed varying in G418 sensitivity status. The color code used for the right part of this figure indicating if the border sequence is followed by pOphis fragment A or pOphis fragment B is based upon sequence information downstream the DNA sequence depicted here.

## Discussion

In this study, we investigated two possible key variables for having an impact on the formation of circular T-DNA structures in recipient eukaryotic cells after VirD2 mediated T-strand transfer; the presence and type of T-DNA border sequences and the genotype of the recipient cells. Our data gathered using the yeast model system for AMT provides compelling evidence that AMT frequency as well as the type of T-DNA circles formed depend on both variables. As will be further discussed below, the available evidence can be used to construct a model where T-strands enter recipient yeast cells preferably in the form of concatemers that are subsequently resolved by the Rad52 dependent homologous recombination machinery. The formation of these concatemers is likely due to strand-transferase activity of VirD2. Additionally, Rad52 could be involved in the formation of T-strands by increasing the mobility of the ends of the T-strand, thereby facilitating the chaining of T-strands or the direct circularization of monomeric T-strands [30]. By analogy, this model is likely to be applicable to conjugative plasmid transfer, further unifying the processes of AMT and conjugation.

Earlier research has shown that for the genomic integration of T-DNA in yeast cells the host DNA repair pathways are necessary [31, 32]. Considering the fact that AMT and conjugative plasmid transfer via a T4SS appear to be mechanistically similar [16, 33-35], we wanted to perform a systematic analysis to see whether yeast DNA repair pathways would also play a role when T-strands would be allowed to maintain themselves as circular plasmid structures.

We deliberately chose to equip the newly constructed series of binary vectors with an *ARS/CEN* origin of replication for yeast cells as the resulting chromosomal type of plasmid replication maintenance ensures that only one copy is present. Hence, once a post AMT circular structure is formed, this should be the founder of all plasmids copies present in a colony of transformed yeast cells, thus facilitating plasmid analysis. In comparison, should we have used 2 micron plasmid type of replication, post-AMT plasmid analysis could easily be biased; not simply for the reason of getting more plasmids per cell, but predominantly since 2 micron plasmids are known to utilize homologous repair to increase their copy number via a concatemeric intermediate [36], thus already implicating a role for the HR pathway.

The experiments performed with combinations of the T-DNA border containing plasmids pOphis RB LB and pOphis RB and the different yeast genotypes clearly showed that the HR proteins Rad52 and Rad51 must be involved in the formation of T-DNA circles, but not the NHEJ protein Ku70 (Figure 3). This result is remarkable since the pOphis vectors used here do not contain any regions of self-homology. *ARS/CEN* plasmid replication itself is not known to depend upon HR and in our hands *rad52Δ* and *rad51Δ* strains maintaining pOphis plasmids grew just as well on selective medium as the *ku70Δ* strain. Therefore, the observation that a HR pathway deficient yeast strain results in a 75% reduction of AMT derived transformed colonies compared to WT yeast cells strongly indicates that the generation of the first replicating plasmid after T-strand transfer is very much aided by HR.

Although hardly any self-homology is present in pOphis RB LB and pOphis RB, one can explain the HR dependency by assuming that repetitive structures are generated during or just after AMT. Thus either concatemeric T-strand intermediates must enter yeast cells or they must be formed in yeast cells out of monomeric T-strands. Be it without clear experimental proof, most models for conjugative plasmid transfer propose that a rolling circle replication-like mechanism in the donor strain feeds the T4SS with single stranded DNA [10, 37-39]. Supposing that relaxase proteins are unable to cleave each single stranded *oriT* sequence that is generated, entry of single stranded concatemers of plasmid units can be a natural consequence of such a rolling circle mechanism [39].

For pOphis RB, with its single *oriT*-like RB sequence, a similar model can explain how VirD2-mediated transfer of T-strands results in the entry of single stranded plasmid concatemers into recipient cells. For pOphis RB LB, border skipping or the joining of cleaved border sequences either before or after passage of the T4SS can be envisaged as yielding repetitive units. Border skipping has been described earlier as a source for complete vector transfer from *Agrobacterium* to plant cells [40, 41]. In any case, in order to become substrates for HR in yeast recipient cells more than a single unit of plasmid or T-strand length needs to be converted into a double stranded DNA molecule. It might very well be that the *ARS* sequence present in T-strands, concatemeric or not, facilitates double stranded DNA formation due to yeast primase-mediated complementary RNA primer formation at this sequence.

Concerning VirD2-mediated T-strand or plasmid transfer, AMT with pOphis RB LB resulted in about 6-fold more transformed yeast colonies compared to pOphis RB (Figure 3). At first sight, it is tempting to postulate that the generation of longer units as required for pOphis RB transfer is more demanding for the system. However, one would then expect that the pOphis RB LB derived plasmids found in WT yeast cells would nearly all be of the processed smaller type instead of just about 30% (Figure 6). A plausible explanation for these observations might be that the T-strands resulting from pOphis RB, which are expected to encompass the pOphis vector in its entirety, including the pVS1 replication origin, are prone to self-closure and complementary strand synthesis already in the *Agrobacterium* cells thereby preventing their own translocation. If indeed a T-strand encompassing the complete binary vector is prone to complementary strand synthesis and self-closure, then stimulating the termination of T-strand formation should lead to elevated transformation levels. Indeed, the addition of a RB sequence upstream the LB sequence of pOphis RB LB led to a significant increase in transformation efficiency compared to pOphis RB LB (Figure 5).

When further investigating plasmid transfer, in particular for plasmids with more than a single border sequence, it was evident that AMT of all plasmids, thus pOphis RB, pOphis RB LB, pOphis RB RB, and pOphis RB RBLB, was about 70-75% reduced when using *rad52Δ* recipient cells. For the *ku70Δ* genotype, the transformation efficiencies were not significantly different from those observed for the WT. Hence, it can be concluded that in our experiments the HR mechanism in yeast cells contributed to the establishment of about 70-75% of plasmid molecules after T4SS mediated T-strand transfer. The contribution of the NHEJ pathway seems to be negligible. The limited amount of plasmids that are still established in the *rad52Δ* genotype could very well be the result of circularization of monomeric T-strands as free 3'OH groups are able to perform a reversed transesterification reaction, catalyzed by the VirD2 moiety that is covalently bound to the incoming T-strands or via the NHEJ pathway after the synthesis of a complementary strand.

Apart from a few plasmids recovered from *rad52Δ* cells, single bordered pOphis RB was always found as a complete vector (Figure 6), apparently resulting from perfect RB processing events in *Agrobacterium* and perfect post transfer RB rejoining in yeast cells. However, as mentioned above, our data indicates that the most prevalent mechanism for generating plasmids in yeast involves HR-mediated resolution of T-strand concatemers. For pOphis RB, this mechanism will lead to the appearance of complete plasmids in yeast cells. The exceptions found in the *rad52Δ* strain, the few percent of plasmids lacking the KanMX marker, most likely reflect rare circularization events of single truncated T-strands.

For the multiple plasmids with two or three border sequences, significant fractions of the plasmid populations found in yeast lacked the KanMX marker gene, indicating that more than just a single border sequence was processed (Figure 6). For pOphis RB RBLB, a majority of about 75% the plasmids found in the different yeast strains lacked the G418

marker. Apparently, the presence of an extra LB introduction, comparing to pOphis RB RB, strongly favors processing in *Agrobacterium*. This complies with other experimental evidence that the addition of LB sequences aids in the termination of T-strand formation [28].

In our experiments we observed a correlation with the extent of apparent plasmid processing and AMT frequency. As indicated above, this might be a consequence of plasmid formation from T-strands in *Agrobacterium*. Additionally, we found that the HR pathway strongly contributes to AMT efficiency for all different plasmids indicative of the presence of concatemeric T-strands in yeast which are subsequently resolved into monomeric units. Since the HR pathway had a positive impact on the formation of T-DNA circles derived from all different binary vectors (Figure 3 and 5), irrespective if these binary vectors are by preference completely transferred or partially (Figure 6), a conclusion can be drawn about the mechanism behind the predicted concatemer formation. When more than one border sequence within a single plasmid is subjected to border processing in *Agrobacterium*, formation of a chain of concatemers via a rolling circle-like mechanism involving the whole plasmid will be prohibited. Hence, in order to explain the positive role of the HR pathway during AMT of yeast, we have to conclude that formation of T-strand concatemers must occur by joining of T-strands, either prior or after transfer via the T4SS.

For maintenance in *Agrobacterium*, pOphis plasmids depend upon the pVS1 origin of replication. Although considered as a low copy origin ([www.cambia.org](http://www.cambia.org)), there will be more copies than just the one or two found for the natural Ti-plasmid. Therefore, it can be envisaged that the numbers of relaxosomes formed on pOphis plasmids outnumber the T4SS pili in *Agrobacterium*. In that manner, plasmid derived T-strands would be formed without a nearby connection to a pilus structure, giving the T-strands ample opportunity to be covalently linked by VirD2 proteins when lingering in the bacterial cytosol; the transesterification reaction performed by VirD2 is easily reversible [36]. When VirD2-mediated end joining results in a covalently closed circle, such a structure won't be transported through the T4SS, but all linear strands that are assembled with VirD2 at the 5' end are in principle substrates for transport. Considering the contribution of the yeast HR mechanism, a large fraction of the transported T-strands that enter the yeast cells must be concatemers of the various T-strands formed in *Agrobacterium*. As soon as the incoming T-strand consists of more than a single copy of a pOphis derived sequence and becomes double stranded, the yeast HR mechanism can resolve the multimeric DNA sequence forming a T-DNA circle (Figure 2).

Sequence analysis of the RB area found in randomly picked plasmids (Figures 7 and 8) verified the phenotype of yeast colonies harboring these plasmids, thus showing the combination of plasmid fragments A and B in case of G418 resistance and only fragment A when lacking it. In almost all cases perfect border fusions were recovered, as if joined perfectly by VirD2. It is therefore highly unlikely that an NHEJ reaction mediated by Ku70 plays an important role in plasmid formation from T-strands in yeast.

Such a mechanism would require that both ends of the T-strand, being the terminal border remnants, are preserved in a double strand structure prior to end joining. Due to the lagging strand problem during DNA replication, double stranded DNA cannot be formed at the original 3' end of the T-strand, resulting in loss of information and incapability of regenerating a perfectly fused border sequence. The fact that the *ku70Δ* genotype had no significant impact on the transformation efficiency when compared with WT agrees with its redundancy regarding the joining of T-strands. When T-DNA circles are formed in yeast independent of an HR mechanism, it indeed seems most likely that transferred VirD2 catalyzes the joining of processed border ends, thereby regenerating perfect border cores from single stranded molecules. As already mentioned above, subsequent double strand formation can easily be initiated via primase activity on the ARS sequence present in the pOphi plasmids.

Provided that *Agrobacterium* VirD2 is involved in the formation of T-DNA circles, Rad52 could, in addition to the formation of T-DNA circles by resolving concatemeric T-strands, play a dual role in T-DNA circle formation. It was recently found that the association of Rad52 with an irreparable double-strand break (DSB) enhances the mobility of DBSs. A process for which the DNA repair protein Rad51 is required [30]. In the context of the results described here, this could indicate that Rad52 facilitates the chaining of T-strands or direct T-DNA circle formation catalyzed by VirD2 by mobilizing the T-strand-ends during homology search. Any concatemeric T-strand structures formed by the catalytic strand-transfer activity of VirD2, either in *Agrobacterium* or the recipient cell, could be resolved by the HR pathway mediated by Rad52 after a complementary DNA strand is synthesized.

When combining all available data, we get to a model as depicted in Figure 9: tandem repeats of T-strands are formed by a chaining reaction catalyzed by VirD2, either in the *Agrobacterium* cell (Figure 9A) or in the recipient yeast cell, that are subsequently resolved by the HR pathway of yeast into T-DNA circles after the synthesis of a complementary DNA strand (Figure 9B). There is already some data available supporting this model: integrated T-DNAs multimers have been recovered from the yeast strain *Kluyveromyces lactis* after AMT[42]. In addition to this, Rad52 could enhance the mobility of the T-strand-ends in the recipient cell, thereby facilitating VirD2 in catalyzing the strand-transfer reaction leading to the formation of T-DNA circles (Figure 9C), much like a snake biting its own tail. Also here, the fusion of the vector ends takes place prior to the synthesis of a complementary DNA strand.

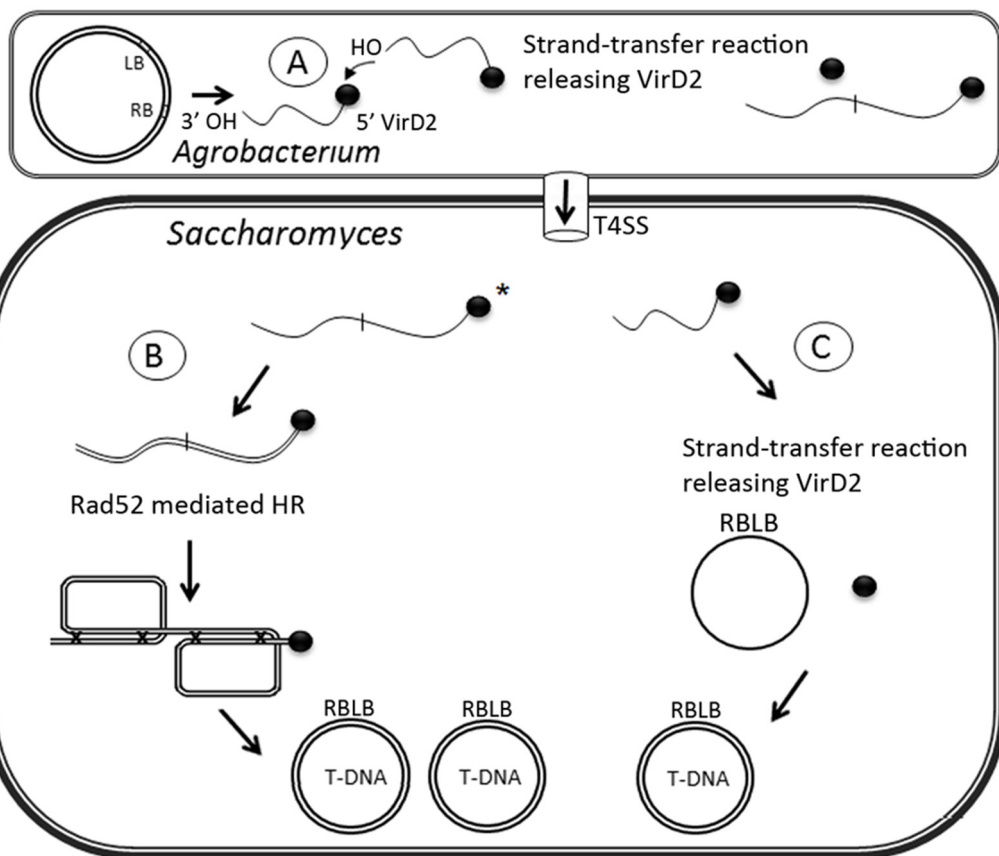


**Figure 9. Modeling the involvement of Rad52 in the formation of T-DNA circles.** (A) T-strand concatemers\* are formed in *Agrobacterium* or in the recipient yeast cell by the catalytic strand-transfer activity of VirD2. (B) After a complementary DNA strand is synthesized in the recipient cell, the T-strand concatemer is a legitimate substrate for the HR pathway which resolves the concatemeric T-strands into T-DNA circles. (C) VirD2 mediated T-DNA circularization of singular T-strands. Rad52 could facilitate the latter process by increasing the mobility of the free 3'-end of the T-strand.

Applying this model to the natural situation, thus with a large Ti-plasmid containing a relatively short T-region, predicts that the concatemers formed will predominantly be repeats representing the T-region; a large strand representing the Ti-plasmid backbone will be relatively rare since after LB skipping such an extended length of plasmid DNA must be unwound. When resolved in the plant cell, by any means of HR, it can be expected that monomeric T-DNA circles encompassing a processed T-DNA will be formed.

T-DNA circles derived from binary vectors have been isolated from plants infected with *Agrobacterium* [18]. It would be very interesting to know if T-DNA circles are also formed in *Agrobacterium* infected plant cells when the T-strands are derived from a Ti-plasmid. The formation of T-DNA circles and other double stranded T-DNA intermediates could be beneficial for *Agrobacterium* since the transformed plant cells can already start producing opines and tumor generating phytohormones from these

T-DNA circles even before the genomic integration of a T-DNA has taken place. Circular dsDNA molecules can be expected to be more stable in recipient cells since they form no substrate for exonucleases in contrast to linear dsDNA. For biotechnological purposes, the formation of T-DNA circles that do not integrate into genomic DNA could offer a means of transiently expressing proteins in plant cells.

## Materials and methods

### Used strains

All cocultivation experiments were performed with the *Saccharomyces cerevisiae* strain YPH250 (*MATa ade2-101 his3-Δ200 leu2-Δ1 lys2-801 trp1-Δ1 ura3-52*) described in [24]. The isogenic *rad52Δ*, *rad51Δ* and *ku70Δ* mutants were constructed earlier by Haico van Atticum [32]. The *Agrobacterium tumefaciens* strain LBA1100 carrying the disarmed octopine Ti-plasmid pTiB6 (pAL1100  $\Delta T_L$ ,  $\Delta T_R$ ,  $\Delta tra$ ,  $\Delta occ$  conferring rifampicin and spectinomycin resistance) was used for all described experiments [16].

### Cocultivation experiments

For the AMT of yeast cells the cocultivation procedure was followed as described by Bundock *et al* [17] with some slight modifications. All cocultivations were induced for a week at 21°C. *Agrobacterium* cells were always precultured on LB medium with appropriate antibiotic selection. Before mixing the *Agrobacterium* cell suspension with the yeast cells the *Agrobacterium* cells were washed with induction medium (IM) without glucose. This step lowers the glucose content of the cell suspension to match that of the IM used for the induction plates.

### Cloning strategy

For the construction of the pOphis series a RB and LB sequence were synthesized identical to the TLlb and TLrb border sequences as are present on the octopine type Ti-plasmid pTi15955 (NCBI reference: NC\_002377.1). The synthesized DNA fragment, stored in pUC57, was designed to contain the following: RB with 322 bp of upstream TL DNA which includes the *overdrive* (OD) sequence, a small spacer sequence consisting of an *NcoI*, *SmaI*, *EcoRI*, *AatII* and *XbaI* restriction sites, LB with 103 bp of upstream TL DNA and 268 bp of downstream TL DNA. Note that the RB and LB sequence are in the same orientation. To retrieve the border sequences from pUC57, two *XhoI* sites directly flanking the border sequences were used resulting in a 793 bp fragment containing both border fragments. The yeast shuttle vector pRS313 [24] was digested with *EcoRV* and *SacI* and religated after a blunting reaction with T4 DNA polymerase. The border sequences were cloned into the resulting vector digested with *XhoI*.

The pVS1 locus was obtained performing a PCR on column purified pCAMBIA. For this the following primers were used:

pVS1 *Bam*HI FW 5'- ATTAGGATCCTTCATTACCGAAGAGATCCG-3'

pVS1 *Sac*II RV 5'- ATTACCGCGGTCCGCTTACAGACAAGCTG-3'

The resulting PCR product was digested with *Bam*HI and *Sac*II and cloned into the likewise digested vector. The KanMX ORF with TEF promoter and terminator sequence was obtained performing a PCR on column purified pSDM8000. For this the following primers were used:

KanMX *Nar*I FW 5'-TAATGGCGCCTTTAGCTTGCCTCGTCCC-3'

KanMX *Sac*II RV 5'-TAAACCGCGGAATCGACAGCAGTATAGCG-3'

The resulting PCR product was digested with *Nar*I and *Sac*II and cloned into the likewise digested vector. The vector resulting from this ligation reaction was renamed into pOphis RB LB. The other pOphis type vectors are all derivatives of pOphis RB LB. Firstly the LB sequence was removed resulting in pOphis RB by digesting pOphis RB LB with *Xba*I and *Spe*I and successively perform a self-ligation reaction with the linearized plasmid. To remove the RB sequence from pOphis RB leading to pOphis Borderless, column purified pOphis RB was used as a template for PCR using the following primers:

RB knock-out FW 5'-ATTAGTCGACGCATCATCACACCAAACC-3'

RB knock-out RV 5'-ATTAGTCGACGTGTGAATAAGTCGCTG-3'

The amplicon was digested with *Sal*I and successively used for a self-ligation reaction. To remove the RB sequence from pOphis RB LB leading to pOphis LB, column purified pOphis RB LB was used as a template for PCR using the following primers:

RB knock-out *Nco*I FW 5'-TGTATCCGTGGCATCCATGG-3

RB knock-out *Nco*I RV 5'-CAGTTCTCACATCACATCCG-3'.

The FW primer introduced a single nucleotide mismatch to create an *Nco*I site directly upstream of the RB. The resulting PCR product was digested with *Nco*I and was ligated into the likewise digested vector backbone of pOphis RB LB (making use of two existing *Nco*I sites: one directly downstream the RB and another at the start of the KanMX ORF) resulting in pOphis LB. To construct pOphis RB RBLB an *Xho*I fragment from pUC57 containing both the RB and LB sequence orientated as direct repeats was cloned into the likewise digested vector backbone of pOphis RB. To construct pOphis RB RB an *Eco*RI fragment of pUC57 containing the RB was cloned into the likewise digested vector backbone of pOphis RB.

To construct pOphis LB LB column purified pUC57 was used as a template for PCR to amplify the LB sequence using the following primers:

TLlb *Sac*II FW: 5'-ATTACCGCGGAATTCGACGTCTCTAG-3'

TLlb *Sac*II RV: 5'-ATTACCGCGGTTCAAATACTTTGCGATCG-3'

The PCR product was digested with *Sac*II and cloned into the likewise digested vector backbone of pOphis LB resulting in pOphis LB LB.

### **Plasmid rescue**

The T-DNA circles were rescued from the transformed yeast cells using the protocol as described by Sing *et al* [43] utilizing a lyticase treatment of the yeast cells followed by a plasmid isolation procedure using the Qiagen Plasmid Miniprep Kit; Qiagen, Valencia.

## References

1. Van Larebeke N, Engler G, Holsters M, Van den Elsacker S, Zaenen I, et al. (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* 252: 169-170.
2. Watson B, Currier TC, Gordon MP, Chilton MD, Nester EW (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol* 123: 255-264.
3. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A* 79: 6322-6326.
4. 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.
5. Durrenberger F, Cramer 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.
6. Jasper F, Koncz C, Schell J, Steinbiss HH (1994) *Agrobacterium* T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proc Natl Acad Sci U S A* 91: 694-698.
7. 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.
8. 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.
9. Ward JE, Akiyoshi DE, Regier D, Datta A, Gordon MP, et al. (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. *J Biol Chem* 263: 5804-5814.
10. Lanka E, Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* 64: 141-169.
11. Pansegrau W, Lanka E (1996) Enzymology of DNA transfer by conjugative mechanisms. *Prog Nucleic Acid Res Mol Biol* 54: 197-251.
12. 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.
13. Lessl M, Lanka E (1994) Common mechanisms in bacterial conjugation and Ti-mediated T-DNA transfer to plant cells. *Cell* 77: 321-324.
14. Fronzes R, Christie PJ, Waksman G (2009) The structural biology of type IV secretion systems. *Nat Rev Microbiol* 7: 703-714.
15. Fronzes R, Schafer E, Wang L, Saibil HR, Orlova EV, et al. (2009) Structure of a type IV secretion system core complex. *Science* 323: 266-268.
16. Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative Transfer by the Virulence System of *Agrobacterium tumefaciens*. *Science* 256: 1324-1327.

17. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *Embo J* 14: 3206-3214.
18. Singer K, Shibolet Y, Li J, Tzfira T (2012) Formation of Complex Extrachromosomal T-DNA Structures in *Agrobacterium tumefaciens*-Infected Plants. *Plant Physiol* 160: 511-522.
19. Bravo-Angel AM, Hohn B, Tinland B (1998) The omega sequence of VirD2 is important but not essential for efficient transfer of T-DNA by *Agrobacterium tumefaciens*. *Mol Plant Microbe Interact* 11: 57-63.
20. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, et al. (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.
21. Lanka E, Barth PT (1981) Plasmid RP4 specifies a deoxyribonucleic acid primase involved in its conjugal transfer and maintenance. *J Bacteriol* 148: 769-781.
22. Merryweather A, Barth PT, Wilkins BM (1986) Role and specificity of plasmid RP4-encoded DNA primase in bacterial conjugation. *J Bacteriol* 167: 12-17.
23. Geibel S, Banchenko S, Engel M, Lanka E, Saenger W (2009) Structure and function of primase RepB' encoded by broad-host-range plasmid RSF1010 that replicates exclusively in leading-strand mode. *Proc Natl Acad Sci U S A* 106: 7810-7815.
24. 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.
25. Schrammeijer B, Beijersbergen A, Idler KB, Melchers LS, Thompson DV, et al. (2000) Sequence analysis of the vir-region from *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *J Exp Bot* 51: 1167-1169.
26. Lee LY, Gelvin SB (2008) T-DNA binary vectors and systems. *Plant Physiol* 146: 325-332.
27. Heeb S, Itoh Y, Nishijyo T, Schnider U, Keel C, et al. (2000) Small, stable shuttle vectors based on the minimal pVS1 replicon for use in gram-negative, plant-associated bacteria. *Mol Plant Microbe Interact* 13: 232-237.
28. Podevin N, De Buck S, De Wilde C, Depicker A (2006) Insights into recognition of the T-DNA border repeats as termination sites for T-strand synthesis by *Agrobacterium tumefaciens*. *Transgenic Res* 15: 557-571.
29. van der Graaff E, den Dulk-Ras A, Hooykaas PJ (1996) Deviating T-DNA transfer from *Agrobacterium tumefaciens* to plants. *Plant Mol Biol* 31: 677-681.
30. Dion V, Kalck V, Horigome C, Towbin BD, Gasser SM Increased mobility of double-strand breaks requires Mec1, Rad9 and the homologous recombination machinery. *Nat Cell Biol* 14: 502-509.
31. van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31: 826-832.
32. van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J* 20: 6550-6558.
33. Cook DM, Farrand SK (1992) The oriT region of the *Agrobacterium tumefaciens* Ti plasmid pTiC58 shares DNA sequence identity with the transfer origins of RSF1010 and RK2/RP4 and with T-region borders. *J Bacteriol* 174: 6238-6246.

34. Hamilton CM, Lee H, Li PL, Cook DM, Piper KR, et al. (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.
35. Llosa M, Zunzunegui S, de la Cruz F (2003) Conjugative coupling proteins interact with cognate and heterologous VirB10-like proteins while exhibiting specificity for cognate relaxosomes. *Proc Natl Acad Sci U S A* 100: 10465-10470.
36. Fitcher AB (1986) Copy number amplification of the 2 micron circle plasmid of *Saccharomyces cerevisiae*. *J Theor Biol* 119: 197-204.
37. de la Cruz F, Frost LS, Meyer RJ, Zechner EL Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* 34: 18-40.
38. Llosa M, Gomis-Ruth FX, Coll M, de la Cruz Fd F (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* 45: 1-8.
39. Becker EC, Meyer R Origin and fate of the 3' ends of single-stranded DNA generated by conjugal transfer of plasmid R1162. *J Bacteriol* 194: 5368-5376.
40. Kononov ME, Bassuner B, Gelvin SB (1997) Integration of T-DNA binary vector 'backbone' sequences into the tobacco genome: evidence for multiple complex patterns of integration. *Plant J* 11: 945-957.
41. Wenck A, Czako M, Kanevski I, Marton L (1997) Frequent collinear long transfer of DNA inclusive of the whole binary vector during *Agrobacterium*-mediated transformation. *Plant Mol Biol* 34: 913-922.
42. Bundock P, Mroczek K, Winkler AA, Steensma HY, Hooykaas PJ (1999) T-DNA from *Agrobacterium tumefaciens* as an efficient tool for gene targeting in *Kluyveromyces lactis*. *Mol Gen Genet* 261: 115-121.
43. Singh MV, Weil PA (2002) A method for plasmid purification directly from yeast. *Anal Biochem* 307:13-17.



*“Isn’t this enough? Just this world? Just this beautiful, complex wonderfully unfathomable world?  
How does it so fail to hold our attention that we have to diminish it with the invention of cheap,  
man-made myths and monsters?”*

Tim Minchin

# Chapter 4

## **Enhanced gene targeting mediated by translocated I-SceI during the *Agrobacterium* mediated transformation of yeast**

4

Martijn Rolloos, Paul J.J. Hooykaas and Bert J. van der Zaal

## Abstract

*Agrobacterium* mediated transformation (AMT) has been embraced by biotechnologists as the technology of choice to introduce or alter genetic traits of plants. However, in plants it is largely impossible to predetermine the integration site of the transferred T-strand unless one is able to generate a double stranded break (DSB) in the DNA at the site of interest. In this study, we wanted to investigate whether transfer of site-specific endonucleases via the type IV secretion system (T4SS), concomitantly with T-DNA transfer, can improve the gene targeting efficiency. In addition to that, the effect of different chromatin states on targeted integration, with or without additional use of nucleases, was investigated. By using yeast as a model system for AMT, it was found that *Agrobacterium* mediated translocation of the homing endonuclease I-SceI has a positive effect on the integration of T-DNA via the homologous repair (HR) pathway. Furthermore, we obtained evidence that nucleosome removal has a positive effect on I-SceI facilitated T-DNA integration by HR. Conversely; inducing nucleosome formation at the site of integration removes the positive effect of translocated I-SceI on T-DNA integration.

## Introduction

*Agrobacterium tumefaciens* is a soil born plant pathogen well known for its ability to genetically alter the cells of the dicotyledonous plants that form its main natural host. These genetically modified plant cells are inclined to produce amino acid derivatives, which can be utilized by *Agrobacterium* as a carbon and nitrogen source. In addition to this, plant tissue infected by *Agrobacterium* shows excessive cell proliferation resulting in tumorous outgrowths. Although for agriculturists these are the symptoms of crown gall disease, for biotechnologists the *Agrobacterium* mediated transformation (AMT) of plants has formed the foundation for new technologies leading the way to crop improvement and the production of medicines or biofuels by plants. The precursor of the DNA that is integrated into the plant genome is located on a tumor-inducing (Ti) plasmid flanked by two imperfect direct 25 bp repeats; the left and right border sequences [1]. This transfer-DNA (T-DNA) is processed by the *Agrobacterium* virulence protein VirD2 which induces a nick at the lower strand of the right and left border sequence [2,3]. A single stranded section of the T-DNA with VirD2 covalently attached to the 5' end, the T-strand, is released and targeted to a type IV secretion system (T4SS) which injects the T-strand into the host cell [4].

One characteristic of AMT is that T-DNA tends to integrate randomly into the genome of the recipient cell [5]. Inherent to the randomness of T-DNA integration is the risk of unintended gene disruption at the integration site or deviating expression levels of the introduced transgene caused by different positional effects at integration sites. To address these concerns related to AMT, it would be highly desirable to utilize the homologous recombination (HR) system of plants to replace a predetermined genomic locus with an

incoming T-strand that comprises a strong homology to that locus. Unfortunately, HR-mediated integration of DNA in plants is a rare event, estimated to occur only once per  $10^4$  to  $10^5$  integration events [6].

A breakthrough regarding the targeted integration of T-DNA via HR was the finding that this is strongly increased by the formation of DNA double strand breaks (DSBs) at the target locus [7,8]. With the development of highly site specific artificial nucleases, it is now theoretically possible to generate a DSB at any particular genomic position of interest [9] and a major bottleneck for site-directed mutagenesis has thus disappeared. However, without directly selectable gain of function mutations after gene targeting, finding the few HR-mediated gene targeting events remains challenging. Although negative selection against randomly integrated T-DNA copies can be instrumental for the recovery of rare gene targeting events as was recently demonstrated in rice [10], it is not yet clear whether this will be generically applicable since success might be related to the particular type of target cells used. Other studies towards enhancing gene targeting have explored down regulation of enzymes involved in the NHEJ pathway [11,12], but this approach is not very efficient yet.

Altogether, despite the abundant use of AMT for plant transformation, reports about its successful application for gene targeting are still scarce and a generally applicable strategy is lacking. Additionally, obvious parameters likely to be important for successful application of AMT for gene targeting have still been insufficiently explored. The first parameter that is subject of the present study concerns the chromatin state of the target locus. Although it is tempting to suppose that a more open chromatin structure will facilitate HR events between a chromosomal locus and an incoming T-strand, experimental proof of this idea is still required. Interestingly, it has been demonstrated that the intended effect of target site specific artificial transcription factors very much depends on their accession to DNA [13,14]. As a more open chromatin structure should allow for better access of site specific nucleases as well, our experiments regarding chromatin state have addressed nuclease accession as well. The second parameter to be studied related to the delivery of the required nuclease to the recipient cells. In the present study, we explored the possibility to deliver nuclease proteins to the recipient cell via *Agrobacterium* mediated translocation concomitant with T-strand delivery.

*Agrobacterium* effector proteins are targeted for translocation to the recipient host cell because of a series of positively charged amino acids at their C terminal end that act as a T4SS translocation signal [15]. An illustrative example of this is the C terminal end of the virulence protein VirF (VirF<sup>CT</sup>). By making translational fusions with VirF<sup>CT</sup> a number of DNA modifying proteins have been shown to translocate to the host cells, like the recombinase Cre [15,16]. Recently, it was demonstrated that complex fusion proteins consisting of a NLS, VirD2, I-SceI and a VirF<sup>CT</sup> moiety were translocated to recipient plant cells [17]. These experiments were performed by means of assaying simultaneous transfer of T-strands coupled to such fusion proteins.

Although these experiments did not result in the recovery of targeted integration events, the occurrence of footprints within the original I-SceI recognition site in the host genome proved that the I-SceI moiety retained its nuclease activity after *Agrobacterium* mediated translocation. Hence, T-strands can be equipped with covalently attached modified VirD2 proteins, combining the relaxase activity of VirD2 with enzymatic activities of choice [18].

A possibility to achieve gene targeting via AMT by introducing particles consisting of a T-strand with target site homology, covalently linked with a nuclease moiety that creates the required DSB to trigger an HR-mediated gene targeting event is well worth exploring. However, even without covalent linkage of a nuclease to a T-strand, AMT offers opportunities to deliver both entities separately but simultaneously via the T4SS; in those cases, the T-strand passes to the recipient cell piloted by an unmodified VirD2 protein bound at its 5'end, while the nuclease of interest is produced in the same *Agrobacterium* cell and can be transported by means of a fusion with a VirF<sup>CT</sup>.

4

Previously, it was shown that AMT of *Saccharomyces cerevisiae* (budding yeast) can be used as a model system to address the roles of HR and NHEJ pathways in T-DNA integration [19,20]. Readily available mutants, time-efficient transformation protocols for the introduction of genomic target loci, as well as convenient transformant screening and subsequent analysis further contributes to making AMT of yeast a valuable asset. Of special interest to our study are the well characterized promoter regions of yeast genes involved in the galactose pathway; depending on the carbon source, open or closed chromatin states can be induced at will [21]. In the present work, we therefore used the yeast model to explore effects of the chromatin state for gene targeting via AMT as well as for seeking possibilities for combined *Agrobacterium* mediated introduction of T-strands and the site-specific nuclease I-SceI to enhance gene targeting. This was done in experimental set-ups conducive for transfer of T-strands covalently attached to NLS-VirD2-I-SceI-VirF<sup>CT</sup> fusion proteins or, alternatively, for the transfer of T-strands covalently attached to an unaltered VirD2 molecule with a NLS-I-SceI-VirF<sup>CT</sup> fusion protein translocated separately, albeit from single *Agrobacterium* cells.

As shown, without introduced nucleases, the chromatin state of a target locus was of little importance for gene targeting frequencies. However, evidence was obtained that the observed elevated efficiency of T-DNA integration by the cotransfer of I-SceI fusion proteins is completely abolished when nucleosome formation is induced at the targeted integration site.

## Results

### Experimental setup

To get a deeper insight into the requirements of T-DNA integration by HR a versatile target locus was developed and introduced into the yeast genome. As can be seen in Figure 1, this target locus had a *URA3* marker gene for positive and negative selection flanked by two 18 bp two I-SceI recognition sites. To enable the modulation of the nucleosome occupancy of the target locus, two different DNA sequences, normally present upstream of galactose inducible genes of the yeast galactose metabolism pathway, were cloned adjacent to the I-SceI sites. The same sequences, *UAS7* and *UAS10,1*, were also introduced into a binary vector as part of the T-DNA sequence flanking a yeast KanMX marker which lacks homology to the yeast genome. During AMT, a T-strand carrying these sequences will enter recipient yeast cells. With the *UAS7* and *UAS10,1* sequences, the T-strand can interact with the identical chromosomal sequences at the target locus. Regarding the target locus, such an interaction can result in *URA3* for KanMX marker exchange by a two-sided HR event. As for the *UAS7* and *UAS10,1* sequences, their nucleosome occupancy is known to be reduced in the presence of galactose, and increased in the presence of glucose, resulting in increased and reduced expression of downstream ORFs, respectively. Detailed nuclease protection patterns for the sequences derived from yeast growing on galactose or glucose have become available [21]. The wide application of *UAS* sequences for carbon-source dependent gene expression provides compelling evidence that their characteristics, including the nucleosome occupancy, are portable traits.

In order to increase HR at the target locus, DSBs can be induced by I-SceI at the cognate restriction sites. We tested whether such DSBs can be induced efficiently by I-SceI translocated from *Agrobacterium* as a VirD2-I-SceI-VirF<sup>CT</sup> fusion protein that is covalently attached to the translocated T-strand (Figure 1A). This protein, as well as the other fusion proteins used in this study, contains an N-terminal FLAG tag for immunodetection followed by an SV40 nuclear localization signal (NLS) to ensure its nuclear entry. Since the full length VirD2 protein contains an active NLS sequence, an extra NLS is not really required [18]. For the I-SceI-VirF<sup>CT</sup> protein, an active NLS is a prerequisite. Expression in *Agrobacterium* was mediated by the *virF* promoter and verified by Western blotting (data not shown).

The I-SceI moiety should induce one or more DSBs at the target locus, thereby enhancing the integration of T-DNA mediated by the HR pathway. Since the Holliday junctions involved in crossing over events during HR pathway require two dsDNA molecules, marker exchange is depicted as occurring between two double stranded DNA molecules (Figure 1B), as if the incoming T-strand has formed a complementary strand prior to integration. Incoming T-strands did not contain a nuclease target site.

The target locus depicted in Figure 1, as well as those mentioned in Table 1, were integrated in the genomic *PDC6* locus. By performing the AMT of these yeast target lines on medium containing galactose or glucose, the respective removal or formation of nucleosomes can be induced and their influence on the integration efficiency can be assessed. For control experiments, alternative T-DNA regions were constructed, aimed to generate T-strands just representing the *KanMX* marker gene, without flanking *UAS* regions or to possess similarly sized flanking *PDC6* sequences instead of *UAS* regions (see Table 1).

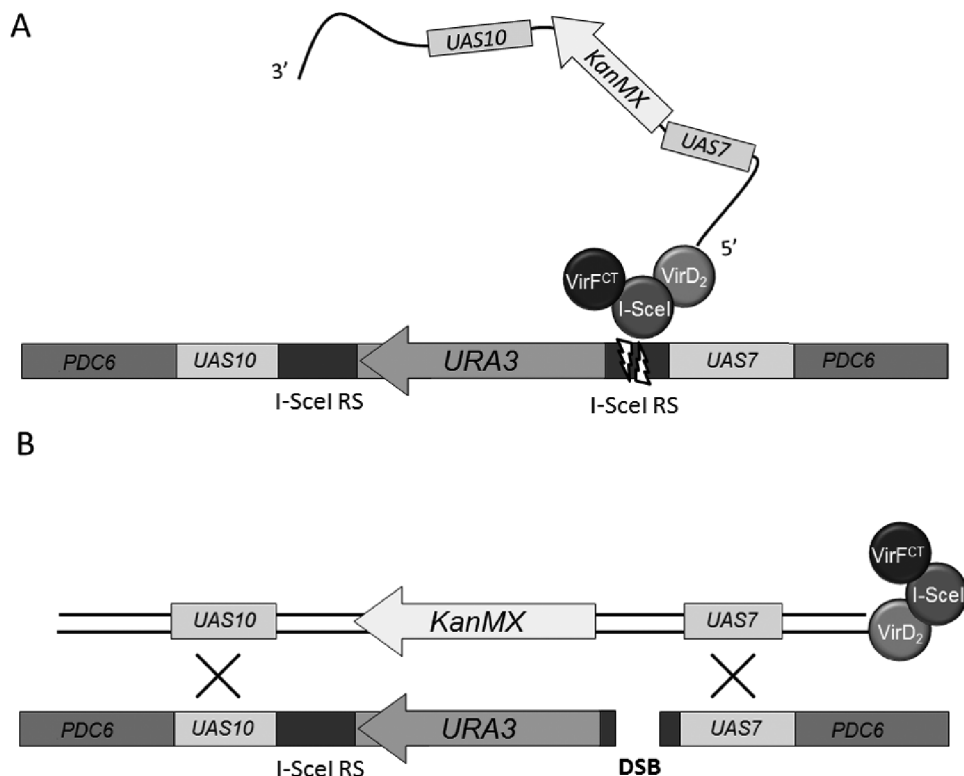


**Figure 1. Experimental set-up for the study of T-DNA integration via HR by the I-SceI mediated DSB formation.** (A) The T-strand entering yeast cells is covalently attached to a fusion protein consisting of an FLAG-NLS module (not depicted in the figure), VirD2, I-SceI, and an *Agrobacterium* translocation signal derived from the VirF C-terminus. The I-SceI moiety induces the formation of a DSB at one or both the I-SceI recognition sites flanking the *URA3* gene, which triggers DNA recombination. (B) Precise HR of the target locus with the incoming T-strand results in the exchange of the *URA3* autotrophy marker with the *KanMX* marker gene leading to G418 resistance as well as uracil auxotrophy. The regions on the T-strand homologous to the *UAS10,1* and *UAS7* sequences in the target locus are approximately 700 bp in length.

AMT experiments with T-strands covalently attached to VirD2 containing fusion proteins, such as FLAG-NLS-VirD2-I-SceI-VirF<sup>CT</sup>, were performed using *Agrobacterium* strain LBA2556 [22] containing a Ti plasmid deleted for the *virD2* gene. For T-strand transfer combined with transfer of non-VirD2 containing proteins, such as FLAG-NLS-I-SceI-VirF<sup>CT</sup>, the isogenic strain LBA1100 was used.

In this strain, a wild type VirD2 protein is still produced. In addition to a target locus that is substrate for I-SceI (Figure 2A), similarly designed target loci were prepared harboring recognition sites of the recombinase Cre (Figure 2B) and a negative control locus lacking any particular recognition sites (Figure 2C). A complete list of all gene constructs used for the experiments and their abbreviations is presented in Table 1.

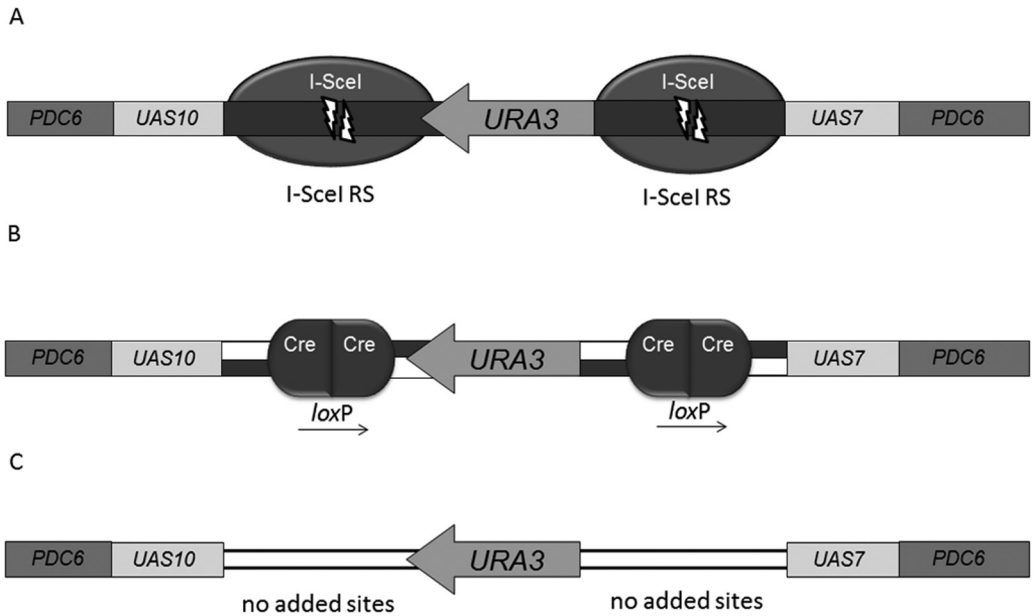


**Figure 2. Target loci used to assay the effect of DSBs and nucleosome occupancy on the integration of T-DNA by HR.** These target loci were all integrated at the *PDC6* locus, all containing sequence homology to the incoming T-strands by two different *UAS* type promoters derived from the yeast galactose pathway or by the flanking *PDC6* fragments. To remove nucleosomes from these promoter sequences, the yeast cells are grown in a medium with galactose as the sole carbon source. Growth in a glucose medium would induce the formation of nucleosomes on the *UAS* typed promoters. The central *URA3* marker is flanked by: (A) recognition sites of the homing endonuclease I-SceI that catalyzes the formation of DSBs as a monomer, (B) recognition sites of the homing recombinase Cre allowing for the removal of the *URA3* marker gene without inducing double stranded break formation, (C) target locus without any added recognition sites but still provided with homology to the incoming T-strands.

**Table 1. Overview of all the constructs mentioned in the experimental setup.** For convenience, the constructs were given a short working name. All uninterrupted ORFs in the description are depicted in bold writing. The abbreviation “I-SceI-RS” indicates recognition sites for I-SceI.

| Name                | Description [reference]                                      | Vector type, organism               |
|---------------------|--------------------------------------------------------------|-------------------------------------|
| pIDU                | pINT: <i>PDC6:URA3:PDC6</i>                                  | Integrative vector, yeast           |
| pID2GU              | pINT: <i>PDC6:UAS10:URA3:UAS7:PDC6</i>                       | Integrative vector, yeast           |
| pID2G2LU            | pINT: <i>PDC6:UAS10:loxP:URA3:loxP:UAS7:PDC6</i>             | Integrative vector, yeast           |
| pID2G2SU            | pINT: <i>PDC6:UAS10:I-SceI RS:URA3:I-SceI RS:UAS7:PDC6</i>   | Integrative vector, yeast           |
| pRS425 I-SceI       | pRS425:pADH: <b>FLAG::NLS::I-SceI::virF<sup>CT</sup></b>     | Expression vector, yeast            |
| pRS425 VirD2-I-SceI | pRS425:pADH: <b>NLS::virD2::I-SceI::virF<sup>CT</sup></b>    | Expression vector, yeast            |
| pRS425 Cre          | pRS425:pADH: <b>FLAG::NLS::Cre::virF<sup>CT</sup></b>        | Expression vector, yeast            |
| pRS425 VirD2-Cre    | pRS425:pADH: <b>FLAG::NLS::virD2::Cre::virF<sup>CT</sup></b> | Expression vector, yeast            |
| pBFF I-SceI         | pBFF: <b>FLAG::NLS::I-SceI::virF<sup>CT</sup></b>            | Expression vector, yeast            |
| pBFF VirD2-I-SceI   | pBFF: <b>FLAG::NLS::virD2::I-SceI::virF<sup>CT</sup></b>     | Expr. vector, <i>Agrobacterium</i>  |
| pBFF Cre            | pBFF: <b>FLAG::NLS::Cre::virF<sup>CT</sup></b> [16]          | Expr. vector, <i>Agrobacterium</i>  |
| pBFF VirD2-Cre      | pBFF: <b>FLAG::NLS::virD2::Cre::virF<sup>CT</sup></b>        | Expr. vector, <i>Agrobacterium</i>  |
| p14-PDC6KX          | pSDM14: <i>rightborder:PDC6:KanMX:PDC6:leftborder</i>        | Binary vector, <i>Agrobacterium</i> |
| p14-2GKX            | pSDM14: <i>rightborder:UAS7: KanMX:UAS10:leftborder</i>      | Binary vector, <i>Agrobacterium</i> |
| pSDM8000            | pSDM14: <i>rightborder:KanMX:leftborder</i> [19]             | Binary vector, <i>Agrobacterium</i> |

### Qualitative assay on enzyme activity of I-SceI and Cre fusion proteins

To test if the expressed fusions proteins were able to process the target loci described in Figure 2, an *in vivo* assay was designed. For this, the ORFs of the VirD2 fusion proteins were transferred to the yeast constitutive expression vector pRS425ADH equipped with a leucine autotrophy marker [23]. This resulted in pRS425ADH expressing NLS-VirD2-I-SceI-VirF<sup>CT</sup> (pRS425 VirD2-I-SceI) and NLS-VirD2-Cre-VirF<sup>CT</sup> (pRS425 VirD2-Cre) to serve as a positive control.

These vectors were transferred to yeast cells carrying the corresponding target loci, or a negative control without any cognate recognition sites. Fifty yeast colonies resulting from each vector transformation with one of the expression vectors were pooled. An aliquot containing at least 1000 suspended yeast cells was plated on selective medium without leucine and uracil and on medium supplemented with 5-Fluoroorotic Acid (5-FOA) and uracil in order to determine the fraction of yeast cells that had lost the *URA3* marker, rendering them 5-FOA resistant. Although the described assay did not provide quantitative data on enzyme activity, it did show that the two VirD2 fusion proteins were able to process their cognate target loci but not the negative control target locus without recognition sites. The expression of I-SceI fusion proteins resulted in a complete loss of the *URA3* marker gene, Cre fusion proteins removed 88% of this marker.

### **The effect of galactose or glucose treatment on AMT transformation efficiency**

It was assessed whether the usage of alternative carbon sources could have an impact on the AMT transformation efficiency. As can be seen in Figure 3, using galactose as carbon source instead of glucose during the AMT cocultivation period resulted in a steep increase in the transformation efficiency for the yeast strain YPH250. In case the positive contribution of galactose containing medium to AMT efficiency would have been limited to HR-mediated integration events due to the *UAS* sequences present in p14 2GKX, it would have supported a model where induced nucleosome depletion allowed for enhanced T-DNA integration. However, the galactose effect was observed with all T-DNA vectors used, thus also for p14 PDC6KX which produces T-strands with *PDC6*-derived flanking sequences instead of *UAS* sequences and for pSDM8000 which completely lacks homology to the target locus. Inverse PCR products representing five randomly picked yeast strains, transformed with pSDM8000 on galactose plates, were sequenced to determine the integration sites of the T-DNA. This demonstrated that the KanMX marker gene was present at five different genomic loci (data not shown), as is to be expected for a randomized integration pattern. None of these sites represented loci involved in the galactose pathway. Therefore, the observed positive effects of galactose on T-DNA integration in strain YPH250 were unlikely to be attributable to differences in nucleosome occupancy at *UAS* sequences.

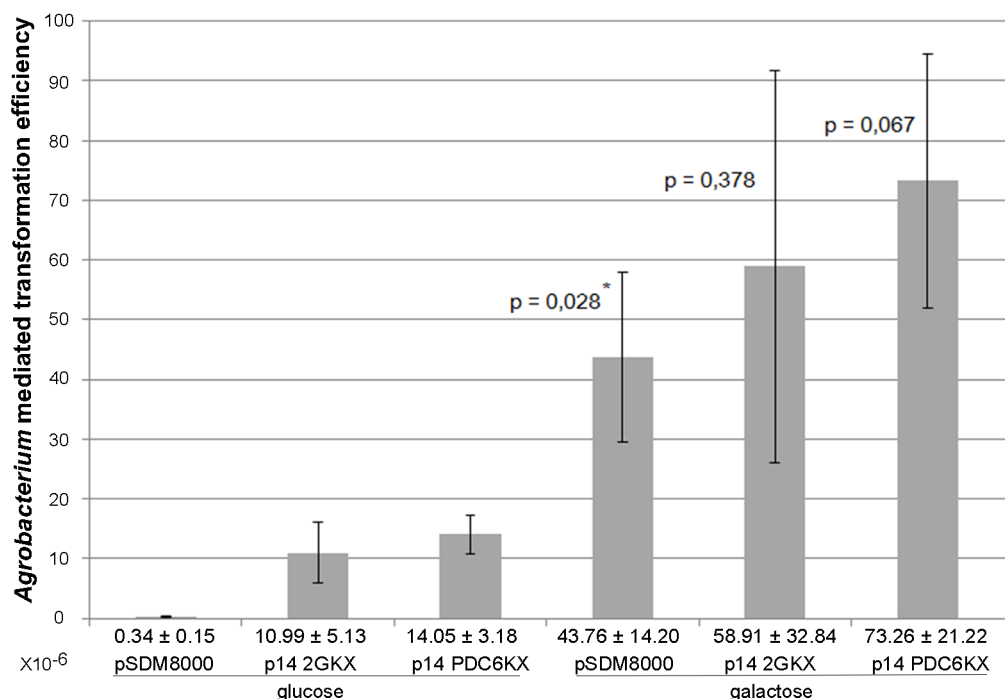


**Figure 3. Effects of the carbon source on AMT efficiency using yeast strain YPH250.** YPH250 was equipped with a pID2GU target locus consisting of two UAS sequences flanking URA3 (see Figure 2C) and transformed with the *Agrobacterium* strain LBA1100 harboring the binary vectors p14 2GKX, p14 PDC6KX or pSDM8000. The transformation frequencies following cocultivation on glucose or galactose containing medium were calculated by dividing the total number of G418 resistant colonies by an estimate of the total number of yeast cells. Each bar represents the average of four independent experiments. Error bars represent the SEM. P-values given represent those found for a two-tailed student's T-test comparing the results on glucose or galactose treatment for a given binary vector. Statistical significance, with  $\alpha$  set to 0.05, is indicated with an asterisk.

To investigate whether galactose stimulation of AMT frequency was a generally occurring phenomenon, a similar experiment was performed using the yeast strain RSY12, which is often used for AMT in our lab. Remarkably, when using RSY12 harboring a pID2GU reporter locus in combination with T-DNA from the binary vector p14-2GKX, the frequency of T-DNA integration was not significantly affected by a change of carbon source (Figure 4). The same was observed for a negative control consisting of a pIDU reporter locus lacking UAS promoter sequences in combination with T-DNAs derived from p14-PDC6KX. Hence, it was evident that the use of RSY12 was preferable for further experimentation with galactose or glucose containing media, thus avoiding unnecessarily skewed representation of AMT frequencies. From Figure 4, it can be seen that the presence or absence of nucleosomes at the UAS containing area of the target locus had no measurable effect on the efficiency of T-DNA integration.

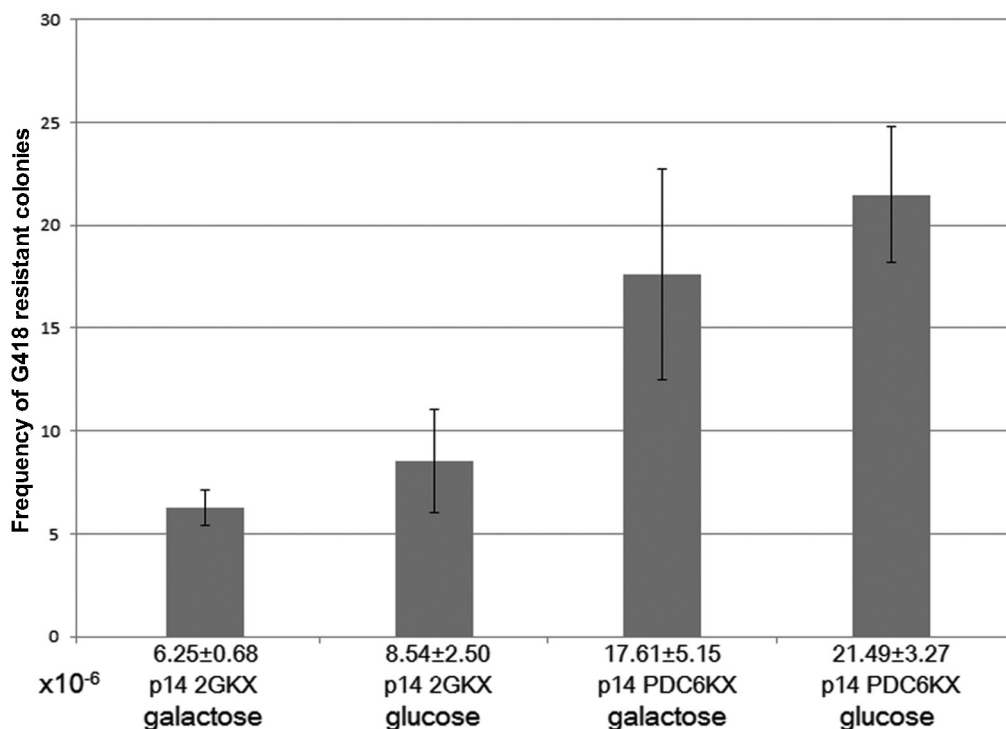


**Figure 4. Effects of carbon source on AMT efficiency using yeast strain RSY12.** The first and second bar represent RSY12 with the pID2GU target locus cocultivated with LBA1100 harboring p14-2GKX. The third and fourth bar represent RSY12 with the pIDU target locus, lacking the UAS promoter sequences, cocultivated with LBA1100 harboring p14-PDC6KX. The average results of ten independent experiments are shown. Error bars indicate the SEM. The hypothesis of equality could not be rejected with  $\alpha$  set to 0.05 when performing a two-tailed student's T-test comparing "galactose" with "glucose".

### A comparison of two approaches to induce DSBs to affect T-DNA integration via HR

Using the yeast strain RSY12 harboring the target locus depicted in Figure 1, it was assessed whether it would be possible to obtain evidence for a promotion of gene targeting efficiency by delivery of nuclease activity during AMT. The first approach was based on translocation of T-strands piloted by a chimerical NLS-VirD2-I-SceI-VirF<sup>CT</sup> protein using an *Agrobacterium* strain with a *virD2* deletion (LBA2556), the second one on T-strand delivery via an isogenic *Agrobacterium* strain containing a WT *VirD2* locus (LBA1100) concomitant with transfer of NLS-I-SceI-VirF<sup>CT</sup>. Apart from the plasmids expressing the proteins of interest, both strains contained the binary T-DNA vector p14-2GKX (Table 1).

The observed AMT frequencies are depicted in Figure 5. In this figure, the third and fourth bars represent cocultivations with LBA2556 harboring p14-2GKX and pBFF VirD2-I-SceI.

The presence of two I-SceI restriction sites flanking *URA3* did not lead to an increase in transformation efficiency using this *Agrobacterium* strain. Different results were obtained when yeast cells were cocultivated with LBA1100 harboring p14-2GKX and pBFF I-SceI. Here, WT VirD2 directed the translocation of the T-strand while NLS-I-SceI-VirF<sup>CT</sup> was translocated separately to induce DSBs at the target locus. Now a significant effect of the presence of the I-SceI recognition sites at the target locus was observed (Figure 5, first two bars). Interestingly, this positive effect of translocated NLS-I-SceI-VirF<sup>CT</sup> on T-DNA integration via HR was found when galactose had been present during the cocultivation period but not when glucose was added to the induction medium. This observation suggests that the efficiency of T-DNA integration is affected by the nucleosome occupancy of the target locus, but merely because it affects the degree of access of translocated nuclease to the target locus. Enhanced nucleosome formation due to addition of glucose apparently leads to a reduced accessibility of the target locus for the I-SceI fusion protein, while galactose has the opposite effect.

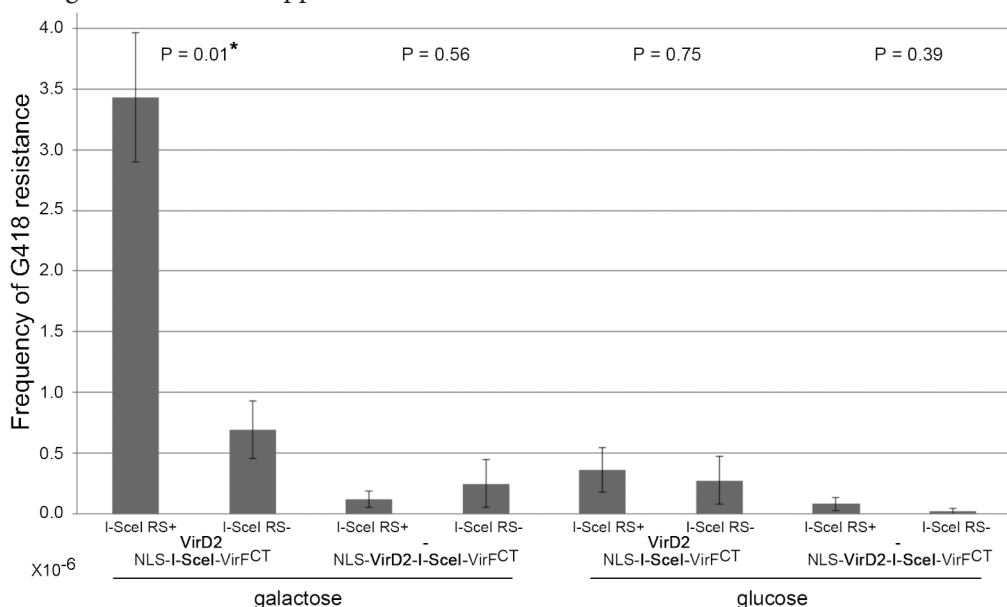


**Figure 5. Influence of galactose or glucose treatment on the I-SceI assisted integration of T-DNA by HR.** The yeast strain RSY12 was cocultivated with LBA1100 harboring p14-2GKX and pBFF I-SceI or LBA2556 p14-2GKX and pBFF VirD2-I-SceI. Different yeast strains with and without I-SceI recognition sites between the uracil autotrophy marker and the *UAS* promoter sequences making up the target locus (Figure 2A, 2C) were compared, here indicated as I-SceI RS<sup>+</sup> and I-SceI RS<sup>-</sup>. A horizontal line is used to mark the presence of galactose (inducing nucleosome removal *UAS* regions in the target locus) or glucose (inducing nucleosome formation). Each bar represents the average of four independent experiments. Error bars represent the SEM. The p-values were calculated performing a two-tailed student's T-test. The asterisk indicates statistical significance with  $\alpha = 0.05$ .

To verify that the majority of the integration events indeed took place at the target locus, G418 resistant colonies were assessed for loss of uracil autotrophy. As shown for colonies resulting from three different cocultivations (Figure 6), the majority of yeast cells lost the

*URA3* marker, indicative for HR-mediated gene replacement. The few colonies that retained uracil autotrophy possibly represent single cross-overs at the target locus or at native *UAS7* or *UAS10,1* promoter sequences that are part of the cognate *GAL* genes.

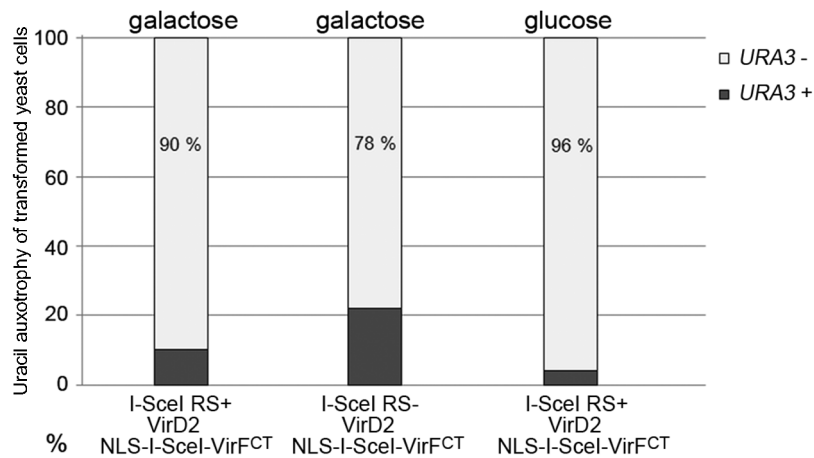


**Figure 6. Assay on uracil auxotrophy of G418 resistant yeast transformants.** As a consequence of the integration of the T-DNA (*UAS10,1:kanMX:UAS7*) mediated by HR, the uracil autotrophy marker of the target locus can be replaced by the KanMX marker gene leading to Ura auxotrophy (here indicated as *URA3*-). The off-target integration of the T-DNA leaves the *URA3* gene intact (here indicated as *URA3*+). Each bar represents the average of fifty G418 resistant colonies assayed for growth on SD-ura.

### Combining protein transfer to yeast with integration of non-homologous T-DNA

Using the tools developed, we wanted to investigate whether NHEJ mediated T-DNA targeting could be achieved by the simultaneous translocation of I-SceI fusion proteins NLS-I-SceI-VirF<sup>CT</sup> or NLS-VirD2-I-SceI-VirF<sup>CT</sup>. Similar to the experiments described above, yeast cells were transformed containing a target locus harboring I-SceI recognition sites. As a control for a possible effect of I-SceI induced DSBs mediated T-DNA targeting, NLS-Cre-VirF<sup>CT</sup> and NLS-VirD2-Cre-VirF<sup>CT</sup> were used in a similar experiment with YPH250 containing the *URA3* marker gene flanked by two *loxP* sites (Figure 2). In this case, Cre transfer might trigger a recombination reaction, causing loss of the *URA3* marker but without generating a DSB. To boost overall AMT frequency, AMT of strain YPH250 was performed using induction medium containing galactose as the sole carbon source (see Figure 3). Direct screening for loss of the *URA3* marker led to background problems when more than 10<sup>5</sup> cells were put on a single 5-FOA containing plate. Considering the much larger numbers of cells that needed to be assayed, cells were first selected for G418 resistance and subsequently for loss of the the *URA3* marker. To get at least 10 colonies per AMT combination, 80 cocultivations were performed. Especially the yeast transformations involving *Agrobacterium* strains harboring VirD2 fusion proteins were of very low efficiency. The results are given in Table 2.

**Table 2.** Screening for T-DNA integration via the NHEJ pathway, possibly facilitated by DSB induction by translocated fusion proteins. The third column shows the percentage of cells for which the *URA3* marker was found inactive. The actual number of uracil auxotrophic colonies divided by the number of analyzed transformant colonies is given in between brackets.

| Fusion protein                      | Target locus                            | % of Ura auxotrophy (nr. Ura/nr. G418 <sup>r</sup> ) |
|-------------------------------------|-----------------------------------------|------------------------------------------------------|
| NLS-I-SceI-VirF <sup>CT</sup>       | I-SceI RS - <i>URA3</i> - I-SceI RS     | 0.44 (4/901)                                         |
| NLS-I-SceI-VirF <sup>CT</sup>       | Empty - <i>URA3</i> - Empty             | 0 (0/24)                                             |
| NLS-VirD2-I-SceI-VirF <sup>CT</sup> | I-SceI RS - <i>URA3</i> - I-SceI RS     | 0 (0/12)                                             |
| NLS-VirD2-I-SceI-VirF <sup>CT</sup> | Empty - <i>URA3</i> - Empty             | 0 (0/38)                                             |
| NLS-Cre-VirF <sup>CT</sup>          | <i>loxP</i> - <i>URA3</i> - <i>loxP</i> | 57 (103/181)                                         |
| NLS-Cre-VirF <sup>CT</sup>          | Empty - <i>URA3</i> - Empty             | 0 (0/62)                                             |
| NLS-VirD2-Cre-VirF <sup>CT</sup>    | <i>loxP</i> - <i>URA3</i> - <i>loxP</i> | 32 (11/34)                                           |
| NLS-VirD2-Cre-VirF <sup>CT</sup>    | Empty - <i>URA3</i> - Empty             | 0 (0/18)                                             |

AMT with the possible concomitant transfer of NLS-I-SceI-VirF<sup>CT</sup> proteins did result in four Ura auxotrophic colonies out of 901 transformants. However, PCR analysis of the I-SceI containing target locus showed that the *URA3* marker had not been removed and that the KanMX marker gene had inserted elsewhere. The nature of the loss of *URA3* activity was not further investigated. When using the *virD2* deletion mutant *Agrobacterium* strain LBA2556 for translocation of NLS-VirD2-I-SceI-VirF<sup>CT</sup> bound to the T-strand, only a few dozen G418 resistant cells were obtained, despite a large number of cocultivations. None of these colonies had lost their uracil autotrophy.

When AMT was performed using yeast cells harboring a target locus consisting of *URA3* flanked by two *loxP* sites, with an *Agrobacterium* strain translocating NLS-Cre-VirF<sup>CT</sup>, more than half of the transformed cells had lost uracil autotrophy. Five colonies varying in their uracil autotrophy status were PCR-analyzed and found to have lost the *URA3* marker at the target locus each time the transformed cells were previously found to be autotrophic for Ura. Insertion of the KanMX marker was not found. AMT via T-strand transfer aided by NLS-VirD2-Cre-VirF<sup>CT</sup> resulted in 32% of the G418 resistant cells losing uracil autotrophy. PCR analysis of ten randomly picked yeast transformants varying in uracil autotrophy status confirmed the presence or absence of the *URA3* marker in the target locus corresponding to their ability to grow on medium without Ura. The KanMX marker was in every case found to be integrated outside the target locus.

## Discussion

Using a yeast model system, we investigated possible effects of different chromatin states on gene targeting by means of AMT. We also assessed the contribution of co-delivered fusion proteins containing I-SceI nuclease domains to facilitate targeted integration of T-DNA during the AMT. Based on our findings, we conclude that enhancing or decreasing nucleosome occupancy of a target locus has no significant impact on the integration efficiency in the absence of a site specific nuclease. However, the presence of nucleosomes may prevent the action of such a specific nuclease, thereby obscuring the positive effect of DSBs on T-DNA integration. The obtained results clearly demonstrated that combined translocation of nuclease proteins and T-strands during AMT facilitated HR mediated integration of T-DNA.

After evaluating the performance of the experimental tools generated, experiments were performed to determine if the nucleosome occupancy of a target locus would have an impact on the efficiency of T-DNA integration via the HR pathway. As described above, we used promoter sequences from genes that are part of the yeast galactose pathway as the regions homologous to the incoming T-strand at the target locus. The nucleosome occupancy of the *UAS* type promoter sequences used has been well characterized. For *UAS10,1* it has been shown that galactose induced association of the transcriptional activator Gal4 rapidly recruits the nucleosome remodeling complex *SWI/SNF* which then removes nucleosomes from this promoter [21]. In another study, a Gal4 binding site was transferred to a histidine promoter located on a yeast episome [24]. On galactose-induced expression of Gal4, nucleosomes adjacent to the Gal4 binding site were removed, demonstrating that Gal4-mediated nucleosome removal was independent of sequences flanking the Gal4 binding site. Due to background signals from the *UAS* sequences already present in the genome, changes in nucleosome occupancy within the introduced target locus could not be addressed adequately in our case (data not shown). Nevertheless, concerning the studies mentioned above, it should be safe to assume that the unaltered *UAS* promoters used still exhibit the reported galactose- and glucose dependent regulation of nucleosome occupancy.

Surprisingly, our experiments showed that HR-mediated integration of T-DNA at the target locus was independent of the presence or absence of nucleosomes. This is an intriguing observation since for yeast cells transformed with linear dsDNA using a LiAc/ssDNA/PEG transformation protocol, the presence of nucleosomes at the genomic locus carrying sequence homology to the incoming DNA fragment is known to negatively affect the transformation efficiency [25]. Regarding AMT, there are several lines of evidence that host chromatin components interact with the T-complex of *Agrobacterium* [26-29], but whether this interaction facilitates genomic integration of T-DNA is presently unclear.

The results regarding the effects of nuclease proteins translocated concomitantly with T-strands into yeast cell showed that of the proteins tested, translocation of NLS-I-SceI-VirF<sup>CT</sup> clearly enhanced T-DNA integration, when I-SceI recognition sites were present at the target locus (Figure 5). The majority of the integration events were correctly targeted and the presence of I-SceI recognition sites at the target locus even led to a slight increase in correctly targeted integration events (Figure 6). These data provide evidence that it is indeed feasible to translocate nucleases to introduce a DSB in the host genome during AMT simultaneously with T-strands. Using yeast cells with an already very effective HR pathway, and ineffective non-homologous T-DNA integration, the improvements regarding gene targeting remained relatively modest. However, in systems where the HR mechanism is much less proficient, delivery of nuclease proteins might very well serve to improve gene targeting. Procedures used thus far to improve T-DNA targeting, have always relied on the introduction of gene constructs encoding nucleases of interest, but our methodology of “protein therapy” in combination with a gene targeting construct has the advantage that no extra nucleic acids need to be introduced besides the gene targeting construct.

Strikingly, as can be observed in Figure 5, the enhanced transformation efficiency that was obtained by the *Agrobacterium* mediated translocation of NLS-I-SceI-VirF<sup>CT</sup> was abolished if glucose was present in the medium instead of galactose. As is illustrated in Figure 4, the presence of galactose or glucose had no effect on the transformation efficiency when using yeast strain RSY12. Therefore, repression of AMT frequency in the presence of glucose must be indicative of glucose induced nucleosome formation at the UAS target sites, thereby reducing the availability of the I-SceI binding sites for digestion by NLS-I-SceI-VirF<sup>CT</sup>. The consequent reduction in DSB formation at the target locus would then account for the reduced transformation efficiency.

While an *in vivo* assay demonstrated that NLS-VirD2-I-SceI-VirF<sup>CT</sup> still possessed nuclease activity (Table 2), attempts to use this chimerical protein for T-strand delivery as well as DSB induction only resulted in very low transformation efficiencies, with too few yeast transformants to give a reliable estimate of the T-DNA targeting efficiency (Figure 5). This could very well be due to reduced T4SS mediated translocation of NLS-VirD2-I-SceI-VirF<sup>CT</sup> compared to WT VirD2 or, alternatively, to reduced relaxase activity of VirD2 when fused to other protein moieties. Both issues would lead to less T-strands being transferred, leading to a lower transformation efficiency. However, the chimerical VirD2 proteins can still support enzyme activity for DNA metabolism as shown for the ability of NLS-VirD2-I-Cre-VirF<sup>CT</sup> to support Cre-mediated recombination on a yeast target locus (Table 4). In parallel experiments performed in plant cells, it was shown that also NLS-VirD2-I-SceI-VirF<sup>CT</sup> can be transferred during AMT, while retaining nuclease activity [17]. Therefore, unless bottlenecks can be adequately addressed, the collected data indicate that the most effective approach to transfer nuclease proteins to induce DSBs concomitant with AMT is the separate translocation of the nuclease proteins rather

than making more complex VirD2 fusion proteins. When only assessing the transfer of a very similar NLS-Cre-VirFCT protein to yeast cells containing a *loxP* flanked locus, it was previously found that up to 1% of yeast target cells exhibited the recombined genotype after AMT [16]. In summary, the data obtained using the yeast model system for AMT indicated that the nucleosome occupancy status of the future integration site was irrelevant for gene targeting. However, site specific nucleases can be introduced concomitantly with T-strand transfer. For a clearly positive contribution to gene targeting, it is of great importance that the nucleases can get access to the intended target site, not hindered by nucleosome occupancy.

## Materials and methods

### The construction of the yeast target lines RSY12 pID2GU and RSY12 pIDU

For the amplification of the *UAS* typed promoter sequences from the yeast galactose pathway, genomic DNA from the *Saccharomyces cerevisiae* strain YPH250 [30] was used as a template. The PCR is performed using the primers listed in figure x. The PCR product from *UAS7* was digested with *Bam*HI and *Xho*I and cloned into pSKNS*gral* which was also digested with *Bam*HI and *Xho*I. The resulting vector was named pOPSL3. The PCR product from *UAS10* was digested with *Spe*I and *Bam*HI and cloned into pSKNS*gral* also digested with *Spe*I and *Bam*HI. The resulting vector was named pOPSL5. The cloning vector pSKNS*gral* has been described by Neuteboom *et al* [31]. After these initial cloning steps pOPSL5 was digested with *Bam*HI and *Spe*I. The resulting fragment containing the *UAS10* fragment was cloned into pOPSL3 digested with *Bgl*II and *Spe*I resulting in a new vector containing both upstream activating sequences; p2Gstart. This vector in turn was digested with *Spe*I and *Xho*I. A PCR fragment of *URA3* digested with *Spe*I and *Xho*I was cloned in between the upstream activating sequences resulting in p2GU. The primers used for the amplification of *URA3* are depicted in table 3.

**Table 3.** Oligos designed for the amplification of the galactose pathway promoter sequences *UAS7* and *UAS10*, using genomic DNA from yeast strain YPH250 as a PCR template. For the amplification of the autotrophy marker gene *URA3*, the yeast shuttle vector pRS316 was used as PCR template. The restriction sites added to the PCR primer sequences are underlined. Each oligo is written in the 5'-3' direction.

| Oligo name                     | Nucleotide sequence                                 |
|--------------------------------|-----------------------------------------------------|
| UASGAL7TTGAL10 <i>Bam</i> HI F | GAGTAGGATCCAATATTCAACTGTTTTTTTTTATCATG              |
| UASGAL7TTGAL10 <i>Xho</i> I R  | ATCGCTCGAGTTGCCAGCTTACTATCCTTC                      |
| UASGAL10-GAL1 <i>Spe</i> I F   | TGACA <u>CTAGT</u> TTGAATTTTCAAAAATTCTTACTTTTTTTTGG |
| UASGAL10-GAL1 <i>Bam</i> HI R  | GCATGGATCCTTGACGTTAAAGTATAGAGGT                     |
| URA <i>Xho</i> FW              | TCCTC <u>TCGAG</u> CTTTTCAATTCAATTCATC              |
| URA <i>Spe</i> RV              | TGGC <u>ACTAGT</u> GGGTAATAACTGATATAATTAAATT        |

The shuttle vector pRS316 [30] was used as template for the PCR. From p2GU a *UAS10-URA3-UAS7 BamHI* fragment was excised and cloned into pINT [32], a vector that integrates into the *PDC6* locus [33] resulting in pID2GU. The intact *PDC6* locus contains a weakly expressed pyruvate decarboxylase (*PDC*) gene that is not required for full *PDC* activity in yeast cells [33,34]. To prepare the pINT vector for this cloning step it was digested with *BamHI* and *SphI* and treated with T4 DNA polymerase to create blunt ends recreating the *BamHI* recognition site which was used to clone the *UAS10-URA3-UAS7 BamHI* fragment into. As a negative control, another target locus was required without the galactose inducible promoter sequences to rule out any down or upstream effects of the UAS nucleosome status on the neighboring *PDC6* locus. For this, the digested PCR product of *URA3* with *XhoI* and *SpeI* overhangs was cloned directly into pINT digested with *SalI* and *SpeI*. The yeast strains RSY12 (MATa leu2-3, 112 his3-11, 15 ura3D::HIS3 [35]) and YPH250 (MATa, ura3-52, lys2-801, ade2-101, trp1-Δ1, his3- Δ 200, leu2- Δ 1 [30]) were transformed with pID2GU and pIDU by LiAc/ssDNA/PEG transformation [36].

#### The construction of the yeast target lines pID2G2SU and pID2G2LU.

The plasmids pID2G2SU and pID2G2LU were constructed by the successive addition of two DNA fragments containing the enzyme recognition sites, directly upstream and downstream the *URA3* autotrophy marker gene of pID2GU. Pairs of complementary oligos depicted in Table 4 were annealed resulting in dsDNA fragments with ss overhangs compatible with pID2GU, successively digested with *XhoI* and *SpeI*. The yeast strains RSY12 [35] and YPH250 [30] were transformed with pID2G2SU and pID2G2LU by LiAc/ssDNA/PEG transformation [36].

**Table 4.** Oligos used for the construction of DNA fragments added to the integrative vector pID2GU resulting in pID2G2SU (equipped with I-SceI recognition sites) and pID2G2LU (equipped with Lox sites). Each oligo is written in the 5'-3' direction.

| Plasmid   | Oligo name  | Nucleotide sequence                      |
|-----------|-------------|------------------------------------------|
| pID2G2SU: | scesalxhoFW | TCGACCTTAAGTTACGCTAGGGATAACAGGGTAATATAAC |
|           | scesalxhoRV | TCGAGTTATATTACCCTGTTATCCCTAGCGTAACTTAAGG |
|           | scespexbaFW | CTAGTACGCTAGGGATAACAGGGTAATATAACGGAATTCT |
|           | scespexbaRV | CTAGAGAATTCCGTTATATTACCCTGTTATCCCTAGCGTA |
| pID2G2LU: | loxsalxhoFW | TCGACATAACTTCGTATAGCATACATTATACGAAGTTATC |
|           | loxsalxhoRV | TCGAGATAACTTCGTATAATGTATGCTATACGAAGTTATG |
|           | loxspexbaFW | CTAGTATAACTTCGTATAGCATACATTATACGAAGTTATT |
|           | loxspexbaRV | CTAGAATAACTTCGTATAATGTATGCTATACGAAGTTATA |

## The construction of the binary and protein translocation vectors for *Agrobacterium*

Two types of vectors were used; binary vectors for the transfer of T-DNA and a set of vectors involved in protein translocation expressing chimeric DNA modifying proteins equipped with a translocation signal. For the construction of the binary vectors p14-2GKX and p14-PDC6KX, the following steps were taken. A *KanMX* marker gene flanked with *SpeI* and *XhoI* restriction sites was generated by PCR using pSDM8000 as DNA template [19]. This fragment was cloned into p2GU also digested with *SpeI* and *XhoI* resulting in p2GKX. The fragment containing the *KanMX* marker gene flanked with the *UAS7* and *UAS10* sequence was cloned into the binary vector pSDM14 [37] as a *KpnI*, *SalI* fragment resulting into p14-2GKX. The PCR primers used for the amplification of *KanMX* are shown in table 7. To assemble a binary vector with two *PDC6* flanks homologous to the *PDC6* sequences present in pID2GU and its derivatives, PCR was performed using the primers depicted in figure 5 using pINT as template. The resulting PCR product was digested with *BglII* and cloned into *BamHI* digested pSDM14, resulting in p14-PDC. A *KanMX* marker gene with TEF1 promoter and terminator was obtained from pSDM8000 as a *BamHI* fragment and was cloned in p14-PDC digested with *BamHI*, resulting into p14-PDC6KX. The binary vectors p14-2GKX and p14-PDC6KX were transferred to the *Agrobacterium* strain LBA1100 or LBA2556 by electroporation. LBA1100 is an C58C1 *Agrobacterium* strain carrying the disarmed octopine Ti-plasmid pTiB6 [38]. LBA2556 is a *virD2* null-deletion mutant, isogenic to LBA1100 [22].

**Table 5. Oligos designed for the amplification of the PDC6 locus and the KanMX marker gene.** The restriction sites added to the PCR primer sequences are underlined. Each oligo is written in the 5'-3' direction.

| Oligo name | Nucleotide sequence                        |
|------------|--------------------------------------------|
| PDC6FW     | TAAT <u>AGATCT</u> AGGTAAATAAATGTGCAGATGCA |
| PDC6RV     | TAAT <u>AGATCT</u> TGAACAAAGGGCGAACTTCGCA  |
| Kan Xho    | TAAT <u>CTCGAG</u> TTTAGCTTGCCTCGTC        |
| Kan Spe    | TAAT <u>ACTAGT</u> TTTCGACACTGGATGG        |

For the construction of all the fusion protein translocation vectors describe here, pBFF was used as the vector backbone. This non-transmissible vector was derived from the broad host-range plasmid pRL662 [39] and was adjusted to enable the translocation of a diversity of fusion proteins [16,18]. Cloning ORFs as *NotI* fragments into the pBFF vector puts protein expression under control of the *VirF* promoter sequence from the octopine typed Ti-plasmid pTiB6. When properly allowing for translational fusions, produced proteins will contain an N-terminal FLAG-tag for immunodetection, a SV40 nuclear localization signal (NLS) to ensure its nuclear entry and a 37 aa C-terminal end of *VirF* allowing for T4SS mediated translocation [15]. The vector pBFF I-SceI and pBFF *VirD2* I-SceI were generated by PCR-amplifying the I-SceI coding region (kind gift of Dr. Holger Puchta, University of Karlsruhe, Germany) and cloning it into the pBFF vector, thus supplying it with the features mentioned above.

## The construction of the yeast expression vectors

The yeast expression vectors were constructed by first cloning a DNA fragment encoding an ATG translational initiation codon followed by the *FLAG::NLS::virF<sup>CT</sup>* sequence into pRS425 ADH [23] resulting into pRS425 ADH *FLAG::NLS::virF<sup>CT</sup>*. New ORFs were subsequently introduced as *NotI* fragments within the same site present directly upstream of *virF<sup>CT</sup>*. The vector pRS425 *FLAG::NLS::virD2::I-SceI::virF<sup>CT</sup>* was thus constructed by insertion of a *NotI* fragment from pBFF VirD2-I-SceI and pRS425 *FLAG::NLS::virD2::Cre::virF<sup>CT</sup>* by insertion of a *NotI* fragment from pBFF VirD2-CRE. The fusion proteins produced in *Agrobacterium* and yeast were derived from identical ORFs.

## Qualitative assay for enzymatic activity of DNA modifying fusion proteins in yeast

4 The pRS425ADH expression vectors encoding NLS-VirD2-I-SceI-VirF<sup>CT</sup> or NLS-VirD2-CRE-VirF<sup>CT</sup> were transferred to yeast cells carrying the different target loci by LiAc-transformation and transformants were selected for leucine autotrophy. After 3 or 4 days, 50 transformed yeast cells were removed from the selective plates and pooled in 500 µl of sterile 0.9% salt solution. The OD<sub>600</sub> of this cell suspension was adjusted to 0.1 and diluted 100-fold. Subsequently, 100 µl of this cell suspension was replated on SD medium selective for leucine and uracil autotrophy, in order to maintain yeast cells carrying an expression vector and an intact target locus. The same aliquot of cells was plated on SD medium selective for leucine autotrophy but containing uracil as well as 5-FOA. The surviving colonies represented the yeast cells that carried the expression vector but lost the uracil selection marker gene.

## Agrobacterium mediated transformation of yeast

For the AMT of yeast cells, we followed an earlier published method [40] with some modifications. Briefly, *Agrobacterium* cells were precultured on LB medium with appropriate antibiotic selection. The transition from medium containing glucose to medium containing galactose has no impact on the growth rate of the *Agrobacterium* strains described here. For reliable results, the yeast strain YPH250 and RSY12 required acclimatization to galactose before and during the preculture step for sufficient growth on IM plates with galactose as the sole carbon source. Before mixing the *Agrobacterium* cell suspension with the yeast cells, the *Agrobacterium* cells were washed with induction medium (IM) without any carbon source. All cocultivations were performed for one week at 21°C, without any antibiotic or autotrophic selection in order to avoid any experimental biases. Whenever galactose induction of the yeast cells was applied, the IM used for preculturing the yeast strains as well as the induction plates were prepared with galactose instead of glucose.

## References

1. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A* 79: 6322-6326.
2. Durrenberger 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.
3. Jasper F, Koncz C, Schell J, Steinbiss HH (1994) *Agrobacterium* T-strand production in vitro: sequence-specific cleavage and 5' protection of single-stranded DNA templates by purified VirD2 protein. *Proc Natl Acad Sci U S A* 91: 694-698.
4. Ward JE, Akiyoshi DE, Regier D, Datta A, Gordon MP, et al. (1988) Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid. *J Biol Chem* 263: 5804-5814.
5. Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J* 51: 779-791.
6. Puchta H (2002) Gene replacement by homologous recombination in plants. *Plant Mol Biol* 48: 173-182.
7. Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res* 21: 5034-5040.
8. Li J, Vaidya M, White C, Vainstein A, Citovsky V, et al. (2005) Involvement of KU80 in T-DNA integration in plant cells. *Proc Natl Acad Sci U S A* 102: 19231-19236.
9. Prieto J, Molina R, Montoya G Molecular scissors for in situ cellular repair. *Crit Rev Biochem Mol Biol* 47: 207-221.
10. Ozawa K A high-efficiency *Agrobacterium*-mediated transformation system of rice (*Oryza sativa* L.). *Methods Mol Biol* 847: 51-57.
11. Nishizawa-Yokoi A, Nonaka S, Saika H, Kwon YI, Osakabe K, et al. Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytol* 196: 1048-1059.
12. Qi Y, Zhang Y, Zhang F, Baller JA, Cleland SC, et al. Increasing frequencies of site-specific mutagenesis and gene targeting in *Arabidopsis* by manipulating DNA repair pathways. *Genome Res* 23: 547-554.
13. Liu PQ, Rebar EJ, Zhang L, Liu Q, Jamieson AC, et al. (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. *J Biol Chem* 276: 11323-11334.
14. Zhang L, Spratt SK, Liu Q, Johnstone B, Qi H, et al. (2000) Synthetic zinc finger transcription factor action at an endogenous chromosomal site. Activation of the human erythropoietin gene. *J Biol Chem* 275: 33850-33860.
15. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, et al. (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.

16. Schrammeijer B, den Dulk-Ras A, Vergunst AC, Jurado Jacome E, Hooykaas PJ (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.
17. Kregten Mv (2011) *Agrobacterium*-mediated delivery of a meganuclease into target plant cells. 104-121 p.
18. van Kregten M, Lindhout BI, Hooykaas PJ, 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.
19. van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J* 20: 6550-6558.
20. van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31: 826-832.
21. Bryant GO, Prabhu V, Floer M, Wang X, Spagna D, et al. (2008) Activator control of nucleosome occupancy in activation and repression of transcription. *PLoS Biol* 6: 2928-2939.
22. Jurado-Jacome E (2003) Translocation of the *Agrobacterium tumefaciens* VirD2 and VirE3 proteins into eukaryotic host cells. Thesis Chapter 2: 39-60.
23. Mumberg D, Muller R, Funk M (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156: 119-122.
24. Yu L, Morse RH (1999) Chromatin opening and transactivator potentiation by RAP1 in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19: 5279-5288.
25. Aslankoohi E, Voordeckers K, Sun H, Sanchez-Rodriguez A, van der Zande E, et al. Nucleosomes affect local transformation efficiency. *Nucleic Acids Res* 40: 9506-9512.
26. Loyter A, Rosenbluh J, Zakai N, Li J, Kozlovsky SV, et al. (2005) The plant VirE2 interacting protein 1. a molecular link between the *Agrobacterium* T-complex and the host cell chromatin? *Plant Physiol* 138: 1318-1321.
27. Zhu Y, Nam J, Humara JM, Mysore KS, Lee LY, et al. (2003) Identification of *Arabidopsis* rat mutants. *Plant Physiol* 132: 494-505.
28. Lacroix B, Loyter A, Citovsky V (2008) Association of the *Agrobacterium* T-DNA-protein complex with plant nucleosomes. *Proc Natl Acad Sci U S A* 105: 15429-15434.
29. Magori S, Citovsky V Epigenetic control of *Agrobacterium* T-DNA integration. *Biochim Biophys Acta* 1809: 388-394.
30. 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.
31. Neuteboom LW, Lindhout BI, Saman IL, Hooykaas PJ, van der Zaal BJ (2006) Effects of different zinc finger transcription factors on genomic targets. *Biochem Biophys Res Commun* 339: 263-270.
32. Meijer AH, Ouwerkerk PB, Hoge JH (1998) Vectors for transcription factor cloning and target site identification by means of genetic selection in yeast. *Yeast* 14: 1407-1415.
33. Hohmann S (1991) Characterization of PDC6, a third structural gene for pyruvate decarboxylase in *Saccharomyces cerevisiae*. *J Bacteriol* 173: 7963-7969.

34. Hohmann S (1991) PDC6, a weakly expressed pyruvate decarboxylase gene from yeast, is activated when fused spontaneously under the control of the PDC1 promoter. *Curr Genet* 20: 373-378.
35. Schiestl RH, Petes TD (1991) Integration of DNA fragments by illegitimate recombination in *Saccharomyces cerevisiae*. *Proc Natl Acad Sci U S A* 88: 7585-7589.
36. Gietz RD, Schiestl RH, Willems AR, Woods RA (1995) Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast* 11: 355-360.
37. Offringa R (1992) Homologous recombination between T-DNAs after co-introduction into tobacco protoplasts via *Agrobacterium*. Thesis Chapter 2: 29-47.
38. Beijersbergen A, Dulk-Ras AD, Schilperoort RA, Hooykaas PJ (1992) Conjugative Transfer by the Virulence System of *Agrobacterium tumefaciens*. *Science* 256: 1324-1327.
39. Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, et al. (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290: 979-982.
40. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *Embo J* 14: 3206-3214.



# Chapter 5

## General discussion

## Introduction

*Agrobacterium tumefaciens* is a Gram-negative phytopathogen which is known to be the causal agent of crown gall disease in a diversity of dicotyledonous plants. Crown gall is characterized by the formation of tumors, usually close to the root crown. Virulent *Agrobacterium* strains are always carrier of a tumor-inducing (Ti) plasmid [1,2]. A section of this plasmid, the transfer DNA (T-DNA) is delimited by two specific DNA sequences: the *right border* (RB) and the *left border* (LB) [3]. *Agrobacterium* is able to perceive sugars and phenolic compounds that form indicators for damaged plant cells leading to the production of a range of virulence (Vir) proteins [4,5]. The T-DNA is nicked by the relaxase VirD2, facilitated by VirD1, at both the RB and LB sequence [6]. VirC1 enhances the nicking of the T-DNA by interacting with VirD1, VirD2 and VirC2 [30]. As a result of the expression of Vir proteins, a single stranded copy of the T-DNA (T-strand) is formed. For this, VirC1 binds to *overdrive* (OD), a DNA sequence located upstream the T-DNA RB [28, 29]. The T-strand remains covalently attached to VirD2, and is transferred via a type four secretion system (T4SS) to the recipient cell. After translocation from the *Agrobacterium* cell to the plant cell, the T-strand can integrate into the genomic DNA of the host cell.

5 Prior to genomic integration, the T-strand must be targeted to the nucleus. Several proteins are involved in the nuclear import of the T-strand. The effector protein VirE2 binds to the T-strand in the recipient cell where it covers the entire length of the T-strand thereby protecting it from degradation by plant nucleases [7-9]. The nucleoprotein complex that is formed in the plant cytosol from the T-strand and VirE2 is termed the T-complex. In *Arabidopsis thaliana* it has been shown that VirE2 also interacts with VIRE2 BINDING PROTEIN 1 (VIP1) [10]. VIP1 is a transcription factor that can be targeted for nuclear import by the mitogen-activated kinase (MAPK) MPK3 that is part of the plant cell defense mechanism against pathogens [11].

Although the natural host range of *Agrobacterium* is restricted to the plant kingdom, a number of research groups have managed to apply AMT to several other organisms ranging from *Streptomyces*[12], yeast[13], filamentous fungi[14] to sea urchin[15]. Because of its short generation time allowing for easy transformant screening and analysis, the AMT of yeast has developed into a highly valued tool to study the role of host factors in the T-DNA integration process [16].

### Towards T-DNA processing and integration in *Saccharomyces cerevisiae*

Cocultivation studies using *Saccharomyces cerevisiae* as a model organism resulted in important new insights regarding the involvement of the DNA repair mechanism in T-DNA integration [16,17]. Adding to this, we started the development of an all-in-yeast system, to allow for a more detailed study of T-DNA processing, T-strand formation, and T-DNA integration. Such a model system would enable a more thorough and reliable screening for host components important for T-DNA integration than is possible when

relying on the AMT of yeast. The protocol for the AMT of yeast introduces variables that can influence the AMT efficiency, like the humidity of the plates or the adhesion of the *Agrobacterium* cells to the yeast cells that could obscure the variables that are the actual target of the experiment. Therefore an all-in-yeast system that removes many hard-to-control variables from the equation could fully employ the readily available collections of yeast strains harboring individual knockout mutations for annotated genes.

For this purpose, a vector harboring an artificial T-DNA substrate, pArT1, as well as two galactose inducible virulence gene expression cassettes were prepared equipped with ORFs expressing the *Agrobacterium* Vir proteins VirD2, VirD1, VirC2, VirC1 and VirE2, as well as the *Arabidopsis* ORF encoding the *Arabidopsis* transcriptional activator VIP1. The latter protein might increase the nuclear targeting of VirE2 in yeast since for yeast cells it has been shown that VirE2 needs to be co-expressed with VIP1 to induce nuclear uptake of VirE2 [18]. These genes were collected as they all are known to function in T-DNA processing or in the nuclear targeting of processed T-DNA. Initially, a fully plasmid-based system for the processing and integration of T-DNA was assembled. The already available mutant yeast lines could then simply be equipped with the newly developed plasmids.

Initially, a fully plasmid-based system for the processing and integration of T-DNA was assembled. The already available mutant yeast lines could then simply be equipped with the newly developed plasmids. To provide the yeast cells with a vector to serve as an artificial T-DNA-like substrate (pArT1) the *ARS/CEN* yeast shuttle vector pRS313 [19] was equipped with an RB and LB sequence, where the RB was preceded by an *overdrive* (OD) sequence. A promoterless KanMX marker gene was placed in between these border sequences. After the VirD2 mediated nicking of both border sequences, the promoterless KanMX ORF could then be released as a T-strand. Integration of this T-strand into a transcriptionally active region of the yeast genome should reconstitute KanMX marker gene expression and thus provide us with a selectable trait for T-DNA integration events. Prior testing for a possible level of kanamycin (G418) resistance in pArT1 containing cells by plating cells on media with different G418 concentrations indicated that cells were killed above a concentration of 50 µg/µl. When performing experiments with larger cell numbers and with two plasmids encoding the Vir proteins it appeared that up to 1-3 % of the plated cells were resistant for G418 concentrations of up to 200 µg/µl. Further analysis of the obtained G418 resistant colonies revealed that the G418 resistance was pArT1 plasmid-based and was not due to insertion of T-DNA into the yeast genome. The acquired G418 resistance was independent of the expression of virulence genes but did correlate with the presence of additional *ARS/CEN* type vectors besides pArT1, even when these plasmids did not carry any inducible ORFs. Since plasmid instability caused by the simultaneous usage of several plasmids apparently frequently resulted into the stochastic expression of the KanMX marker on pArT1 another approach was required. To reduce the occurrence of G418 resistance, two measures were taken.

Firstly, an extra piece of stopcodon-rich DNA was inserted into pArT1, just upstream the RB sequence. This modification resulted in the vector pArT1stop which was supposed to prevent any translational read through into the KanMX ORF from upstream translation starts. Transcription at this area of the plasmid could be enhanced by unforeseeable epigenetic changes. Secondly, the virulence gene expression cassettes were integrated into the yeast genome, thus preventing the presence of similar plasmids that might recombine with pArT1stop. These measures strongly reduced the background levels of G418 resistance to values in principle low enough to detect even very small positive contributions of Vir proteins to T-DNA integration. In that manner, it seemed to be very promising that the conditions allowing for the expression of VirD2 resulted into increased numbers of cells that acquired G418 resistance compared to the empty vector control (Figure 11, Chapter 2). The addition of ORFs encoding other virulence proteins, including the VirD2 assisting protein VirD1 [6], did not further elevate the frequency of G418 resistance.

A Southern blot of DNA derived from G418 resistant yeast colonies revealed additional bands that might be indicative for potential T-DNA integration events (Chapter 2, Figure 11). However, it finally turned out that the G418 resistance was plasmid-borne in all cases; transferring the plasmids isolated from the G418 resistant strains to a fresh WT yeast strain always resulted into new G418 resistance and Southern blot analysis of genomic DNA isolated from cured plasmid-free strains proved that the KanMX ORF had disappeared. Evidence for genomic integration of T-DNA produced in yeast had thus not been obtained.

The molecular explanation regarding the origin of G418 resistance found in the experiments with pArT1stop and integrated expression cassettes has remained unclear. It is tempting to speculate that the association of VirD2 with single stranded border sequences during plasmid replication challenges accurate DNA replication, resulting in the generation of aberrant plasmids which then can express the KanMX gene. Since any further insights regarding the source of these events would not address the research questions we originally started out with, it seemed not prudent to further analyze the plasmids that conferred G418 resistance. Instead, it was concluded that the tools developed thus far did not allow for a sufficiently high level of all-in-yeast T-DNA production to support the intended investigations regarding the role of yeast factors in T-DNA integration

Not all parameters that are of importance to the development of an all-in-yeast T-DNA production system could be investigated in detail; by lack of antibodies it was for instance not possible to verify the expression of unmodified Vir proteins in yeast. However, GFP fusions demonstrated that all expression cassettes were in principle active. In the cases that there could be a problem due to a clear difference in codon usage between *Agrobacterium* and yeast, yeast optimized codons were inserted. The minimal border sequences used in Chapter 2 proved to support a high level of DNA transfer, comparable

to the larger border sequences that were used for the cocultivation studies described in Chapter 3. It is conceivable that changes in orientation and position of border sequences in respect to the origin of plasmid replication could affect the possibility for VirD2-mediated border processing. Results from experiments addressing this issue, showed changes in the background frequency of G418 resistance, but no evidence was found indicating that successful T-DNA formation and integration had occurred (data not shown). One other experiment that was performed led to the conclusion that it is unlikely that *in yeast* border processing by *Agrobacterium* Vir proteins takes place while using plasmids equipped with an *ARS/CEN* replication origin as a substrate. This experiment will be addressed below in the paragraph discussing Chapter 3 “Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens* mediated transformation of *Saccharomyces cerevisiae*”.

Hence, without clear cues on how to improve upon the construction of an effective and flexible all-in-yeast T-DNA processing and integration system, it was decided to focus on cocultivation based studies instead to get more insights in particular aspects of AMT. During these studies it was possible to use tools and insights that were obtained during the attempts to generate the all-in-yeast system.

### **Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens* mediated transformation of *Saccharomyces cerevisiae***

In a pioneering study performed by our laboratory circular T-DNAs could be recovered after AMT of yeast cells [13]. In a recent publication, similar T-DNA circles were isolated from *Agrobacterium* infested plant tissue [20]. The recovery of T-DNA circles from yeast cells as well as from plant cells is indicative that T-DNA circle formation might be a general trait of AMT in addition to genomic T-DNA integration. From a biological point of view, the formation of T-DNA circles could benefit *Agrobacterium*, giving it the opportunity to express genes from the T-DNA circles even before genomic integration of the T-DNA is accomplished. Since the T-DNA circles do not contain free 5' and 3'-ends, the T-DNA should be insensitive to exonuclease mediated degradation in the recipient host cell. The formation of T-DNA circles could also be of great interest for plant biotechnologists that are searching for new techniques that allow transient expression of genes in plant cells without the requirement to disrupt the genome.

Here, we investigated the impact of the presence as well as the type of T-DNA border sequences on a binary vector upon the occurrence of T-DNA circles. Additionally, we investigated the effects of mutations in the DNA repair mechanisms present in the recipient yeast cells, since previous research demonstrated that genomic integration of T-DNA can be mediated via the different DNA repair mechanisms [16,17]. Involvement of the DNA repair machinery could very well also be relevant to T-DNA circle formation. A new series of binary vectors was constructed, the pOphis series, all provided with a yeast *ARS/CEN* origin of replication.

Any circular T-DNA that would be formed in yeast could thus be maintained as a chromosomal type of plasmid. The pOphis series consisted, among others, of pOphis RB LB, pOphis RB, pOphis RB RBLB and a borderless vector pOphis.

The results that were obtained using the different yeast genotypes clearly indicated that the homologous repair (HR) DNA repair proteins Rad52 and Rad51 are involved in T-DNA circle formation, but not the nonhomologous end-joining (NHEJ) protein Ku70 (Figure 3 of Chapter 3). At first sight, these results were surprising since most of the pOphis vectors that were used did not contain sufficiently long regions of internal homology to support HR or no internal homology at all. However, these results would make perfect sense when multimeric T-DNA structures are formed, either in *Agrobacterium* or in yeast, which are consequently resolved into T-DNA circles by the HR pathway of the recipient cell. Current models for conjugative plasmid transfer assume that a rolling circle replication-like mechanism in the donor strain feeds the T4SS with single stranded DNA [21-24]. Experiments were therefore performed to test if such a mechanism could underlie the formation of concatemeric T-DNA structures. The obtained results indicated that following AMT of all plasmids, thus pOphis RB, pOphis RB LB, pOphis RB RB, and pOphis RB RBLB, a similar reduction of 70-75% in transformation efficiency was found when comparing AMT with *rad52Δ* recipient cells to AMT with the congenic WT strain. After analysis of T-DNA circles that were rescued from transformed yeast cells, it appeared as if pOphis RB was always translocated in its entirety while for vectors with two or more borders the transfer was mostly restricted to the region between two border sequences (Figure 6, Chapter 3). Since the transformation efficiencies of yeast cells that received partial vectors and complete vectors were similarly reduced by the *rad52Δ* mutation, therefore both types of entering T-strands were supposedly concatemeric. This observation prompted us to discard the rolling circle replication-like mechanism as a source for T-DNA concatemers since a rolling circle replication-like mechanism cannot be the source of multimers of partial vector sequences.

Sequence analysis of the RB area of T-DNA circles rescued from transformed yeast cells (Figures 7 and 8, Chapter 3) showed that in virtually all cases perfect border fusions were present, as if joined perfectly by VirD2. Interestingly, in addition to the reconstitution of RB sequences, also processed RB and LB sequences were found to be perfectly fused. Since the NHEJ reaction mediated by Ku70 was shown to be of negligible importance to T-DNA circle formation, a model for T-DNA circle formation was constructed wherein T-DNA concatemers are formed by a chaining reaction catalyzed by VirD2 and are subsequently resolved by the HR pathway of yeast into T-DNA circles (Figure 9, Chapter 3). As a follow up to the results of Chapter 2 “Towards T-DNA processing and integration in *Saccharomyces cerevisiae*”, pOphis RB LB was modified to test if *in yeast* T-DNA processing mediated by *Agrobacterium* Vir proteins could occur. By the addition of an *URA3* gene to section B of pOphis RB LB, it should be possible to select for a Vir protein mediated processing event.

After recirculation of the T-strand derived from pOphis RB LB a smaller plasmid could be reconstituted lacking the *URA3* gene. A positive selection for the formation of these smaller circles is available in the form of 5-FOA that selects against cells that produce uracil. A major benefit of this approach is that for the detection of T-DNA processing it does not depend on the integration of T-DNA in transcriptional active genomic DNA. A proof of principle experiment yielded very few yeast colonies that apparently had inactive uracil synthesis pathways (data not shown). However, a slightly higher amount of Ura auxotrophic colonies was also retrieved from the negative controls that did not express any *Agrobacterium* virulence genes. The latter observation prompted us to discard the idea that yeast episomes equipped with *Agrobacterium* border sequences form a suitable substrate for virulence protein mediated T-DNA processing. As was explained previously in the discussion of Chapter 2, differences in the plasmid topology between yeast and prokaryotic cells like *Agrobacterium* and *E. coli* [25] could account for the observation that T-DNA present on pOphis RB LB is processed in *Agrobacterium* but no prove for T-DNA processing was obtained using yeast.

### **Enhanced gene targeting mediated by translocated I-SceI during the *Agrobacterium* mediated transformation of yeast**

Random integration of the T-DNA is a natural consequence of using AMT for plant transformation [26]. The inherent unpredictability of the future integration site of the translocated transgene is commonly regarded as a major limitation of AMT, as it can lead to disruption of host genes or to the presence of undesirable regulatory elements at the genomic integration site. To avoid random integration of T-DNA, the HR system, that is naturally present in plants, might offer a means to target T-DNA to a predetermined genomic locus that comprises a strong homology to the incoming T-strand. Unfortunately, HR-mediated integration of DNA molecules are usually rare events in plants, estimated to occur once per  $10^4$  to  $10^5$  integration events [27,28]. Fortunately, the levels of HR can be boosted by the induction of a DNA double strand break (DSB) at the target locus [29-31].

To further characterize possible variables regarding targeted T-DNA integration, the effect of different chromatin states was investigated. Additionally, it was assessed if DSBs can be induced during cocultivation by the co-delivery of fusion proteins containing I-SceI nuclease domains to facilitate HR mediated T-DNA integration during AMT. The latter would be a very useful addition to the techniques that are currently used to induce DSBs to achieve T-DNA targeting via HR in plants since no extra DNA encoding such a nuclease needs to be present in the recipient cell.

The obtained results indicated that a decrease in the nucleosome occupancy of the target locus had no measurable impact on the efficiency of T-DNA integration at this locus.

However, the combined translocation of nuclease proteins and T-strands during AMT greatly facilitated HR-mediated integration of T-DNA (Figure 5, Chapter 4). In cases where *Agrobacterium* mediated translocation of specific nucleases for DSB induction at the target locus occurred to cells with an elevation of the nucleosome occupancy of that target locus, it was found that the presence of nucleosomes severely reduced the positive effect of the nuclease on the transformation efficiency. Apparently the nucleosome occupancy influenced the level of access that was granted to the nuclease to the target locus.

Surprisingly, during the course of performing the experiments mentioned above, it was also found that the agent that was used to induce nucleosome removal from the target locus, i.e. growth on galactose instead of glucose, had a severe impact on the AMT efficiency of the yeast strain YPH250 (Figure 3, Chapter 4). Especially T-DNA integration via the NHEJ pathway was more efficient if the cocultivation was performed on galactose containing agarose plates. The yeast strain RSY12 proved to be largely unaffected by this “galactose effect” and was therefore used to perform the experiments that required manipulation of the nucleosome occupancy of the target locus. Especially because growth on glucose resulted in an improvement on the AMT efficiency of the NHEJ binary vector pSDM8000, which usage usually results in very low transformation efficiencies, a deeper knowledge of the mechanisms behind the “galactose effect” might be beneficial. The T-DNA integration itself seemed not to be of a different nature than AMT of yeast performed on glucose plates: sequencing data of five inverse PCR products showed a random integration pattern typical for T-DNA integration. When comparing growth curves of RSY12 and YPH250 on galactose, YPH250 is clearly disadvantaged, hardly showing any growth on galactose medium. However the growth curves for glucose were similar between the two yeast strains.

Possibly difficulties in the uptake of galactose forms the basis for the observed differences between the AMT of YPH250 and the AMT of RSY12. There are also some indications from literature that the yeast strain S288C, which is ancestral to YPH250, had a defect in the galactose permease *GAL2* [32]. An alignment of the amino acid sequence of the galactose transporter protein *GAL2* from RSY12 and YPH250 showed that two amino acids were varying between the two strains, of which one was predicted to be located at the inside of the Gal2 pore. Possibly, these variations influence the capacity of these yeast strains to gather galactose from their environment. Because of time limitations no experiments have been performed to link the mutations in the YPH250 *GAL2* gene to the ability to grow on galactose medium. Although not supported by solid evidence, the reduced growth we observed for YPH250 on galactose medium could be instrumental for the enhanced levels of NHEJ mediated T-DNA integration we observed. Possibly starvation has a yet unknown impact on some of the mechanisms important for T-DNA-integration like the nuclear uptake of DNA or DNA repair.

## Double stranded DNA formation as a limiting factor for transkingdom AMT

Although useful and reliable protocols have been developed for the AMT of yeast [13] and filamentous fungi [14], the transformation efficiencies that could be achieved have always remained limited. This is especially true if the T-DNA is forced to integrate via the NHEJ pathway, thus in cases of absence of sequence homology between the T-DNA and the genome of the recipient cell [16,17].

Several steps of the AMT pathway could be considered to be potentially rate limiting, already starting with the adhesion of *Agrobacterium* to the intended recipient cells. Observations at the start of AMT of yeast cells on a cocultivation plate, using a confocal microscope show that only a small fraction of the yeast cells is actually bound to the excess of *Agrobacterium* cells that is usually present (P.A. Sakalis, personal communication). Later on in the AMT process, it can be envisaged that one or more of the different Vir proteins that are transported to recipient cells might not exert the required AMT supporting function, either due to limiting concentration or to lack of a required interaction partner within the recipient cell. Considering the transferred effector proteins VirE2, VirE3, VirD5 and VirF [33,34], an example of the latter limitation is the absence of a yeast homolog of VIP1 from *Arabidopsis thaliana* [18]. Another example of a deficiency in a recipient cell could be the observation that the absence or presence of the virulence protein VirF bears no impact on the AMT efficiency in yeast [13] although VirF proved to be of importance to the AMT of tomato and a number of other plant species [35-37]. Strikingly, although considered of vital importance to the AMT of plants, the absence of some of the natural interactors of *Agrobacterium* effector proteins in yeast cells [18,38-40] does apparently still allow for nuclear import of the T-strand or T-complex and the consequent genomic T-DNA integration [13].

Apart from the issues regarding the Vir proteins and potential interactors mentioned above, several thus far unnoticed aspects of AMT might provide ample opportunities to improve upon AMT efficiency. The galactose induced elevation of NHEJ-dependent AMT shown in Figure 3 of Chapter 4, could very well be an example of such a hitherto unknown experimental variable. Furthermore, during the finalizing steps that are required for the T-DNA integration mediated by the DNA repair mechanisms of the recipient cell there could be multiple moments where very important determinants for the efficiency of AMT need to be operational. Support for the latter hypothesis might be found when further exploring the phenomenon that when performing AMT of yeast using the KanMX marker gene for the selection of transformants often many miniscule colonies are visible, besides the larger colonies which clearly represent stable G418 resistant transformants. Although for lack of time these smaller colonies were never thoroughly analyzed, it might very well be that they represent cells that only transiently obtained G418 resistance. This would mean that a T-strand, containing the KanMX marker gene, has entered these yeast cells but was degraded before it was incorporated into the genome.

Nevertheless, due to transient expression of double stranded T-DNA in the original recipient cell as well as some of its descendants, growth of a small colony might be supported around relatively few G418 detoxifying cells. In that manner, the observation of minuscule colonies during experiments aimed at T-DNA integration via the NHEJ-pathway, might indicate that in the majority of cases the incoming T-strands are expressed but do not end up in the genome, thereby pointing out that instead of T4SS mediated translocation or the proper functioning of effector proteins, it are the final steps that take place just prior to the genomic incorporation of the T-DNA that are apparently rate limiting.

A comparison of all the transformation efficiencies resulting from the different experimentations described in the different Chapters of this thesis shows that the highest transformation efficiencies were generally obtained with the *ARS/CEN* typed T-DNA circle forming binary vectors that are described in Chapter 3. AMT with these binary vectors is more efficient than when using binary vectors that can integrate via HR and are even several orders of magnitude more efficient than when using binary vectors that must rely on NHEJ for integration of T-DNA. Strikingly, the smaller colonies mentioned above were generally absent from the transformant selection plates used with T-DNA circle forming vectors. This could indicate that when incoming T-strands can be maintained as episomal structures, AMT frequency would approach the maximal potential as far as it is related to the frequency of successful contact and T4SS mediated transport of macromolecules. For AMT of yeast, this frequency is between 1 and 5 transformants per 1000 yeast cells under the most optimal conditions thus far identified.

For the repair of DSBs via the HR or the NHEJ DNA repair machinery, the synthesis of a complementary DNA strand is required. Since the involvement of the host DNA repair mechanisms in T-DNA integration is well documented [16,17,31] I would like to postulate that the formation of a complementary strand is a prerequisite for T-DNA integration, thereby forming an important rate limiting factor in the AMT of yeast and possibly also other organisms. The *ARS/CEN* typed binary vectors would then have an advantage since they are equipped with a yeast replication origin and are thereby automatically a target for yeast primases leading to an efficient conversion to dsDNA.

## Future perspectives of the application of genetically modified plants

Much has been said and written about the application of AMT of plants in plant biotechnology. Meanwhile, AMT has become the method of choice for the genetic modification of crops like corn and soybeans. In the United States, genetically modified foods have been on the menu for over 16 years. How much more can be expected from this small soil born bacterium, more than a 100 years after *Agrobacterium tumefaciens* was first described in Science by Erwin Smith and Charles Townsend ?[41] With further research, AMT could be put to use to meet some of the needs that are expected to become more prevalent in a future with a growing world population and limited resources.

Some of these applications that are worth looking forward to, are the development of drought resistant plants that can be grown at localities now unavailable for agriculture and the optimization of the maximal amount of yieldable biomass to make full use of the available agricultural lands. From an environmental point of view, the development of pest resistant plants would also be beneficial since such plants could effectively reduce the amount of pesticides that are used now. The sympathetic “Golden rice” project, that resulted in genetically engineered rice plants that produce beta-carotene, a precursor of vitamin A, proved that it is possible to perform very challenging research without a commercial drive [42]. However, the strict regulations concerning the application of genetically engineered plants have severely delayed the application of this technology [43].

Looking at the potential of applying genetically modified (GM) plants to solve actual problems, one could wonder why research institutes that could be developing new applications of GM plant lines are mostly focusing on other technologies instead. Possibly, the negative associations that both the general public and policymakers have regarding GM plants play a major part in this. The few GM crops that are put to use were developed purely for commercial reasons, adding little to the general interest. However, the general reluctance to embrace this technology cannot simply be disposed of as a critique on commercialism or as technophobia. There can be a strong willingness to accept new technology or commercialism if the resulting products bear sufficient attraction to the consumer. The wide range of new technologies that led to the development of smartphones had no difficulties at all to become widely accepted and even admired. Adding to this view, an example of a GM organism that did receive considerable popularity is the “GloFish” from Yorktown Technologies, a transgenic fluorescent fish that comes in several bright colors and is available in pet shops throughout the United States. Clearly, apparent usefulness is not a relevant selling point for new technology.

By performing more research on the subject of AMT of plants, with a special focus on potential problem solving applications, it will become apparent that when it comes to the application of GM plant lines, it is prudent to go against popular beliefs and fears in order to accomplish something very worthwhile.

## References

1. Van Larebeke N, Engler G, Holsters M, Van den Elsacker S, Zaenen I, et al. (1974) Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature* 252: 169-170.
2. Watson B, Currier TC, Gordon MP, Chilton MD, Nester EW (1975) Plasmid required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol* 123: 255-264.
3. Yadav NS, Vanderleyden J, Bennett DR, Barnes WM, Chilton MD (1982) Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A* 79: 6322-6326.
4. Stachel SE, Zambryski PC (1986) *virA* and *virG* control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell* 46: 325-333.
5. Ankenbauer RG, Nester EW (1990) Sugar-mediated induction of *Agrobacterium tumefaciens* virulence genes: structural specificity and activities of monosaccharides. *J Bacteriol* 172: 6442-6446.
6. 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.
7. Das A (1988) *Agrobacterium tumefaciens virE* operon encodes a single-stranded DNA-binding protein. *Proc Natl Acad Sci U S A* 85: 2909-2913.
8. Citovsky V, Wong ML, Zambryski P (1989) Cooperative interaction of *Agrobacterium VirE2* protein with single-stranded DNA: implications for the T-DNA transfer process. *Proc Natl Acad Sci U S A* 86: 1193-1197.
9. Christie PJ, Ward JE, Winans SC, Nester EW (1988) The *Agrobacterium tumefaciens virE2* gene product is a single-stranded-DNA-binding protein that associates with T-DNA. *J Bacteriol* 170: 2659-2667.
10. Tzfira T, Rhee Y, Chen MH, Kunik T, Citovsky V (2000) Nucleic acid transport in plant-microbe interactions: the molecules that walk through the walls. *Annu Rev Microbiol* 54: 187-219.
11. Djamei A, Pitzschke A, Nakagami H, Rajh I, Hirt H (2007) Trojan horse strategy in *Agrobacterium* transformation: abusing MAPK defense signaling. *Science* 318: 453-456.
12. Kelly BA, Kado CI (2002) *Agrobacterium*-mediated T-DNA transfer and integration into the chromosome of *Streptomyces lividans*. *Mol Plant Pathol* 3: 125-134.
13. Bundock P, den Dulk-Ras A, Beijersbergen A, Hooykaas PJ (1995) Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *Embo J* 14: 3206-3214.
14. de Groot MJ, Bundock P, Hooykaas PJ, Beijersbergen AG (1998) *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat Biotechnol* 16: 839-842.
15. Bulgakov VP, Kiselev KV, Yakovlev KV, Zhuravlev YN, Gontcharov AA, et al. (2006) *Agrobacterium*-mediated transformation of sea urchin embryos. *Biotechnol J* 1: 454-461.
16. van Attikum H, Hooykaas PJ (2003) Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 31: 826-832.

17. van Attikum H, Bundock P, Hooykaas PJ (2001) Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J* 20: 6550-6558.
18. Tzfira T, Vaidya M, Citovsky V (2001) VIP1, an *Arabidopsis* protein that interacts with *Agrobacterium* VirE2, is involved in VirE2 nuclear import and *Agrobacterium* infectivity. *Embo J* 20: 3596-3607.
19. 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.
20. Singer K, Shibolet Y, Li J, Tzfira T (2012) Formation of Complex Extrachromosomal T-DNA Structures in *Agrobacterium tumefaciens*-Infected Plants. *Plant Physiol* 160: 511-522.
21. de la Cruz F, Frost LS, Meyer RJ, Zechner EL Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* 34: 18-40.
22. Llosa M, Gomis-Ruth FX, Coll M, de la Cruz Fd F (2002) Bacterial conjugation: a two-step mechanism for DNA transport. *Mol Microbiol* 45: 1-8.
23. Lanka E, Wilkins BM (1995) DNA processing reactions in bacterial conjugation. *Annu Rev Biochem* 64: 141-169.
24. Becker EC, Meyer R Origin and fate of the 3' ends of single-stranded DNA generated by conjugal transfer of plasmid R1162. *J Bacteriol* 194: 5368-5376.
25. Mayan-Santos MD, Martinez-Robles ML, Hernandez P, Krimer D, Schwartzman JB (2007) DNA is more negatively supercoiled in bacterial plasmids than in minichromosomes isolated from budding yeast. *Electrophoresis* 28: 3845-3853.
26. Kim SI, Veena, Gelvin SB (2007) Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J* 51: 779-791.
27. Offringa R, de Groot MJ, Haagsman HJ, Does MP, van den Elzen PJ, et al. (1990) Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *Embo J* 9: 3077-3084.
28. Puchta H (2002) Gene replacement by homologous recombination in plants. *Plant Mol Biol* 48: 173-182.
29. de Pater S, Pinas JE, Hooykaas PJ, van der Zaal BJ ZFN-mediated gene targeting of the *Arabidopsis* protoporphyrinogen oxidase gene through *Agrobacterium*-mediated floral dip transformation. *Plant Biotechnol J* 11: 510-515.
30. Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res* 21: 5034-5040.
31. Li J, Vaidya M, White C, Vainstein A, Citovsky V, et al. (2005) Involvement of KU80 in T-DNA integration in plant cells. *Proc Natl Acad Sci U S A* 102: 19231-19236.
32. Mortimer RK, Johnston JR (1986) Genealogy of principal strains of the yeast genetic stock center. *Genetics* 113: 35-43.
33. Vergunst AC, van Lier MC, den Dulk-Ras A, Stuve TA, Ouwehand A, et al. (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.
34. Vergunst AC, Schrammeijer B, den Dulk-Ras A, de Vlaam CM, Regensburg-Tuink TJ, et al. (2000) VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science* 290: 979-982.

35. Hooykaas PJ, Hofker M, den Dulk-Ras H, Schilperoort RA (1984) A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid, and an Ri plasmid by complementation analysis of *Agrobacterium tumefaciens* mutants. *Plasmid* 11: 195-205.
36. Schrammeijer B, Risseuw E, Pansegrau W, Regensburg-Tuink TJ, Crosby WL, et al. (2001) Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. *Curr Biol* 11: 258-262.
37. Regensburg-Tuink AJ, Hooykaas PJ (1993) Transgenic *N. glauca* plants expressing bacterial virulence gene *virF* are converted into hosts for nopaline strains of *A. tumefaciens*. *Nature* 363: 69-71.
38. Tzfira T, Vaidya M, Citovsky V (2004) Involvement of targeted proteolysis in plant genetic transformation by *Agrobacterium*. *Nature* 431: 87-92.
39. 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.
40. 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.
41. Smith EF, Townsend CO (1907) A Plant-Tumor of Bacterial Origin. *Science* 25: 671-673.
42. Beyer P, Al-Babili S, Ye X, Lucca P, Schaub P, et al. (2002) Golden Rice: introducing the beta-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *J Nutr* 132: 506S-510S.
43. Potrykus I Lessons from the 'Humanitarian Golden Rice' project: regulation prevents development of public good genetically engineered crop products. *N Biotechnol* 27: 466-472.





# Chapter 6

## Nederlandse samenvatting

*Agrobacterium tumefaciens* is een Gram-negatieve plant pathogene bacterie, bekend als veroorzaker van kroongal (wortelhalsknobbel) in diverse soorten dicotyle planten. Kroongal wordt gekarakteriseerd door de vorming van tumoren, meestal nabij de wortelhals. Virulente *Agrobacterium* stammen zijn altijd drager van een tumor-inducerend (Ti) plasmide [1, 2]. Een plasmide is een circulair stuk DNA dat onafhankelijk van het chromosomaal DNA wordt gerepliceerd. Een sectie van het Ti plasmide, het transfer DNA (T-DNA) wordt begrensd door twee specifieke DNA sequenties, de rechter border (RB) en de linker border (LB) [3]. Wanneer *Agrobacterium* suikers en fenol-achtige componenten waarneemt, die voor deze bodembacterie een indicator vormen voor beschadigde plantencellen [4, 5], wordt een enkele streng van het T-DNA vrijgemaakt van het Ti plasmide. Dit wordt bewerkstelligd door het relaxase VirD2 die hierin gefaciliteerd wordt door VirD1 [6]. VirD2 katalyseert de vorming van een enkelstrengsbreuk in zowel de RB als de LB. De hieruit resulterende DNA-streng, die aan één zijde gekoppeld blijft aan VirD2, wordt de T-streng genoemd. De T-streng kan worden overgedragen naar een gastheercel om vervolgens opgenomen te worden in het genomische DNA dat zich in de celkern bevindt. Deze *Agrobacterium* gemedieerde transformatie (AMT) zet plantencellen aan tot de productie van opines, gemodificeerde aminozuren die *Agrobacterium* kan gebruiken als een energierijke koolstof- en stikstofbron [7]. Ook worden de plantencellen aanzet om de productie van specifieke plantenhormonen te verhogen die op hun beurt getransformeerde plantencellen aanzetten tot verhoogde celdeling waardoor de totale productie van opines nog verder wordt opgevoerd.

6 Ongeveer 30 jaar geleden werd ontdekt dat het DNA tussen de RB en LB vervangen kan worden door een DNA fragment naar keuze wat vervolgens door de bacterie naar planten overgedragen kan worden [8]. De mogelijkheid om plantencellen op deze wijze te transformeren leidde tot bloei van de plantenbiologie en plantenbiotechnologie. Er werden speciaal voor dit doeleinde handige kleine plasmides gemaakt die een RB en LB bevatten om (trans)genen tussen te plaatsen [9]. Deze plasmiden werden binaire vectoren genoemd omdat zij voor de expressie van VirD2 en VirD1 en de andere Virulentie-eiwitten afhankelijk zijn van het Ti plasmide. Behalve het overdragen van nieuwe genen, is het ook mogelijk met AMT bestaande plantengenen uit te schakelen. Door het bestuderen van nieuw verkregen eigenschappen van genetisch gemodificeerde planten is er veel wetenschappelijke informatie beschikbaar gekomen over de werking van plantengenen. Behalve planten, bleek het ook mogelijk om onder gecontroleerde omstandigheden gist [10, 11] en filamenteuze schimmels [12] te transformeren met *Agrobacterium*. Door de korte generatietijd van gist en de mogelijkheid om efficiënt getransformeerde cellen te kunnen isoleren, opkweken en analyseren, kon er met behulp van het gistmodel onder andere gekeken worden naar de rol die gastheerfactoren spelen in de opname van T-DNA in het genetisch materiaal van de gastheer. Uit dergelijk onderzoek bleek dat de DNA reparatiemechanismen van de gastheer onontbeerlijk zijn voor de genomische opname van T-DNA [13, 14].

Een meer uitgebreid overzicht van de op dit moment bekende processen die leiden tot de productie, het transport en de uiteindelijke genomische integratie van het *Agrobacterium* T-DNA in de gastheercel is te vinden in **hoofdstuk 1**.

Om de resultaten die zijn beschreven in dit proefschrift te verkrijgen is veelvuldig gebruik gemaakt van het gistmodel. Voor de AMT van gist moeten gist en *Agrobacterium* cellen gezamenlijk worden gecultiveerd. Tijdens deze cocultivatperiode worden de virulentiegenen van *Agrobacterium* geïnduceerd door toevoeging van de natuurlijke fenolische verbinding acetosyringon. Voor sommige onderzoeksvragen zou het erg aantrekkelijk zijn wanneer de cocultivatie achterwege kon blijven omdat dit proces een aantal zeer lastig te controleren variabelen met zich meebrengt. Dit is zeker het geval bij het bestuderen van de rol van gastheerfactoren in de genomische integratie van T-DNA.

In **hoofdstuk 2** is daarom gezocht naar een mogelijk alternatief voor cocultivatie door te proberen een *Agrobacterium*-vrij systeem te genereren waarbij een selectie van *Agrobacterium* virulentiegenen in gist, gecontroleerd tot expressie zou kunnen worden gebracht. Door in die giststammen ook een plasmide te plaatsen voorzien van een artificieel T-DNA (pArT), zouden in principe T-strengen geproduceerd kunnen worden door de virulentiegenen tot expressie te laten komen. Deze T-strengen zouden dan weer elders kunnen integreren in het genoom. Het in het onderzoek gebruikte T-DNA bevatte een merker-gen dat inactief was gemaakt door het verwijderen van de promoter-sequentie die normaal gesproken de activiteit van het merker-gen bepaalt. T-strengen met dit inactief gemaakte merker-gen, zouden toevallig achter een promoter-sequentie in het gistgenoom kunnen integreren en daardoor zou het merker-gen weer actief worden, wat zou leiden tot kanamycine resistentie. Uit de resultaten bleek dat de expressie van VirD2 een positieve invloed had op het aantal gistcellen dat herstelde merker-gen activiteit vertoonden. Verdere analyse van deze resistente giststammen toonde echter aan dat het geactiveerde merker-gen zich niet in het genomische DNA bevond. In plaats daarvan vond er tijdens de inductie van de gistcellen een herschikking van DNA fragmenten plaats in pArT die er toe leidde dat het merker-gen, op de oorspronkelijke positie, in meer of mindere mate actief werd. Wellicht vormde de associatie van VirD2 met pArT een fysieke barrière tijdens de plasmide replicatie waardoor in sommige gevallen een verandering in de plasmide-structuur optrad. Er zijn geen aanwijzingen gevonden voor T-DNA mobiliteit. Wanneer aangenomen mag worden dat de verschillende virulentie eiwitten op de juiste manier in gist tot expressie komen zou het achterwege blijven van aantoonbare T-streng productie kunnen worden verklaard doordat integratie van DNA in het gistgenoom via het verbinden van niet homologe uiteinden (NHEJ) slechts met zeer lage efficiëntie plaatsvindt. Een andere mogelijke oorzaak kan zijn dat de torsie van de superhelix-structuur van plasmide-DNA in gist enigszins anders is dan bij bacteriën [15]. Uit een eerdere studie is gebleken dat de torsie van het DNA dat het T-DNA bevat, beperkend kan zijn voor de relaxase activiteit van VirD2 [6].

In **hoofdstuk 3** is er gekeken naar verschillende variabelen die invloed hebben op de vorming van circulair T-DNA in de gastheercel na AMT. De vorming van dergelijke T-DNA-cirkels is inmiddels aangetoond in zowel gist [10] als planten [16]. Door het uitvoeren van cocultivatatie experimenten met binaire vectoren die zijn voorzien van een *ARS/CEN* sequentie voor stabiele replicatie in gist kon worden aangetoond dat Rad52 een belangrijke rol speelt in de vorming van T-DNA-cirkels, terwijl Ku70 hieraan geen aantoonbare bijdrage leverde. Rad52 speelt een sleutelrol in DNA-herstel door homologe recombinatie (HR), Ku70 is noodzakelijk voor DNA-herstel middels NHEJ. Omdat de gebruikte binaire vectoren slechts een verwaarloosbare zelfhomologie hebben, kan de afhankelijkheid van Rad52 voor de vorming van T-DNA-cirkels alleen verklaard worden door de vorming van T-DNA-multimeren die vervolgens door HR worden opgelost in plasmiden. Een nieuwe serie binaire vectoren is geconstrueerd waarbij gevarieerd is met het aantal en de soort toegepaste bordersequenties. Het toepassen van deze pOphis serie, naar het Oudgriekse *ὄφης* (“slang”) maakte het mogelijk na AMT onderscheid te maken tussen volledige vectoroverdracht, zoals dat voor zou moeten komen bij gebruik van een enkele border, en partiële overdracht, wanneer twee borders worden gebruikt. In deze experimenten werd gevonden dat zowel in het geval van volledige of partiële vectoroverdracht Rad52 noodzakelijk was voor de vorming van T-DNA cirkels. Hiermee kon een “rolling circle replication like” mechanisme voor de vorming van T-DNA-multimeren worden uitgesloten. Uit het bepalen van de DNA sequenties van T-DNA-cirkels geïsoleerd uit getransformeerde gistcellen bleek dat er steeds perfecte borderfusies waren ontstaan, dus alsof VirD2 de fusie tussen de verschillende vector-uiteinden had gekatalyseerd. Deze resultaten leidden tot een model waarbij losse T-strengen, bestaande uit een partiële of uit de complete pOphis sequentie, gekoppeld kunnen worden door VirD2. Na de synthese van een complementaire DNA-streng kunnen de ontstane T-DNA-multimeren worden opgelost tot T-DNA-cirkels.

In **hoofdstuk 4** is onderzocht of het door *Agrobacterium* laten overdragen van de plaatsspecifieke endonuclease I-SceI, gezamenlijk met een T-streng zou kunnen leiden tot een verhoogde efficiëntie in T-DNA integratie via HR. Ook werd de invloed van de nucleosoombezetting van het chromatine op T-DNA integratie via HR onderzocht. Een nucleosoom is een complex dat wordt gevormd uit dubbelstrengs DNA en specifieke eiwitten. De positionering van nucleosomen speelt een bepalende rol in de expressie van genen. Efficiënte T-DNA integratie via HR zou voor gebruik in planten een zeer welkome technologie zijn die het mogelijk maakt een genomisch DNA fragment te vervangen door een T-DNA naar keuze. Vooralsnog is deze technologie nog niet breed toepasbaar doordat T-DNA integratie via HR in planten zeer infrequent plaatsvindt [17, 18]. Eerder uitgevoerd onderzoek toonde aan dat het induceren van een breuk in beide DNA strengen (DSB) van het genomische DNA op een specifieke plaats via I-SceI de integratie van T-DNA op die plaats bevorderde [19]. Een genconstruct met sequentie-homologie ten opzichte van een inkomende T-streng is daarom voor dit onderzoek geconstrueerd en in een gistchromosoom geplaatst.

Doordat dit doel-locus werd geflankeerd door twee I-SceI herkenningsequenties zouden hier via geïntroduceerd I-SceI specifieke DSBs kunnen worden aangebracht tijdens de cocultivatatie. Om de nucleosoombezetting te kunnen manipuleren tijdens de cocultivatatie zijn sequenties toegevoegd aan het doel-locus waarvan bekend is dat daarop nucleosomen worden gevormd wanneer glucose aanwezig is in het groeimedium. Deze nucleosoomstructuur wordt verwijderd zodra glucose wordt vervangen door galactose in het groeimedium. Uit de uitgevoerde experimenten bleek dat de efficiëntie waarmee T-DNA integratie plaatsvindt via HR niet werd beïnvloed door de aan- of afwezigheid van nucleosomen. De efficiëntie van T-DNA integratie via HR werd wel verhoogd door *Agrobacterium* gemedieerde overdracht van I-SceI tijdens de cocultivatatie. Wanneer echter tijdens dit experiment nucleosoomvorming werd geïnduceerd verdween het positieve effect van I-SceI op de transformatie efficiëntie. Een plausibele verklaring voor deze waarneming is dat de vorming van nucleosomen de inductie van DNA breuken verhindert door de bereikbaarheid van het DNA voor I-SceI te verlagen. Het is dus zeer wel mogelijk om T-DNA integratie via HR efficiënter te laten verlopen door gebruik te maken van *Agrobacterium* gemedieerde overdracht van specifieke nucleasen tijdens de cocultivatatie. Het succes hiervan is echter wel afhankelijk van de nucleosoombezetting van het voorgenomen integratie locus.

**Hoofdstuk 5** is behalve een samenvatting van het totale proefschrift ook een beschouwend hoofdstuk waarin een aantal van de bevindingen die zijn behandeld in het proefschrift in een bredere context worden bekeken. Daarnaast wordt een visie op toekomstige toepassingen van *Agrobacterium* gegeven waarbij wordt benadrukt dat een groot deel van het potentieel van AMT om maatschappelijk verantwoorde genetisch gemodificeerde gewassen te maken nog onbenut is.

## Referenties

1. Van Larebeke, N., et al., Large plasmid in *Agrobacterium tumefaciens* essential for crown gall-inducing ability. *Nature*, 1974. 252(5479): p. 169-70.
2. Watson, B., et al., Plasmid required for virulence of *Agrobacterium tumefaciens*. *J Bacteriol*, 1975. 123(1): p. 255-64.
3. Yadav, N.S., et al., Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc Natl Acad Sci U S A*, 1982. 79(20): p. 6322-6326.
4. Stachel, S.E. and P.C. Zambryski, virA and virG control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*. *Cell*, 1986. 46(3): p. 325-33.
5. Ankenbauer, R.G. and E.W. Nester, Sugar-mediated induction of *Agrobacterium tumefaciens* virulence genes: structural specificity and activities of monosaccharides. *J Bacteriol*, 1990. 172(11): p. 6442-6.
6. Scheiffele, P., W. Pansegrau, and E. Lanka, 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*, 1995. 270(3): p. 1269-76.
7. Montoya, A.L., et al., Octopine and nopaline metabolism in *Agrobacterium tumefaciens* and crown gall tumor cells: role of plasmid genes. *J Bacteriol*, 1977. 129(1): p. 101-7.
8. Caplan, A., et al., Introduction of genetic material into plant cells. *Science*, 1983. 222(4625): p. 815-21.
9. Hoekema, Binary plant vector strategy based on separation of vir- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature*, 1983. 303: p. 179-180.
10. Bundock, P., et al., Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *Embo J*, 1995. 14(13): p. 3206-14.
11. Piers, K.L., et al., *Agrobacterium tumefaciens*-mediated transformation of yeast. *Proc Natl Acad Sci U S A*, 1996. 93(4): p. 1613-8.
12. de Groot, M.J., et al., *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. *Nat Biotechnol*, 1998. 16(9): p. 839-42.
13. van Attikum, H., P. Bundock, and P.J. Hooykaas, Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *Embo J*, 2001. 20(22): p. 6550-8.
14. van Attikum, H. and P.J. Hooykaas, Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic Acids Res*, 2003. 31(3): p. 826-32.
15. Mayan-Santos, M.D., et al., DNA is more negatively supercoiled in bacterial plasmids than in minichromosomes isolated from budding yeast. *Electrophoresis*, 2007. 28(21): p. 3845-53.
16. Singer, K., et al., Formation of Complex Extrachromosomal T-DNA Structures in *Agrobacterium tumefaciens*-Infected Plants. *Plant Physiol*. 160(1): p. 511-22.

17. Offringa, R., et al., Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *Embo J*, 1990. 9(10): p. 3077-84.
18. Puchta, H., Gene replacement by homologous recombination in plants. *Plant Mol Biol*, 2002. 48(1-2): p. 173-82.
19. Puchta, H., B. Dujon, and B. Hohn, Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. *Nucleic Acids Res*, 1993. 21(22): p. 5034-40.



## ***Curriculum Vitae***

Arie Martijn Rolloos was born on the 8<sup>th</sup> of April, 1978 in Papendrecht, The Netherlands. In September 1998 he started the study “Laboratory Techniques” at the AHS Larenstein, Velp. Resulting from this, he received the bachelor degree in microbiology/biochemistry in September 2002. Subsequently, he studied Molecular Sciences at Wageningen University and Research Centre on the subject of biological chemistry and obtained the Master degree in November, 2005. Starting from June, 2006 until November, 2007 he did research at the *Helicobacter* group of the Laboratory of Gastroenterology and Hepatology of the Erasmus Medical Centre in Rotterdam, The Netherlands under supervision of Dr. Arnoud H. M. van Vliet and Prof. Dr. Ernst J. Kuipers. In December, 2007 he started as PhD student at the department of Molecular Developmental Genetics of the Institute of Biology Leiden under supervision of Dr. Ir. Bert J. van der Zaal and Prof. Dr. Paul J. J. Hooykaas. This research was focused on the *Agrobacterium* mediated translocation of proteins and transfer DNA (T-DNA) to budding yeast. During the last two years he did research in the same group on a project for the Consortium for Improving Plant Yield which received funding from the “Nederlands Genomics Initiative”. For this project, *Arabidopsis thaliana* plants transformed with a library of zinc fingers fused to a transcriptional repressor domain were screened for enhanced growth traits.

