The handle [http://hdl.handle.net/1887/22213](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
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 tumorgenesis 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
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 µm distance.](image)
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 µm 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 µl of a yeast cell suspension with an OD\text{\textsubscript{600}} of 10\textsuperscript{-3} was plated on YPD medium with G418 concentrations ranging from 2 µg/ml to 150 µg/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 µg/ml. The yeast cells harboring pArT1prom still divided and formed readily developing colonies even when challenged with a G418 concentration of 150 µg/ml. For all the experiments described below, G418 was applied at a concentration of 200 µg/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 µl of a yeast cell suspension with an OD\text{\textsubscript{600 nm}} adjusted to 10\textsuperscript{-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 acetylsyringone (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 acetylsyringone induction indicated here as AS+ and AS-. 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.
In the presence of glucose, this band was not observed. The apparent molecular weight of VirD2 on the Western blot is slightly larger than its predicted molecular weight of 47.5 kDa. This observation is in accordance with earlier work of Durrenberger et al. These results proved that the VirD2 protein was properly expressed in the inducible manner that could be expected. Although we lacked means to directly verify galactose-inducible expression of unmodified VirD1, VirC1, VirC2, and VirE2, the results regarding expression of GFP fusion proteins mentioned above (Figure 2 and 3) can be taken as an indication that the used expression cassettes are functional. Lack of proof for galactose-inducible T-DNA processing and mobilization thus seemed to be caused by other factors than faulty Vir protein expression.

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 µl of a yeast cell suspension with an OD_{600 nm} adjusted to 10^{-3} 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 pdc6 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**

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 outnumbe 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 *NcoI*-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 *NcoI*-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 pArT1stop 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 there 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 Southern blot 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 were 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 were 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 (MATα, ura3-52, lys2-801, ade2-101, trp1-Δ1, his3- Δ 200, leu2- Δ 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.

<table>
<thead>
<tr>
<th>Name primers</th>
<th>Sequence 5’-3’ Target cleavage site</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>VirD1 BamHII F</td>
<td>ATTAGGATCCAAAAATGTCAAAAACACCCAGAGTC</td>
<td>pOPSL 2</td>
<td>virD1</td>
<td>2</td>
</tr>
<tr>
<td>VirD1 NheI R</td>
<td>ATTAGCCTAGCTAAAGCAGCCCTCTAATAGC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VirC2 Xbal F</td>
<td>TAAAGTCGACCTTCACCAATCCTCGATGG</td>
<td>pOPSL 4</td>
<td>virC2</td>
<td>4</td>
</tr>
<tr>
<td>VirC2 XbaI R</td>
<td>ATTATCTAGAAAAATGGAATTCGCAAGCCCG</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VirC1 Xbal F</td>
<td>ATATGTGATCAAGGCTCCAGAATTGTTGCAATGAG</td>
<td>pOPSL 6</td>
<td>virC1</td>
<td>6</td>
</tr>
<tr>
<td>VirC1 BglII R</td>
<td>TAAAGATCTAAAAATGGAACATTGGACCTCATATGC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VirD2 BglII F</td>
<td>TAAATGATCTAAAAATGGAACATTGGACCTCATATGC</td>
<td>pOPSL 8</td>
<td>virD2</td>
<td>8</td>
</tr>
<tr>
<td>VirD2 XhoI R</td>
<td>TAAATGATCTAAAAATGGAACATTGGACCTCATATGC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VirE2 BamHII F</td>
<td>ATTAGGATCCAAAAATGGAATTCGCAATGAG</td>
<td>pOPSL 9</td>
<td>virE2</td>
<td>9</td>
</tr>
<tr>
<td>VirE2 NheI R</td>
<td>TCACTGCTAGCAAGAAGAGCTCAAAAGC</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VIPI SpeI F</td>
<td>ATTAAGAAGAACGAGGAGGAGAAGAGGA</td>
<td>pOPSL 11</td>
<td>VIPI</td>
<td>11</td>
</tr>
<tr>
<td>VIPI XhoI R</td>
<td>TAAATGATCAAAAGGAGGAAGAGGAGAAGAGGA</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Name primers</th>
<th>Sequence 5’-3’ Target cleavage site</th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>TTGAL7 Xbal F</td>
<td>CAGCTCTTAGAAGATATAAGAGTATGTGCG</td>
<td>pOPSL 1</td>
<td>TTgal7</td>
<td>1</td>
</tr>
<tr>
<td>TTGAL7 NheI R</td>
<td>ATGCGCTAGCGAAGATATAAGAGTATGTGCG</td>
<td></td>
<td>Not excised</td>
<td></td>
</tr>
<tr>
<td>UASGAL7 TTGAL10 BamHII F</td>
<td>GAGTAGGATCAATACCTACGTTTTTTTTTGATG</td>
<td>pOPSL 3</td>
<td>UASgal7, TTgal10</td>
<td>3</td>
</tr>
<tr>
<td>UASGAL7 TTGAL10 Xhol R</td>
<td>ATCGCTCGAGTTTGCAGCTTACTATCCTTC</td>
<td></td>
<td>BamHII, Xhol</td>
<td></td>
</tr>
<tr>
<td>UASGAL10-GAL1 SpeI F</td>
<td>TGACACTAGTTGGATATTTTTTTTTTTTTTTTTTTG</td>
<td>pOPSL 5</td>
<td>UASgal10,1</td>
<td>5</td>
</tr>
<tr>
<td>UASGAL10-GAL1 BamHII R</td>
<td>GCGATGATCTGGACCTGATTAAAGTAGAGGT</td>
<td></td>
<td>SpeI, BamHII</td>
<td></td>
</tr>
<tr>
<td>TTGAL1 XhoI R</td>
<td>GATGGTGCAGCTTTTTCACCTTTTACCTGCAAGAC</td>
<td>pOPSL 7</td>
<td>TTgal1</td>
<td>7</td>
</tr>
<tr>
<td>TTGAL1 SpeI F</td>
<td>TAAATGAGCGATAGCTGATAAGAC</td>
<td></td>
<td>Not excised</td>
<td></td>
</tr>
</tbody>
</table>
Cloning strategy

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 BamHI and XhoI and fragment 2 was inserted into this vector. The resulting plasmid was digested with NheI and BamHI. Fragment 3 was inserted into this vector. Independent of these cloning steps, pOPSL1 was digested with BglII and SalI and similarly digested fragment 6 was inserted into this vector. The resulting plasmid was digested with BglII and SpeI 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 Scal and XhoI. 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” which was renamed as Agrobacterium virulence cassette I (Figure 1).

To generate Agrobacterium virulence gene expression cassette II (Figure 1) pOPSL1 was digested with BamHI and XhoI and fragment 2 was inserted into this vector. The resulting plasmid was digested with NheI and BamHI and similarly digested fragment 9 was inserted. Independent of these cloning steps, pOPSL7 was digested with BglII and SalI and similarly digested fragment 8 was inserted. The resulting plasmid was digested with BglII and SpeI, 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 Scal and XhoI 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 resulted in plasmid “pSKN-SgraI with TTgal7:virE2:UASgal7:TTgal10:UAS10,1:virD2:TTgal1” which was renamed as Agrobacterium tumefaciens virulence cassette II (Figure 1).
“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 were 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 APTI marker gene was excised by digesting pINT with HindIII. The resulting 1401 bp fragment that carried a part of the pdc6 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-GFP6H 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’ GTCAGGTAGTAAAGGAGAAGAACTTTTCACTGG 3’) and VirD1-GFP RV (5’ GTGCCTTTGTATAGTTCATCCATGGCCTG 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’ ATTACGGACCGTGGTGAAGAGGAAGAAGAAAATTTTCACTGG 3’) and VirD2-GFP RV (5’ ATTACGGACCGATTTGTATAGTTCATCCATGGCCTG 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).

\[
\begin{align*}
\text{LB-FW:} & \quad 5' - \text{GCTAGCGTCGcgcagctattcatatcttgaattGGCTTCATGTCCCGGAAATCTACATG} \\
& \quad \text{GATCAGCAATGAGTATGCAGCT} -3' , \\
\text{LB-RV:} & \quad 5'-\text{ATCCATGTAGATTTCCCGGACATGAAGCCatttacaattgaatatcttgccgCGACGCTAGCGAGCT} -3' , \\
\text{RB-FW:} & \quad 5'-\text{TCGACTGACgtggcaggatatatcttgtaattTGAGCTCGTGTGAATAAGTCGCTGTGT} \\
& \quad \text{ATGTGGTTGTGTGTGCTAGC} -3' , \\
\text{RB-RV:} & \quad 5'-\text{TCGAGCTAGCACAATCAAACAAACATACACACAGCGACTTATTCACACAG} \\
& \quad \text{AGCTCAaattacaagcttatatcttgccagTCAG} -3'.
\end{align*}
\]

The KanMX marker gene used for vector construction originates from the \textit{E. coli} transposon Tn903. To obtain this gene, the binary vector pSDM8000 [22] was digested with \textit{NcoI} and \textit{SalI}. This DNA fragment containing the \textit{KanMX} gene with \textit{TEF1} terminator was cloned into pSKN-SgraI. A \textit{SalI} site was removed by blunting \textit{SalI} 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’-TCGACTAGTCTAGACTAGTGTGATCATTAAGTGATCATTAAGC-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 ºC 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 H₂O. The yeast cells were transferred to 50 ml liquid SD medium containing 2% glucose or 2% galactose with the OD₆₀₀ 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 ºC, 200 rpm. The cell suspensions were centrifuged for two minutes at 2500 rpm and the pellets were washed with 50 ml sterile H₂O. 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 30ºC, 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 ºC, 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 OD₆₀₀ 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 ºC. 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 NcoI 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’ GCCGTTTCTGTAAATGAGGAG 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).

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


