Involvement of host and bacterial factors in Agrobacterium-mediated transformation
Shao, S.

Citation

Shao, S. (2019, November 28). *Involvement of host and bacterial factors in Agrobacterium-mediated transformation*. Retrieved from https://hdl.handle.net/1887/80841

Cover Page

Universiteit Leiden

The following handle holds various files of this Leiden University dissertation:
http://hdl.handle.net/1887/80841

Author: Shao, S.
Title: Involvement of host and bacterial factors in Agrobacterium-mediated transformation
Issue Date: 2019-11-28

Involvement of host and bacterial factors in *Agrobacterium*-mediated transformation

Shuai Shao

邵帅

Involvement of host and bacterial factors in *Agrobacterium*-mediated transformation

Proefschrift

Ter verkrijging van

de graad van Doctor aan de Universiteit Leiden,

op gezag van Rector Magnificus prof. mr. C.J.J.M. Stolker,

volgens besluit van het College voor Promoties

te verdedigen op donderdag 28 November 2019

klokke 11:15 uur

door

Shuai Shao

Geboren te Jiangxi, China

17 Augustus 1989

To Lingjuan and my family

Table of contents

Chapter 1
General Introduction

Shuai Shao

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Agrobacterium tumefaciens, a gram-negative plant pathogen belonging to the family *Rhizobiaceae*, is the causative agent of crown gall disease, which can affect many plant species including agronomically important ones. It induces tumor formation in plants by transferring a segment of its tumor-inducing plasmid (Ti-plasmid) to plant cells. This transferred DNA (T-DNA) contains genes encoding enzymes involved in the synthesis of auxin (*iaaM, iaaH*), cytokinin (*ipt*), opines resulting in uncontrolled proliferation of cells producing opines and *plast* genes (such as *b, c', d, e, 5, 6a, 6b*) for phenotypic plasticity. Under laboratory conditions it is also able to transform other eukaryotes such as yeast *Saccharomyces cerevisiae* and other fungi (Bundock, et al., 1995; de Groot et al., 1998; Lacroix, et al., 2006). Hence, it was developed and is now extensively used as a vector to create transgenic plants and fungi. In fact *Agrobacterium*-mediated transformation (AMT) has become the preferred method of transformation of these organisms over the past decades.

Agrobacterium genus

Empirically, the genus *Agrobacterium* was divided into several species based on the disease phenotype and their host range (reviewed by Gelvin, 2003). *A. tumefaciens* provokes crown gall tumors on dicotyledonous plant species; *Agrobacterium rhizogenes* induces hairy root disease, characterized by root proliferation from infected sites; *Agrobacterium rubi* causes crown gall disease on raspberries; *Agrobacterium vitis* causes gall formation but only restricted to grapevines; *Agrobacterium radiobacter* is not pathogenic. Since the tumorigenic properties of *Agrobacterium* are largely determined by the presence of a transmissible tumor-inducing (Ti) or root-inducing (Ri) plasmid, the above classification does not reflect relatedness of the bacteria. Therefore, *Agrobacterium* organisms were alternatively grouped into three biovars based on biochemical and physiological properties (Keane et al., 1970). Biovar I embraced most of the best characterized strains such as A6, Ach5, B6, and C58 and included both tumorigenic and avirulent bacteria. Biovar II contained the bacteria inducing hairy root disease but also avirulent bacteria such as the biological control agent K84. Biovar III comprised mostly strains with a narrow host range for *Vitis vinifera*. An initial proposal to include all these biovars into the genus *Rhizobium* on the basis of limited 16S rRNA divergence led to controversy (Farrand et al., 2003; Young et al., 2003). Extensive taxonomic analyses of *Rhizobiaceae* has now led to the proposal to classify biovar I as the genus *Agrobacterium*, biovar II as the species *Rhizobium rhizogenes* within the genus *Rhizobium* and biovar III as the species *Allorhizobium vitis* (Lindström and Young, 2011; Mousavi et al., 2015). Bacteria of the genus *Agrobacterium* (biovar I) are distinguished from other bacteria of the family by possessing a lin chromosome besides a circ chromosome (Ramirez-Bahena et al., 2014). Nowadays, conventional sequencing is common and easy to perform with low price and this contributes to the classification of *Agrobacterium* species. Phylogenetic analysis using sequences of the *recA* gene led to the proposal to distinguish 13 species (genomovars) within the *Agrobacterium* genus representing specific ecological adaptation (Costechareyre et al., 2010). However, many of these species have not yet obtained an official name and therefore in this thesis we shall use the generic name *Agrobacterium tumefaciens* to refer to the bacteria of this genus, which we have used in the experiments described in this thesis.

The Ti-plasmid or Ri-plasmid present in *Agrobacterium* cannot be used for taxonomic purposes, because they are transmissible and may be present in different *Agrobacterium*

species and even in certain of the other species of the *Rhizobiaceae* family. While strains with a Ti-plasmid induce the formation of tumor/crown gall, Ri-plasmid containing bacteria induce hairy roots. This is due to differences in their T-DNA genes, Ri T-DNAs containing a set of *rolABCD* genes instead of the *iaa* and *ipt* genes. Still, some Ri plasmids such as agropine pRiA4 have a second TR-region which contains *iaaM* and *iaaH* (Camilleri and Jouanin, 1991). Otherwise Ti- and Ri-plasmid share similar virulence (*vir*)–genes, conjugative transfer (*tra*, *trb*)-genes, and replication (*repABC*) genes. When strains are cured of their Ti/Ri-plasmid, the plasmid-less strain will become nonpathogenic. Upon re-introduction of the plasmid, virulence is restored. When a Ti-plasmid is introduced into a strain already carrying a Ri-plasmid, the resulting strain with both a Ti- and a Ri-plasmid induces the symptoms of both crown gall and hairy root disease (Costantino et al., 1980).

Ti- and Ri-plasmids not only induce tumor and hairy root formation but also the synthesis of various amino acid and sugar phosphate derivatives in transformed plant cells named opines (Petit et al., 1970; Dessaux et al., 1993). Conversely, Ti-and Ri-plasmids carry opine catabolism gene clusters which enable the host strains to take up and utilize opines as nutrient source. According to the different opines found in the tumor, *A. tumefaciens* strains and their Ti-plasmids were originally classified into three groups: nopaline-type, octopine-type and null-type. Later new crown gall opines were found, including agrocinopines, agropine, chrysopine, leucinopine, mannopine, succinamopine and vitopine (Chilton et al., 1984; Chilton et al., 1985a; 1985b; Chilton et al., 1995; Chilton et al., 2001). In hairy roots agrocinopines, agropine and mannopine were found, but also the novel opines cucumopine (Davioud et al., 1988; Szegedi et al., 1988). More information can be found at https://en.wikipedia.org/wiki/Opine. The classification scheme of Ti-and Ri-plasmids based on opine types is still widely used. Ti/Ri-plasmids have matching genes for the uptake and catabolism of the opines they induce in the tumors/hairy roots.

Several widely spread opine biosynthesis genes are listed in Table 1, such as nopaline synthase (*nos*), octopine synthase (*ocs*), agropine synthase (*ags*), mannopine synthases (*mas1, mas2*), agrocinopine synthase (*acs*), succinamopine/leucinopine synthase (*sus, les*), chrysopine synthase (*chs*). Opine utilization genes include genes encoding proteins involved in transport of a specific opine into the cell besides catabolic enzymes (Table 1). The biosynthesis and catabolism of opines involve similar reactions performed in opposite directions. However, the genes involved are not homologous in most cases. In relation to agropine and mannopine, however, they were pairwise homologous. The *ags* gene for agropine synthase is homologous to the catabolic *agcA* gene, while the mannopine synthesis genes *mas1* and *mas2* are homologous to the catabolic genes *mocC* and *mocD*.

The wide use of next-generation sequencing techniques has generated a revolution in genomics by obtaining the whole genome sequence at a lower cost and faster than in the past. With the now common practice of sequencing whole bacterial genomes, large genomic data sets are easily acquired and released in the public databases. The average nucleotide identity (ANI) calculation is therefore becoming more and more popular to elucidate bacterial relatedness (Konstantinidis and Tiedje, 2005; Ormeno-Orrillo et al., 2015). To date, more than 50 *Agrobacterium* strains have been sequenced and their genome assembly levels range from contigs to complete genomes. All complete sequences of *Agrobacterium* strains released by April 2019 are listed in Table 2. Sequences of *Agrobacterium* strains released after that date

are available in the NCBI database. As demonstrated in the list, all *Agrobacterium* strains harbor one circ chromosome, one lin chromosome, at least one extrachromosomal plasmid and the virulent strains always carry a Ti-plasmid or Ri-plasmid. This chromosomal organization is distinct from those found in biovar II and III. For instance, *Rhizobium* strains including *R.rhizogenes* contain two circular chromosomes.

Table 1. The common known opines and related genes.

Opines	Biosynthesis	Regulator	Transporter	Degradation
Nopaline	*nos*	*nocR*	*nocP, T, Q, M*	*noxA,B,C* (nopaline oxidase); *arc* (arginase)
Octopine	*ocs*	*occR*	*occP, M, Q, T*	*ooxA, B* (octopine oxidase); *ocd* (ornithine cyclodeaminase)
Agropine	*ags*	*moaR*	*agtA, B, C, D* (agropine); *agaD, B, C, A* (agropinic acid)	*agcA* (agropine); *agaE* (agropinic acid)
Mannopine	*mas1/mas2*	*mocR*	moaA, B, C, D (mannopine and mannopinic acid)	mocC,D (mannopine); agaF,G (mannopinic acid)
Agrocinopine	*acs*	*accR*	*accA, B, C, D, E*	*accF* (agrocinopine phophodiesterase); *accG* (arabinose-phosphate phosphatase)
Succinamopine	*sus*	*sacR*	*sacA, B, C, D*	*sacE, F, G, H*
Chrysopine	*chsA, B, C*	*chcR*	*chtA, B, C* (chrysopine); *sclA, B, C, E* (dfg); *dfpA, B, C, D* (dfop)	*chcA, D, E* (chrysopine); *sclD* (dfg); *dfpF, G, H, I* (dfop)
Leucinopine	*les*	*lecR*	*lecA, B, C, D*	*lecE, F, G, H*

Table 2. The published complete genome sequences of *Agrobacterium* strains (as of April 2019), and those of the published *R.rhizogenes* and *A.vitis* strain.

Organism	Size (Mb)	GC%	Replicons/Accession number	Replicons	Genes
H13-3	5.57	58.5	chrom cir:CP002248; chrom lin:CP002249; pAspH13-3a:CP002250	3	5365
RAC06	4.96	61.1	chrom cir:CP016499; pBSY240_1:CP016500	2	4735

C58	5.67	59.1	chrom cir:AE007869; chrom lin:AE007870; pAtC58:AE007872; pTiC58:AE007871	4	5459
1D132	5.55	59.0	chrom cir:CP033022; chrom lin:CP033023; pAt1D132a:CP033024; pAt1D132b:CP033025; pTi1D132:CP033026	5	5415
Ach5	5.67	58.5	chrom cir:CP011246; chrom lin:CP011247; pAtAch5:CP011248; pTiAch5:CP011249	4	5404
15955	5.87	58.5	chrom cir:CP032917; chrom lin:CP032918; pAt15955:CP032919; plasmid pTi15955:CP032920	4	5626
S33	5.48	59.2	chrom cir:CP014259; chrom lin:CP014260	2	5278
12D13	5.42	59.2	chrom cir:CP033034; chrom lin:CP033035; pAt12D13a:CP033036; pAt12D13b:CP033037; pAt12D13c:CP033038	5	5330
1D1609	5.99	59.5	chrom cir:CP026924; chrom lin:CP026925; pAt1D1609a:CP026927; pAt1D1609b:CP026928; pTi1D1609:CP026926	5	5862
1D1108	5.77	58.5	chrom cir:CP032921; chrom lin:CP032922; pAt1D1108a:CP032923; pAt1D1108b:CP032924; pTi1D1108:CP032925	5	5555
1D1460	5.68	59.3	chrom cir:CP032926; chrom lin:CP032927; pAt1D1460:CP032928; pTi1D1460:CP032929	4	5629
12D1	5.45	59.5	chrom cir:CP033031; chrom lin:CP033032; pTi12D1:CP033033	3	5241
A6	5.94	58.4	chrom cir:CP033027; chrom lin:CP033028; pAtA6:CP033029; pTiA6:CP033030	4	5695
K599/NCPPB2659	5.48	59.7	chrom cir:CP019701; chrom lin:CP019702; pRi2659:CP019703	3	5268
Rhizobium rhizogenes K84	7.27	59.9	chrom cir:CP000628; chrom cir:CP000629; pAgK84:CP000632; pAtK84b:CP000630; pAtK84c:CP000631	5	6941
Allorhizobium vitis S4	6.32	57.5	chrom cir:CP000633; chrom cir:CP000634; pTiS4:CP000637; pAtS4a:CP000639; pAtS4b:CP000635; pAtS4c:CP000636; pAtS4e:CP000638;	7	5820

Likewise, the complete sequences of Ti-plasmids have been acquired by whole genome sequencing. Compared to the difficulties in genome assembly of the chromosomes, especially the lin chromosome which contains tandem repeats at both ends, the complete sequence of the much simpler and smaller Ti-plasmids can relatively easily be acquired. In Table 3, all known complete sequences of Ti-plasmids are presented including the newly sequenced Ti-

plasmids in this thesis. Several Ti-plasmids (nopaline-type pTiSAKURA, octopine-type pTi15955 and agropine-type pTiBO542) (Suzuki et al., 2000; Oger et al., 2001; Zhu et al., 2000) were previously sequenced by genome walking and construction of cosmid libraries. From three *Agrobacterium* strains (nopaline-type C58, octopine-type LBA4213 and Ach5) (Goodner et al., 2001; Wood et al., 2001; Henkel et al., 2014; Huang et al., 2015) Ti-plasmid sequences were obtained after sequencing of the whole genome by next-generation sequencing.

Table 3. All known complete sequences of Ti-plasmids (as of April 2019).

Strain	Plasmid	Accession number	Size (bp)	GC%	Genes
C58	pTiC58	AE007871	214,233	56.7	199
SAKURA	pTi-SAKURA	AB016260	206,479	56.0	196
Ach5	pTiAch5	CP011249	194,264	54.7	182
Bo542	pTiBo542	DQ058764	244,978	55.1	223
186	pTi186	CM008377	177,704	56.1	178
1D132	pTi1D132	CP033026	177,577	56.1	179
1D1609	pTi1D1609	CP026926	166,117	54.9	164
1D1108	pTi1D1108	CP032925	176,213	56.1	177
1D1460	pTi1D1460	CP032929	214,233	56.7	207
15955	pTi15955	CP032920	194,263	54.7	184
12D1	pTi12D1[*]	CP033033	160,006	59.1	155
A6	pTiA6	CP033030	194,263	54.7	184
183	pTi183	CP029048	192,674	56.1	194
LBA4213	pTiLBA4213	CP007228	205,997	55.0	196
Chry5	pTiChry5[#]	KX388536	197,268	54.5	219
EU6	pTiEU6[#]	KX388535	176,375	56.1	195
T37	pTiT37[#]	MK439386	203,781	55.9	198
Kerr27	pTiKerr27[#]	MK439385	243,905	57.2	233
Kerr108	pTiKerr108[#]	MK439384	220,307	56.8	207
Sule1	pTiSule1[#]	MK439381	217,820	56.8	202
CFBP1935	pTiCFBP1935[#]	MK439383	213,092	56.8	196
CFBP2178	pTiCFBP2178[#]	MK439382	217,821	56.8	202
C5.7/C6.5[&]	pTiC5.7/C6.5	MF511177/MK318986	218,413	56.8	213
1724	pRi1724	AP002086	217,594	57.3	173
2659	pRi2659Δ[§]	EU186381	185,462	58.1	146
2659	pRi2659	CP019703	202,302	57.3	181
S4	pTiS4	CP000637	258,824	56.7	249

Note: [*], no virulence region; [#], all newly sequenced Ti-plasmids in this thesis. [&], Both *R.rhizogenes* C5.7 and C6.5 were isolated from the same crown gall tumor and harbor almost identical Ti-plasmids (Kuzmanović and Puławska, 2019). [§], pRi2659Δ is disarmed by deleting its T-DNA.

Ti-plasmids

Ti-plasmids have been comprehensively studied because they are the molecular basis of *Agrobacterium*-mediated genetic transformation (reviewed by Gordon and Christie, 2014). The T-DNA region of the Ti-plasmid is surrounded by two direct repeat sequences of 25 bp termed T-borders. The genes on the T-DNA are mainly related to two functions. Some can lead to the over-production of an auxin (*iaaM, iaaH*) and a cytokinin (*ipt*) leading to the formation of crown gall tumors. Several other related to cell growth and tumor formation such as *6b* and gene *5*. Others are involved in the synthesis of opines as mentioned above. Close to the right T-border outside the T-DNA region genes are located which are responsible for opine uptake and catabolism, enabling *A. tumefaciens* to utilize opines as nitrogen, carbon and energy sources. The traditional classification of Ti-plasmids is defined by the T-DNA region and corresponding opine catabolism region. Besides, all Ti-plasmids share four main gene clusters (Figure 1). The *repABC* operon, which is responsible for the control of plasmid replication, partition and maintenance. The *tra* operon, which is associated with conjugative DNA processing of the Ti-plasmid and the *trb* operon encoding the Type IV secretion system (T4SS) necessary for mate pair formation. The virulence region carries a set of *vir* genes expressing a range of virulence proteins that are necessary for the transfer of T-DNA into host cells. Common to all Ti-plasmids, the *tra* operon is located near the region encoding catabolism of the conjugative opines or other unknown compounds and separated by a large distance from the *trb/rep* region (Wetzel et al., 2015). Notwithstanding these common backbone elements, there still are significant differences in the genetic organization among Ti-plasmids (Figure 1). Moreover, while nopaline and succinamopine Ti-plasmids contain
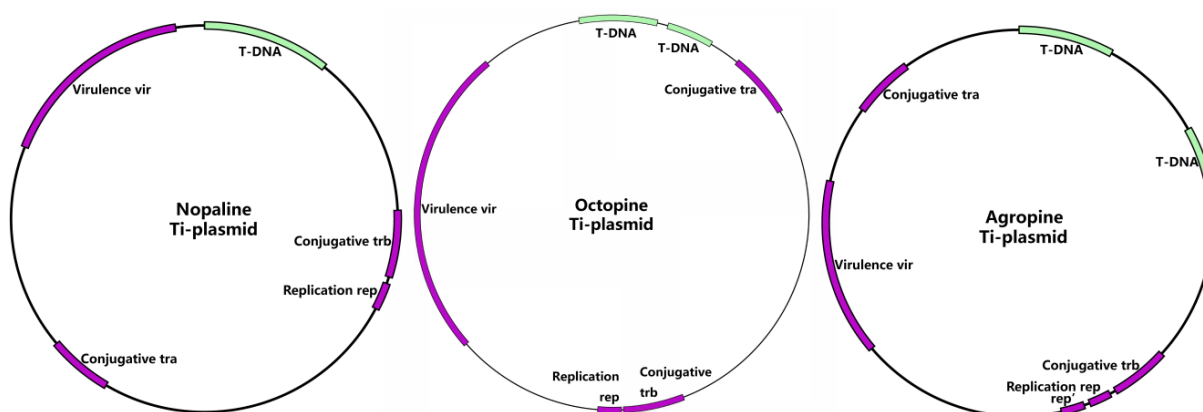
one T-DNA region, octopine-type, chrysopine-type and agropine-type Ti-plasmids carry two separate T-DNA regions which can be transferred individually into the host. Hence, the number of T-DNA regions can also be used as a marker to discriminate the Ti-plasmids.

Figure 1. Schematic structure of common Ti-plasmids. From left to right, schematic structure of nopaline Ti-plasmid pTiC58, of octopine Ti-plasmid pTiAch5, and of agropine Ti-plasmid pTiBo542. All Ti-plasmids embrace several conserved regions: transferred DNA (T-DNA region); conjugation region *tra* and *trb*; replication origin *rep*; and virulence region *vir*.

Virulence proteins involved in *Agrobacterium*-mediated transformation

Agrobacterium mediated transformation can be depicted as a process composed of multiple stages (reviewed by Gelvin, 2017). The early events that occur in *Agrobacterium,* have been extensively studied. After wounding, which is necessary for infection, several signals, such as phenolic compounds and neutral and acidic sugars, are released from the wounded plant tissue infected by *Agrobacterium.* These compounds can trigger a two-component sensory-response system (VirA-VirG) to stimulate the expression of the virulence genes on Ti-plasmids (Winans, 1992). Subsequently, the virulence protein VirD2, together with VirD1, nicks at 25-bp direct repeat sequences called left border (LB) and right border (RB) to generate a single-stranded copy of T-DNA called T-strand (Stachel et al., 1986). During this process, VirD2 covalently binds to the 5' end of the T-strand at the RB and eventually guides it into the nucleus of host cells (Ward et al., 1988). Simultaneously, a variety of other virulence proteins are expressed as well. The VirC1 and VirC2 proteins help VirD2 in nicking at the border repeats (Atmakuri et al., 2007). The VirB1-11/VirD4 proteins couple together to establish a dedicated TypeIV-Secretion System (T4SS) to translocate the T-strand along with some specific virulence effector proteins VirE2, VirE3, VirD5, and VirF to the host cells (Vergunst et al., 2000; 2005). The effector VirE2 plays an important role and can coat the T-strand to prevent exonuclease digestion in the host cell and may participate in targeting the T-strand into the nucleus (Citovsky et al., 1992). VirE3 is a transcription factor which induces the expression of host genes including VBF which has similar function as VirF (Schrammeijer et al., 2001; Garcia-Rodriguez et al., 2006; Niu et al., 2015); VirD5 can cause chromosome instability in the host and therefore probably provide more opportunities for T-DNA integration (Zhang et al., 2017; 2019); VirF is an F-box protein, which can contribute to the removal of VirE2 protein from T-strands in host cells (Schrammeijer et al., 1998; Lacroix and Citovsky, 2015).

T-DNA integration

Over time, several mechanisms have been proposed for T-DNA integration into the plant chromosome, but the precise mechanism of T-DNA integration has not yet been uncovered in all details. The principle reason is no T-DNA integration assay currently exists *in vitro*. With the rapid development of sequencing technology, the chromosomal T-DNA integration sites, especially in plants, have been analyzed and the results enhanced our understanding of T-DNA integration (Kim et al., 2007; Kleinboelting et al., 2015; Shilo et al., 2017). The sequences around the integration sites support the notion that T-DNA integrates into the plant genome randomly by non-homologous recombination. However, in yeast T-DNA integrates preferably by homologous recombination (HR) (Bundock et al., 1995; van Attikum et al., 2001). Even when homology is provided integration by HR in plants remains an extremely rare event (Offringa et al., 1990).

Taking a glance at the integration sites of T-DNA in plants, they are generally not "precise" and "clean". Often deletions, insertions, and rearrangements can be found such as deletions close to the target sites, LB-LB or RB-RB T-DNA insertions and filler DNA from unknown sources (reviewed in Gelvin, 2017). Similar structures were present in extrachromosomal T-circles, which were recovered prior to integration using a plasmid-rescue approach (Singer et al., 2012). The sequences of end-joining junctions included T-

DNA border fusions, T-DNA truncations, binary plasmid sequences, and filler DNA sequences derived from the T-strand or plant genome, indicating that the formation of extrachromosomal T-circles probably exploits the same DNA repair pathways as T-DNA integration. Such T-circles are probably not stably maintained, but may be responsible for early (transient) expression of the T-DNA. In contrast, T-DNA can be circularized and stably maintained as an extrachromosomal plasmid in yeast when the T-DNA contains a replication origin (Bundock et al., 1995; Soltani, 2009; Rolloos et al., 2014; Ohmine et al., 2016). Lacking the need of integration, such plasmid-like T-DNAs were transferred with relatively high efficiency. Strikingly, the HR key protein Rad52, rather than NHEJ crucial protein Ku70, was identified to be involved in the formation of T-circles (Rolloos et al., 2014; Ohmine et al., 2016). To date there is still no direct evidence to support the idea that T-circles can eventually integrate into host genome even though T-circle formation shares a similar mechanism as T-DNA integration. Moreover, we have to be cautious on the role of VirD2 as it may somehow be involved in the formation of T-circles because of its ability to reverse the nicking reaction (Pansegrau et al., 1993). Therefore, the mechanism of complex T-circle formation should be investigated in detail and this understanding may give more clues to uncover the mechanism of T-DNA integration.

The T-DNA itself does not encode any proteins required for integration. In summary, the fate of the T-DNA is determined to a large extent by host cells. In yeast, the integration process is preferentially mediated by HR. T-DNA integration by HR was absent in *rad52Δ* deletion mutants (van Attikum and Hooykaas, 2003; Rolloos et al., 2014; Ohmine et al., 2016), whereas in fact the integration efficiency through NHEJ was slightly increased in the absence of Rad52 (van Attikum et al., 2001). In *ku70/80Δ* and *lig4Δ* deletion mutants which genes are essential for DNA repair by NHEJ, non-homologous T-DNA integration was impossible in yeast (van Attikum et al., 2001). The Mre11-Rad50-Xrs2 complex was taken into consideration as well because of its role in end-joining and the efficiency of non-homologous T-DNA integration was decreased in the corresponding deletion mutants in yeast (van Attikum et al., 2001). This suggested that T-DNA integration occurred at genomic double strand breaks (DSBs) by a process of DSB-repair. However, in plants mutation of Ku70, Ku80 or Lig4 did not abolish T-DNA integration (reviewed by Gelvin, 2017). This indicated that other uncharacterized DNA repair pathways are involved in T-DNA integration in plants. Recently, DNA polymerase theta (Pol θ), which is evolutionary conserved in plants and animals, but not in fungi and which is involved in theta-mediated end-joining (TMEJ), was unraveled to mediate random T-DNA integration in plants (van Kregten et al., 2016). T-DNA integration was completely abolished after mutating Pol θ. Hence, summarizing, integration by HR only occurs at extremely low efficiency in plants and NHEJ might play a role, but is not essential in contrast to the essential role of TMEJ.

Several genome-wide screens have been performed to investigate the role of host factors involved in the transformation process in yeast (Soltani, 2009; Ohmine et al., 2016). By using the collection of mutants with deletions in non-essential genes of the yeast *S. cerevisiae*, more than 200 genes were identified of which deletions result in an at least 2-fold increased or reduced AMT efficiency (Soltani, 2009; Soltani et al., 2009). Interestingly, the deletions of genes (*EAF7*, *NGG1*, *YAF9* and *GCN5*) encoding subunits of ADA, SAGA and NuA4 transcriptional regulatory histone acetyltransferase (HAT) complexes highly increased

AMT efficiency, whereas the deletions of genes (*HDA2/3* and *HST4*) related to histone deacetylase (HDACs) complexes led to a strongly decreased efficiency. Besides, some other host factors in yeast were identified by similar screens and may also be important for AMT such as *SRS2* (encoding a DNA helicase), *SMI1* (a cell wall regulator) and *ERG28* (a membrane sterol scaffold protein) (Ohmine et al., 2016). Similar screens have been performed in the model plant *Arabidopsis thaliana* to identify mutants resistant to *Agrobacterium* transformation by screening a library of T-DNA insertion mutants (Zhu et al., 2003) and utilizing RNAi technology to silence or decrease expression of specific chromatin-related genes (Crane and Gelvin, 2007). In these screens, several genes were found to be important for T-DNA integration, including chromatin structure and remodeling genes (*HTA1*, *HDT1*, *HDT2*, and *SGA1*), nuclear-targeting genes encoding importin, cytoskeleton genes encoding actin, cell wall structural and metabolism genes. Chromatin components may mediate T-DNA targeting to the host genome or modify chromatin structure to allow access of T-DNA to integration sites.

The role of VirD2 in AMT

The VirD2 protein plays a crucial role in the *Agrobacterium*-mediated transformation of plants and yeast. It is involved in the formation of the T-strand, in the delivery of the T-strand into the host cell by the T4SS, in the targeting of the T-strand into the host cell's nucleus and may also have a role in the integration of the T-DNA, for instance by protecting the 5' end of T-strand and interaction with histones. The role of VirD2 in the first three processes is rather well characterized, whereas a putative role of VirD2 in the integration process is still unclear. The N-terminal region of VirD2 contains endonuclease activity to nick at the T-DNA borders to release the T-strand (Yanofsky et al., 1986 and Young et al., 1988). A tyrosine residue (Tyr 29) participates in this cleavage and the T-strand remains covalently attached to the residue. Sequences within the C-terminal part of VirD2 are responsible for the transfer of the VirD2/T-DNA complex through the T4SS (van Kregten et al., 2009). A nuclear localization signal sequence (NLS) present close to the C-terminus contributes to the delivery of the VirD2/T-strand complex into the host nucleus (Howard et al., 1992).

To date, a role of VirD2 in the integration of the T-DNA into the host chromosomal DNA remains obscure. Like other relaxases, VirD2 can reverse the nicking reaction and in this way a single-stranded T-circle may be formed, which may become double-stranded in a next step in host cells. Such T-circles have indeed be found in yeast (Bundock et al., 2005; Rolloos et al., 2014; Ohmine et al., 2018). From plants T-circles have been captured as well in early stage of infection (Singer et al., 2012), but they are not formed by precise circularization, but seem to be formed by a process of non-homologous end-joining. It might be possible that VirD2 could capture sequences with similarity to the border repeat in the host genome, but no evidence has been obtained for such reaction (Ziemienowicz et al., 2000). It was published that VirD2 mutated at Arg129 site leads to integrations with more truncations at the 5'-end, suggesting that VirD2 helps maintain the integrity of the 5'-end end during integration (Tinland et al., 1995).

In yeast Two-Hybrid screens the NLS binding protein importin-α (α-karyopherin) AtKAPα and cyclophilins CyPs from *A. thaliana* were identified to interact with VirD2 (Ballas and Citovsky, 1997; Deng et al., 1998). The conserved cyclin-dependent kinase-activating

kinase CAK2Ms and a type 2C serine/threonine protein phosphatase (PP2C) were also reported to interact with VirD2 (Bako et al., 2003; Tao et al., 2004) affecting nuclear import of the T-strand. The S-adenosyl-L-homocysteine hydrolase (involved in DNA methylation) and a MYST-like histone acetyltransferase 2 were found to interact with VirD2 by screening a cDNA library *in planta* (Lee et al., 2012). In the yeast *S. cerevisiae*, VirD2 was identified to bind to histone proteins and the interactions between VirD2 and the histones H2A, H2B, H3 and H4 were revealed (Wolterink-van Loo et al., 2015). VirD2 may thus play a role in targeting the T-strand to chromatin.

The function of Ada2 protein in yeast

The Spt-Ada-Gcn5 acetyltransferase (SAGA) complex is an evolutionary conserved, multifunctional co-activator complex and is organized into separate modules with distinct functions: the core structural (SPT) module, the histone acetyltransferase (HAT) module, the histone deubiquitinase (DUB) module, and the activator-binding (TAF) module (Lee et al., 2011; Helmlinger and Tora, 2017)(Figure 2). The independent HAT module composed of Gcn5, Ada2/3, and Sgf29 is connected to other parts of the SAGA complex with the help of Ada2. On the other hand, this module can also form the ADA complex with participation of two other proteins, Ahc1/2 (Eberharter et al., 1999; Lee et al., 2011). The HAT module within the SAGA and ADA complexes shares the core crucial proteins and can acetylate lysine residues of histone H3 to activate gene expression by opening the chromatin structure and by generating docking surfaces for other regulatory factors (Helminger and Tora, 2017)). In contrast to the HAT and SAGA complexes, the ADA complex can be targeted to gene promoters independent of Tra1 (Brown et al., 2001), which is thought to be the primary target for transcriptional regulation by recruiting the SAGA complex (Saleh et al., 1998; Berg et al., 2018). Both of these complexes are involved in the post-translational modifications of histones that are crucial for chromatin-dependent functions and the regulation of numerous cellular processes in response to environmental cues (Sterner et al., 2002). In addition, SAGA complexes act as transcription factors that can promote RNA polymerase II transcription (Baptista et al., 2017) and the expression of approximately 10% of all the yeast genes is dependent on SAGA for expression (Huisinga and Pugh, 2004). In human cells, the SAGA complex and especially the DUB module was found to be required for genome stability by promoting the use of the sister chromatid for DSB repair (Evangelista et al., 2018).

Ada2 is evolutionarily conserved among eukaryotes and has been described for several organisms, including *Arabidopsis* (Hark et al, 2009), *Drosophila* (Muratoglu et al., 2003) and human cells (Gamper et al., 2009). Remarkably, there are two paralogous Ada2a and Ada2b proteins in higher eukaryotes. Whereas Ada2a was identified to be included in Ada Two A Containing (ATAC) complex, Ada2b is specific to the HAT module like Ada2 in yeast (Kusch et al., 2003; Muratoglu et al., 2003). The function of Ada2 in this complex is conserved. Ada2 is thought to interact with Gcn5 directly to increase its HAT activity which preferentially acetylates histone H3 and histone H2B (Grant et al., 1997; Hoke et al., 2008). In *Arabidopsis*, Ada2b physically associate with Gcn5 and enhances its HAT activity to regulate gene expression under environmental stress conditions such as cold, drought and salt stress (Hark et al, 2009). In *Drosophila*, Ada2b associates with chromatin independently of the SAGA complex (Soffers et al., 2019). An additional function of Ada2, independent of Gcn5, was

identified in yeast (Jacobson and Pillus, 2009). It was found to promote transcriptional silencing at telomeres through binding to Sir2 to prevent the inward spread of heterochromatin regions.
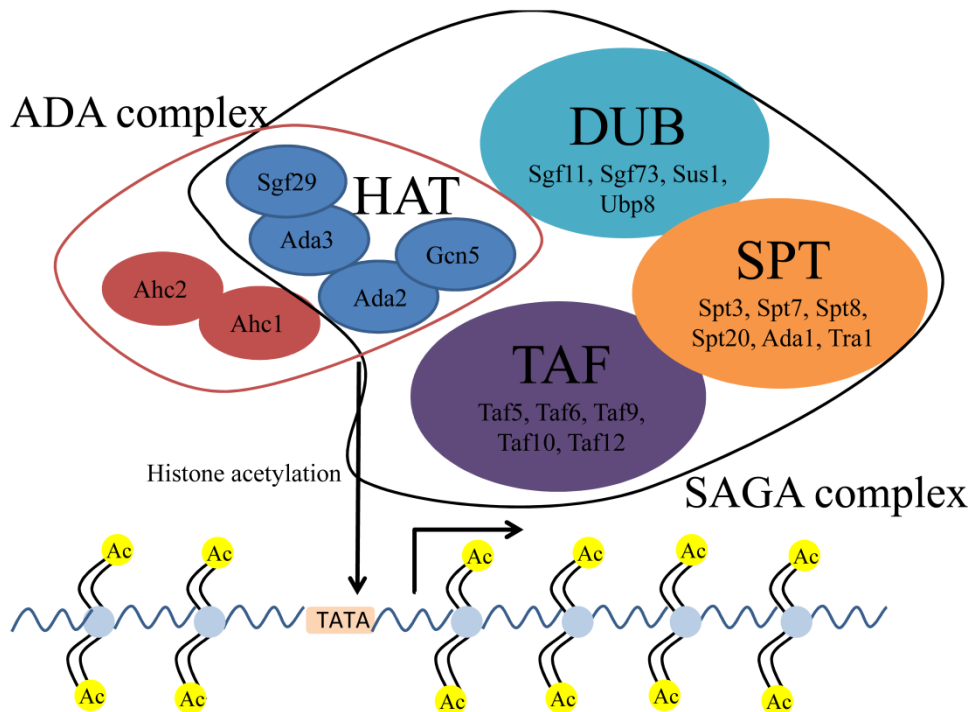

Figure 2. Schematic presentation of the structure of the SAGA and ADA complexes. Different colors are used to represent the different modules, such as the histone acetyltransferase (HAT) module, the histone deubiquitinase (DUB) module, the core structural (SPT) module, and the activator-binding (TAF) module. The HAT module of SAGA and ADA complex acetylates histones and associates with promoter regions, thus enabling chromatin remodeling and regulation of transcription.

Previous results in our group demonstrated that deletion of *ADA2* can increase AMT efficiency in yeast (Soltani, 2009) and combined with the fact that T-DNA integration makes use of DNA repair pathways, the role of *ADA2* in DNA repair should be considered. Whereas Ada2 is well known to be associated with the DNA damage response by various genome-wide screens (McKinney et al., 2013; Muñoz-Galván et al., 2013), its precise role and mechanism of actions still remain obscure in the context of DNA repair. It is not difficult to imagine the complicated situation because the expression of 2.5% of genes, in particular of RNA *PolII* genes, were found to be affected at least 2-fold in its deletion mutant (Hoke et al., 2008). Analogously, chromatin modifications are linked with DNA repair and expected to generate a favorable chromatin environment for the accession of proteins which are responsible for DNA repair (House et al., 2014). The histone acetyltransferase Gcn5 was shown to be important for DNA repair (Lee et al., 2010) and this ability of Gcn5 is facilitated and enhanced by Ada2 (Balasubramanian et al., 2002; Sun et al., 2018).

Recently, results obtained from research in *Arabidopsis* revealed that Ada2b not only binds to histone proteins to participate in chromatin modification but also interacts with the SMC5/6 complex to promote its recruitment to DSBs for DNA repair (Lai et al., 2018).

Meanwhile, SAGA was reported to maintain the monoubiquitinated H2B balance for efficient DNA repair through HR (Evangelista et al., 2018). In this case, the HAT module containing Ada2b is not directly involved in the maintenance of H2B mediated by SAGA, but the integrity of SAGA seems to be prerequisite for its ability.

Outline of the thesis

Agrobacterium-mediated transformation has been widely used for transformation of various eukaryotic cells, not only of plants, the natural host of *Agrobacterium*, but also of yeast and fungi. The yeast *S. cerevisiae* as a model host can provide new insights into the mechanism of AMT to contribute to the development of highly efficient transformation methods in the plant field. Hence, gaining more knowledge on the role of host and *Agrobacterium* factors in this process is important and useful.

In **Chapter 2** we showed that the absence of Ada2 in yeast can strongly enhance AMT efficiency with vectors that rely on HR for integration as well as those that depend on NHEJ. The *ada2Δ* deletion mutant has a reduced growth rate and was shown to be more susceptible for DNA damaging agents. Overexpression of *SFP1*, encoding a transcriptional regulator, rescued the growth deficiency and lowered the AMT efficiency of the *ada2Δ* deletion mutant.

Following this, in **Chapter 3** we developed a new approach to recover circularized extrachromosomal T-DNA structures from yeast and plants, especially from *ada2Δ* deletion mutant. The various T-circles were investigated for the fusion of left and right border repeat sequences and the presence of filler sequences originating either from T-DNA vectors or from the host genome.

In **Chapter 4** the Degron protein tag system was used to specifically degrade *Agrobacterium* virulence protein VirD2 when present in the plant or yeast nucleus. A putative role of VirD2 within the nucleus was investigated by comparing the AMT efficiency after degradation of VirD2 specifically in the host cell nucleus.

In **Chapter 5 and 6** the first chrysopine (pTiChry5) and succinamopine (pTiEU6) Ti-plasmids were sequenced and characterized. The development of new opine profiles may have conferred evolutionary advantage on their host bacteria in some specific environments. Both of these plasmids turned out to be chimeric due to recombination with other related plasmids.

References

Atmakuri, K., Cascales, E., Burton, O. T., Banta, L. M., & Christie, P. J. (2007). *Agrobacterium* ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *The EMBO Journal*, 26, 2540-2551.

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

Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A., & Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *Journal of Biological Chemistry*, 277, 7989-7995.

Ballas, N., & Citovsky, V. (1997). Nuclear localization signal binding protein from *Arabidopsis* mediates nuclear import of *Agrobacterium* VirD2 protein. *Proceedings of the National Academy of Sciences*, 94, 10723-10728.

Baptista, T., Grünberg, S., Minoungou, N., Koster, M. J., Timmers, H. M., Hahn, S., & Tora, L. (2017). SAGA is a general cofactor for RNA polymerase II transcription. *Molecular cell*, 68, 130-143.

Berg, M. D., Genereaux, J., Karagiannis, J., & Brandl, C. J. (2018). The pseudokinase domain of *Saccharomyces cerevisiae* Tra1 is required for nuclear localization and incorporation into the SAGA and NuA4 complexes. *G3: Genes, Genomes, Genetics*, 8, 1943-1957.

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

Camilleri, C., & Jouanin, L. (1991). The TR-DNA region carrying the auxin synthesis genes of the *Agrobacterium rhizogenes* agropine-type plasmid pRiA4: nucleotide sequence analysis and introduction into tobacco plants. Molecular plant-microbe interactions, 4, 155-162.

Chilton, W. S., Tempé, J., Matzke, M., & Chilton, M. D. (1984). Succinamopine: a new crown gall opine. *Journal of bacteriology*, 157, 357-362.

Chilton, W. S., Hood, E., Rinehart Jr, K. L., & Chilton, M. D. (1985). L, L-Succinamopine: an epimeric crown gall opine. *Phytochemistry*, 24, 2945-2948.

Chilton, W. S., Hood, E., & Chilton, M. D. (1985). Absolute stereochemistry of leucinopine, a crown gall opine. *Phytochemistry*, 24, 221-224.

Chilton, W. S., Stomp, A. M., Beringue, V., Bouzar, H., Vaudequin-Dransart, V., Petit, A., & Dessaux, Y. (1995). The chrysopine family of Amadori-type crown gall opines. *Phytochemistry*, 40, 619-628.

Chilton, W. S., Petit, A., Chilton, M. D., & Dessaux, Y. (2001). Structure and characterization of the crown gall opines heliopine, vitopine and rideopine. *Phytochemistry*, 58, 137-142.

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

Citovsky, V., Zupan, J., Warnick, D., & Zambryski, P. (1992). Nuclear localization of *Agrobacterium* VirE2 protein in plant cells. *Science*, 256, 1802-1805.

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

Costechareyre, D., Rhouma, A., Lavire, C., Portier, P., Chapulliot, D., Bertolla, F., & Nesme, X. (2010). Rapid and efficient identification of *Agrobacterium* species by recA allele analysis. *Microbial Ecology*, 60, 862-872.

Crane, Y. M., & Gelvin, S. B. (2007). RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *Proceedings of the National Academy of Sciences*, 104, 15156-15161.

Davioud, E., Petit, A., Tate, M. E., Ryder, M. H., & Tempé, J. (1988). Cucumopine—a new T-DNA-encoded opine in hairy root and crown gall. *Phytochemistry*, 27, 2429-2433.

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

Dessaux, Y., Petit, A., & Tempe, J. (1993). Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry*, 34, 31-38.

Eberharter, A., Sterner, D. E., Schieltz, D., Hassan, A., Yates, J. R., Berger, S. L., & Workman, J. L. (1999). The ADA complex is a distinct histone acetyltransferase complex in *Saccharomyces cerevisiae*. *Molecular and cellular biology*, 19, 6621-6631.

Evangelista, F. M., Maglott-Roth, A., Stierle, M., Brino, L., Soutoglou, E., & Tora, L. (2018). Transcription and mRNA export machineries SAGA and TREX-2 maintain monoubiquitinated H2B balance required for DNA repair. *Journal of cell biology*, 217, 3382-3397.

Farrand, S. K., van Berkum, P. B., & Oger, P. (2003). *Agrobacterium* is a definable genus of the family *Rhizobiaceae*. *International Journal of Systematic and Evolutionary Microbiology*, 53, 1681-1687.

Gamper, A. M., Kim, J., & Roeder, R. G. (2009). The SAGA subunit ADA2b is an important regulator of human GCN5 catalysis. *Molecular and cellular biology*, 29, 266-280.

García-Rodríguez, F. M., Schrammeijer, B., & Hooykaas, P. J. (2006). The *Agrobacterium* VirE3 effector protein: a potential plant transcriptional activator. *Nucleic acids research*, 34, 6496-6504.

Gelvin, S. B. (2003). *Agrobacterium*-Mediated Plant Transformation: the Biology behind the "Gene-Jockeying" Tool. *Microbiology and Molecular Biology Reviews*, 67, 16-37.

Gelvin, S. B. (2017). Integration of *Agrobacterium* T-DNA into the plant genome. *Annual review of genetics*, 51: 195-217.

Goodner, B., Hinkle, G., Gattung, S., Miller, N., Blanchard, M., Qurollo, B., Mullin, L, 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science*, 294, 2323-2328.

Gordon, J. E. and P. J. Christie (2014). The *Agrobacterium* Ti Plasmids. *Microbiology Spectrum*, 2, 18.

Grant, P. A., Duggan, L., Côté, J., Roberts, S. M., Brownell, J. E., Candau, R., & Berger, S. L. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & development*, 11, 1640-1650.

Hark, A. T., Vlachonasios, K. E., Pavangadkar, K. A., Rao, S., Gordon, H., Adamakis, I. D., & Triezenberg, S. J. (2009). Two *Arabidopsis* orthologs of the transcriptional coactivator ADA2 have distinct biological functions. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1789, 117-124.

Helmlinger, D., & Tora, L. (2017). Sharing the SAGA. *Trends in biochemical sciences*, 42, 850-861.

Henkel, C. V., den Dulk-Ras, A., Zhang, X., Hooykaas, P. J, 2014. Genome sequence of the octopine-type *Agrobacterium tumefaciens* strain Ach5. *Genome announcements*, 2, e00225-14.

Hoke, S. M., Genereaux, J., Liang, G., & Brandl, C. J. (2008). A conserved central region of yeast Ada2 regulates the histone acetyltransferase activity of Gcn5 and interacts with phospholipids. *Journal of molecular biology*, 384, 743-755.

Hooykaas, P. J. J., van Heusden, G. P. H., Niu, X., Roushan, M. R., Soltani, J., Zhang, X., & van der Zaal, B. J. (2018). *Agrobacterium*-Mediated Transformation of Yeast and Fungi, 349-374.

House, N., Koch, M. R., & Freudenreich, C. H. (2014). Chromatin modifications and DNA repair: beyond double-strand breaks. *Frontiers in genetics*, 5, 296.

Howard, E. A., Zupan, J. R., Citovsky, V., & Zambryski, P. C. (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.

Huang, Y. Y., Cho, S. T., Lo, W. S., Wang, Y. C., Lai, E. M., Kuo, C. H, 2015. Complete genome sequence of *Agrobacterium tumefaciens* Ach5. *Genome announcements*, 3, e00570-15.

Huisinga, K. L., & Pugh, B. F. (2004). A genome-wide housekeeping role for TFIID and a highly regulated stress-related role for SAGA in *Saccharomyces cerevisiae*. *Molecular cell*, 13, 573-585.

Jacobson, S., & Pillus, L. (2009). The SAGA subunit Ada2 functions in transcriptional silencing. *Molecular and cellular biology*, 29, 6033-6045.

Jia, Q., Bundock, P., Hooykaas, P. J. J., & De Pater, B. S. (2012). *Agrobacterium tumefaciens* T-DNA integration and gene targeting in *Arabidopsis thaliana* non-homologous end-joining mutants. *Journal of Botany*, 2012.

Jia, Q., den Dulk-Ras, A., Shen, H., Hooykaas, P. J. J., & de Pater, B. S. (2013). Poly (ADP-ribose) polymerases are involved in microhomology mediated back-up non-homologous end joining in *Arabidopsis thaliana*. *Plant molecular biology*, 82, 339-351.

Keane, P. J., Kerr, A., & New, P. B. (1970). Crown gall of stone fruit II. Identification and nomenclature of *Agrobacterium* isolates. *Australian Journal of Biological Sciences*, 23, 585-596.

Kim, S. I., Veena, & Gelvin, S. B. (2007). Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *The Plant Journal*, 51, 779-791.

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

Konstantinidis, K. T., & Tiedje, J. M. (2005). Genomic insights that advance the species definition for prokaryotes. *Proceedings of the National Academy of Sciences*, 102, 2567-2572.

Kusch, T., Guelman, S., Abmayr, S. M., & Workman, J. L. (2003). Two *Drosophila* Ada2 homologues function in different multiprotein complexes. *Molecular and cellular biology*, 23, 3305-3319.

Kuzmanović, N., & Puławska, J. (2019). Evolutionary Relatedness and Classification of Tumour-Inducing and Opine-Catabolic Plasmids in Three *Rhizobium rhizogenes* Strains Isolated from the Same Crown Gall Tumour. *Genome biology and evolution*.

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

Lacroix, B., & Citovsky, V. (2015). Nopaline-type Ti plasmid of *Agrobacterium* encodes a VirF-like functional F-box protein. *Scientific reports*, 5, 16610.

Lai, J., Jiang, J., Wu, Q., Mao, N., Han, D., Hu, H., & Yang, C. (2018). The transcriptional coactivator ADA2b recruits a structural maintenance protein to double-strand breaks during DNA repair in plants. *Plant physiology*, 176, 2613-2622.

Lee, H. S., Park, J. H., Kim, S. J., Kwon, S. J., & Kwon, J. (2010). A cooperative activation loop among SWI/SNF, γ-H2AX and H3 acetylation for DNA double-strand break repair. *The EMBO journal*, 29, 1434-1445.

Lee, K. K., Sardiu, M. E., Swanson, S. K., Gilmore, J. M., Torok, M., Grant, P. A., & Washburn, M. P. (2011). Combinatorial depletion analysis to assemble the network architecture of the SAGA and ADA chromatin remodeling complexes. *Molecular systems biology*, 7, 503.

Lee, L. Y., Wu, F. H., Hsu, C. T., Shen, S. C., Yeh, H. Y., Liao, D. C., Fang, M. J., Liu, N. T., Yen, Y. C., Dokladal, L., Sýkorová, E., Gelvin, S. B., & Lin, C. S. (2012). Screening a cDNA library for protein–protein interactions directly in planta. *The Plant Cell*, tpc-112.

Lindström, K., & Young, J. P. W. (2011). International Committee on Systematics of Prokaryotes Subcommittee on the taxonomy of *Agrobacterium* and *Rhizobium*. *International journal of systematic and evolutionary microbiology*, 61, 3089-3093.

Mestiri, I., Norre, F., Gallego, M. E., & White, C. I. (2014). Multiple host-cell recombination pathways act in *Agrobacterium*-mediated transformation of plant cells. *The Plant Journal*, 77, 511-520.

McKinney, J. S., Sethi, S., Tripp, J. D., Nguyen, T. N., Sanderson, B. A., Westmoreland, J. W., & Lewis, L. K. (2013). A multistep genomic screen identifies new genes required for repair of DNA double-strand breaks in *Saccharomyces cerevisiae*. *BMC genomics*, 14, 251.

Mousavi, S. A., Willems, A., Nesme, X., de Lajudie, P., & Lindström, K. (2015). Revised phylogeny of *Rhizobiaceae*: proposal of the delineation of *Pararhizobium* gen. nov., and 13 new species combinations. *Systematic and applied microbiology*, 38, 84-90.

Muñoz-Galván, S., Jimeno, S., Rothstein, R., & Aguilera, A. (2013). Histone H3K56 acetylation, Rad52, and non-DNA repair factors control double-strand break repair choice with the sister chromatid. *PLoS genetics*, 9, e1003237.

Muratoglu, S., Georgieva, S., Pápai, G., Scheer, E., Enünlü, I., Komonyi, O., Cserpan, I., Lebedeva, L., Nabirochjina, E., Udvardy, A., Tora, L., & Boros, I. (2003). Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Molecular and cellular biology*, **23,** 306-321.

Nishizawa-Yokoi, A., Nonaka, S., Saika, H., Kwon, Y. I., Osakabe, K., & Toki, S. (2012). Suppression of Ku70/80 or Lig4 leads to decreased stable transformation and enhanced homologous recombination in rice. *New Phytologist*, 196, 1048-1059.

Niu, X., Zhou, M., Henkel, C. V., van Heusden, G. P. H., & Hooykaas, P. J. J. (2015). The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. *The Plant Journal*, 84, 914-924.

Ohmine, Y., Kiyokawa, K., Yunoki, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2018). Successful Transfer of a Model T-DNA Plasmid to *E. coli* Revealed Its Dependence on Recipient RecA and the Preference of VirD2 Relaxase for Eukaryotes Rather Than Bacteria as Recipients. *Frontiers in Microbiology*, 9.

Oger, P., Reich, C., Olsen, G. J., & Farrand, S. F. (2001). Complete nucleotide sequence and analysis of pTiBo542: what genomics tells us about structure and evolution of plasmids in the family *Rhizobiaceae*. *Plasmid*, 45: 169-170.

Ohmine, Y., Satoh, Y., Kiyokawa, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2016). DNA repair genes *RAD52* and *SRS2*, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene *ERG28* are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. *BMC microbiology*, **16:** 58.

Ormeno-Orrillo, E., Servín-Garciadueñas, L. E., Rogel, M. A., González, V., Peralta, H., Mora, J., & Martínez-Romero, E. (2015). Taxonomy of *rhizobia* and *agrobacteria* from the *Rhizobiaceae* family in light of genomics. *Systematic and applied microbiology*, 38, 287-291.

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. *Proceedings of the National Academy of Sciences*, 90,11538–42.

Petit, A., Delhaye, S., Tempé, J., & Morel, G. (1970). Recherches sur les guanidines des tissus de crown gall. Mise en evidence d'une relation biochimique specifique entre les souches d'Agrobacterium tumefaciens et les tumeurs qu'elles induisent. *Physiol. vég*, 8, 205-213.

Ramírez-Bahena, M. H., Vial, L., Lassalle, F., Diel, B., Chapulliot, D., Daubin, V., Nesme, X., & Muller, D. (2014). Single acquisition of protelomerase gave rise to speciation of a large and diverse clade within the *Agrobacterium/Rhizobium* supercluster characterized by the presence of a lin chromid. *Molecular phylogenetics and evolution*, 73, 202-207.

Rolloos, M., Dohmen, M. H., Hooykaas, P. J. J., & van der Zaal, B. J. (2014). Involvement of Rad52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae*. *Molecular microbiology*, 91, 1240-1251.

Saleh, A., Schieltz, D., Ting, N., McMahon, S. B., Litchfield, D. W., Yates, J. R., & Brandl, C. J. (1998). Tra1p is a component of the yeast Ada Spt transcriptional regulatory complexes. *Journal of Biological Chemistry*, 273, 26559-26565.

Schrammeijer, B., Dulk-Ras, A. D., Vergunst, A. C., Jurado Jácome, E., & Hooykaas, P. J. J. (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 research*, 31, 860-868.

Schrammeijer, B., Hemelaar, J., & Hooykaas, P. J. J. (1998). The presence and characterization of a *virF* gene on *Agrobacterium vitis* Ti plasmids. *Molecular plant-microbe interactions*, 11, 429-433.

Schrammeijer, B., Risseeuw, E., Pansegrau, W., Regensburg-Tuïnk, T. J., Crosby, W. L., & Hooykaas, P. J. J. (2001). Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. *Current Biology*, 11, 258-262.

Shilo, S., Tripathi, P., Melamed-Bessudo, C., Tzfadia, O., Muth, T. R., & Levy, A. A. (2017). T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. *PLoS genetics*, 13, e1006875.

Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. (2012). Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant physiology*, 160, 511-522.

Singer, K. (2018). The Mechanism of T-DNA Integration: Some Major Unresolved Questions. *Agrobacterium* Biology: From Basic Science to Biotechnology, 287-317.

Slater, S. C., Goldman, B. S., Goodner, B., Setubal, J. C., Farrand, S. K., Nester, E. W., & Otten, L. (2009). Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *Journal of bacteriology*, 191, 2501-2511.

Soffers, J. H., Li, X., Saraf, A., Seidel, C. W., Florens, L., Washburn, M. P., & Workman, J. L. (2019). Characterization of a metazoan ADA acetyltransferase complex. *Nucleic acids research*, 47, 3383-3394.

Soltani, J. (2009). Host genes involved in *Agrobacterium*-mediated transformation. Ph.D. thesis, Leiden University, Leiden, The Netherlands.

Soltani, J., Van Heusden, G. P. H., Hooykaas, P. J. J. (2009). Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*. *FEMS microbiology letters*, 298: 228-233.

Stachel, S. E., Timmerman, B., & Zambryski, P. (1986). Generation of single-stranded T-DNA molecules during the initial stages of T-DNA transfer from *Agrobacterium tumefaciens* to plant cells. *Nature*, 322, 706-712.

Sun, J., Paduch, M., Kim, S. A., Kramer, R. M., Barrios, A. F., Lu, V., & Tan, S. (2018). Structural basis for activation of SAGA histone acetyltransferase Gcn5 by partner subunit Ada2. *Proceedings of the National Academy of Sciences*, 115, 10010-10015.

Suzuki, K., Hattori, Y., Uraji, M., Ohta, N., Iwata, K., Murata, K., Kato, A., & Yoshida, K. (2000). Complete nucleotide sequence of a plant tumor-inducing Ti plasmid. Gene, 242(1-2), 331-336.

Szegedi, E., Czakó, M., Otten, L., & Koncz, C. S. (1988). Opines in crown gall tumours induced by biotype 3 isolates of *Agrobacterium tumefaciens*. *Physiological and molecular Plant pathology*, 32, 237-247.

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

Tinland, B., Schoumacher, F., Gloeckler, V., Bravo-Angel, A. M., & Hohn, B. (1995). The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *The EMBO journal*, 14, 3585-3595.

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

van Attikum, H., & Hooykaas, P. J. J. (2003). Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic acids research*, 31, 826-832.

van Kregten, M., Lindhout, B. I., Hooykaas, P. J. J., & van der Zaal, B. J. (2009). *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Molecular plant-microbe interactions*, 22, 1356-1365.

van Kregten, M., de Pater, B. S., Romeijn, R., van Schendel, R., Hooykaas, P. J. J., & Tijsterman, M. (2016). T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nature plants*, 2, 16164.

Vergunst, A. C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C. M., Regensburg-Tuïnk, T. J., & Hooykaas, P. J. J. (2000). VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science*, 290, 979-982.

Vergunst, A. C., van Lier, M. C., den Dulk-Ras, A., Stüve, T. A. G., Ouwehand, A., & Hooykaas, P. J. J. (2005). Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proceedings of the National Academy of Sciences*, 102, 832-837.

Ward, E. R., & Barnes, W. M. (1988). VirD2 protein of Agrobacterium tumefaciens very tightly linked to the 5'end of T-strand DNA. *Science*, 242, 927.

Wetzel, M. E., Olsen, G. J., Chakravartty, V., & Farrand, S. K. (2015). The *repABC* plasmids with quorum-regulated transfer systems in members of the *rhizobiales* divide into two structurally and separately evolving groups. *Genome biology and evolution*, 7, 3337-3357.

Winans, S. C. (1992). Two-way chemical signaling in *Agrobacterium*-plant interactions. *Microbiology and Molecular Biology Reviews*, 56, 12-31.

Wolterink-van Loo, S., Ayala, A. A. E., Hooykaas, P. J. J., & van Heusden, G. P. H. (2015). Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. *Microbiology*, 161, 401-410.

Wood, D. W., Setubal, J. C., Kaul, R., Monks, D. E., Kitajima, J. P., Okura, V. K., Woo, L, (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science*. 294, 2317-2323.

Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M., & Chilton, M. D. (1982). Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proceedings of the National Academy of Sciences*, 79, 6322-6326.

Yanofsky, M. F., Porter, S. G., Young, C., Albright, L. M., Gordon, M. P., & Nester, E. W. (1986). The virD operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell*, 47: 471-477.

Young, C., & Nester, E. W. (1988). Association of the virD2 protein with the 5'end of T strands in *Agrobacterium tumefaciens*. *Journal of bacteriology*, 170, 3367-3374.

Young, J. M., Kuykendall, L. D., Martinez-Romero, E., Kerr, A., & Sawada, H. (2003). Classification and nomenclature of *Agrobacterium* and *Rhizobium*–a reply to Farrand et al. *International journal of systematic and evolutionary microbiology*, 53, 1689-1695.

Zhang, X., van Heusden, G. P. H., & Hooykaas, P. J. J. (2017). Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. *Proceedings of the National Academy of Sciences*, 114, 10238-10243.

Zhang, X., & Hooykaas, P. J. J. (2019). The *Agrobacterium* VirD5 protein hyper-activates the mitotic Aurora kinase in host cells. *New Phytologist*, 1-10.

Zhu, Y., Nam, J., Humara, J. M., Mysore, K. S., Lee, L. Y., Cao, H., ... & Hwang, H. H. (2003). Identification of *Arabidopsis* rat mutants. *Plant physiology*, 132, 494-505.

Zhu, J., Oger, P. M., Schrammeijer, B., Hooykaas, P. J. J. , Farrand, S. K., & Winans, S. C. (2000). The bases of crown gall tumorigenesis. *Journal of Bacteriology*, 182, 3885-3895.

Ziemienowicz, A., Tinland, B., Bryant, J., Gloeckler, V., & Hohn, B. (2000). Plant enzymes but not *Agrobacterium* VirD2 mediate T-DNA ligation *in vitro*. *Molecular and cellular biology*, 20, 6317-6322.

Chapter 2
Analysis of the function of *ADA2* in double strand break repair and *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*

Shuai Shao, Daniela M. d'Empaire Altimari, Paul J. J. Hooykaas, G. Paul H. van Heusden

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Abstract

Agrobacterium tumefaciens is the causative agent of crown gall disease and has been extensively used as a vector to create transgenic plants. *Agrobacterium*-mediated transformation (AMT) has become the preferred method for genetic transformation of plants and some fungi and plays critical roles in worldwide agriculture. Under laboratory conditions the yeast *Saccharomyces cerevisiae* can be transformed as well, hence, the mechanism of AMT has been extensively studied using *S. cerevisiae* as model host. Here we report that the absence of *ADA2*, encoding a histone acetyltransferase, results in increased efficiency of AMT by both targeted and random integration. The phenotype of the *ada2Δ* deletion mutant was characterized and we showed that *ADA2* participates in regulation of the cell cycle and DNA damage responses. Overexpression of *SFP1* which is involved in the activation of transcription of ribosomal protein genes, can rescue the methyl methanesulfonate (MMS) sensitivity of the *ada2Δ* deletion mutant. Furthermore, overexpression of *SFP1* attenuates the increased efficiency of AMT of the *ada2Δ* deletion mutant. These results confirm an important role of histone acetylation-deacetylation in AMT. However, more research is needed to elucidate the exact mechanism and to understand the role of *SFP1* in AMT.

Introduction

The soil pathogen *Agrobacterium tumefaciens* is renowned for its ability to transform a broad range of plant species in nature and a lot of non-plant eukaryotic hosts including the yeast *Saccharomyces cerevisiae* under laboratory conditions (Bundock et al., 1995; De Groot et al., 1998). This unique ability of *Agrobacterium* and the high efficiency of *Agrobacterium*-mediated transformation (AMT) made this bacterium essential for plant biology research and made AMT the preferred technology for plant transformations in agricultural industry (for review see: Nester et al., 1984; Gelvin, 2003; Tzfira and Citovsky, 2006; Păcurar et al., 2011; Christie and Gordon, 2014; Gelvin, 2017).

AMT consists of several steps to transfer a DNA segment into the host cell nucleus. This transferred DNA (T-DNA), derived from the *A. tumefaciens* tumor-inducing plasmid (Ti-plasmid), contains genes involved in the synthesis of auxin, cytokinin and opines resulting in uncontrolled cell proliferation and production of opines. Bacterial factors involved in AMT have been characterized in detail (for review see: Christie and Gordon, 2014; Gelvin, 2017). A set of essential proteins encoded by virulence genes (*vir*) present on Ti-plasmids are employed to facilitate this process. Initially, a two-component sensory-response system (VirA/VirG) is triggered to induce the expression of *vir* genes (Winans, 1992). VirD2, together with VirD1, nicks at two 25bp direct repeat border sequences (RB, right border and LB, left border) which flank the T-DNA (Yadav et al., 1982), yielding a single-stranded copy of T-DNA (T-strand) covalently attached to VirD2 at its 5' end (Ward and Barnes, 1988). This complex is transported into the host through a type IV secretion system (T4SS) composed of the VirB1-11/VirD4 proteins (Zhu et al., 2000). Ultimately the T-DNA is integrated into chromosomal DNA via non-homologous recombination (Offringa et al., 1990) by theta-mediated end-joining (TMEJ) with involvement of Polymerase θ (van Kregten et al., 2016). In yeast the T-DNA is preferentially integrated by homologous recombination (HR) (Bundock et al., 1995; van Attikum and Hooykaas, 2003). Alternatively, T-DNA molecules can form complex extrachromosomal structures such as T-circles (Singer et al., 2012; Rolloos et al., 2014) and maintained if possessing a replicator (Bundock et al., 1995). In parallel, several effector proteins (VirE2, VirE3, VirD5 and VirF) are delivered into host cells independently of the T-DNA and contribute to the transformation process (Tzfira et al., 2001; Niu et al., 2015; Schrammeijer et al., 2001; Zhang et al., 2017).

The involvement of host factors in AMT is still not completely understood. Several chromatin components or chromatin-modifying enzymes were identified which play a role in T-DNA integration in plants, such as histone H2A (Mysore et al., 2000) and H3 (Anand et al., 2007) and the histone deacetylases HDT1 and HDT2 (Crane and Gelvin, 2007). The yeast *S. cerevisiae* has been used as an excellent model host to investigate the role of host factors during AMT. More and more host factors involved in AMT have been uncovered by using the yeast complete non-essential gene deletion mutant collections. Histone acetyltransferases (Gcn5, Ngg1, Yaf9 and Eaf7) and deacetylases (Hst4, Hda2 and Hda3) involved in chromatin modification, were identified as factors affecting AMT (Soltani et al., 2009). *ARP6* encoding an actin-related protein that is part of the SWR1 chromatin remodeling complex, negatively regulates AMT (Luo et al., 2015). The cell wall synthesis regulator Smi1 and membrane sterol synthesis scaffold protein Erg28 are also important for efficient AMT (Ohmine et al., 2016). In addition, some proteins can not only affect the efficiency of AMT but also play important roles

in the maintenance of extrachromosomal T-DNA structures. It has been reported that *RAD52*, *RAD51* and *SRS2* are involved in the formation of T-circles in yeast and are necessary for their maintenance (Rolloos et al., 2014; Ohmine et al., 2016).

The Ada2 protein is the chromatin-binding subunit of the SAGA (Spt–Ada–Gcn5 acetyltransferase) histone acetyltransferase complex. This complex is involved in the post-translational modifications of histones that are crucial for chromatin-dependent functions and the regulation of numerous cellular processes in response to environmental cues (Sterner et al., 2002). Ada2 can interact with Gcn5 directly to increase its histone acetyltransferase (HAT) activity which preferentially acetylates histone H3 and histone H2B (Grant et al., 1997; Hoke et al., 2008). Ada2 is evolutionarily conserved among eukaryotes and has been described in several organisms, including *Arabidopsis* (Hark et al, 2009) and *Drosophila* (Muratoglu et al., 2003). In *Arabidopsis*, the orthologs of Ada2 physically associate with Gcn5 and enhance its HAT activity to regulate gene expression under environmental stress conditions such as cold, drought and salt stress (Hark et al, 2009). In 2009, an additional function of Ada2, independent of Gcn5, was identified in yeast. The novel role of Ada2 was to promote transcriptional silencing at telomeres through binding to Sir2 and to prevent the inward spread of heterochromatin regions (Jacobson and Pillus, 2009).

Previous results in our group demonstrated that the deletion of *ADA2* can increase AMT efficiency (Soltani, 2009). In this chapter, we studied the role of *ADA2* in AMT in more detail. To this end, we further characterized the phenotype of the *ada2Δ* deletion mutant and searched for extragenic suppressors of the *ada2Δ* deletion and showed that overexpression of *SFP1*, encoding a transcriptional regulator, suppressed the increased AMT efficiency of *ada2Δ* deletion mutants.

Methods

Agrobacterium strains and growth conditions

Agrobacterium tumefaciens strain LBA1100, a Ti-plasmid-cured derivative of strain C58 containing a *vir* helper plasmid, was used in this chapter (Beijersbergen et al., 1992). All the binary vector plasmids were introduced into *Agrobacterium* by electroporation (den Dulk-Ras and Hooykaas, 1995). *A. tumefaciens* strains were grown and maintained at 29°C in LC medium (10 g/L tryptone, 5 g/L yeast extract and 8 g/L NaCl) containing appropriate antibiotics carbenicillin (100 µg/mL), kanamycin (100 µg/mL) or rifampicin (20 µg/mL) when required.

Yeast strains and growth conditions

S. cerevisiae strains used in this study can be found in Table 1. The deletion mutants were constructed using the PCR-mediated one-step gene disruption method (Baudin et al., 1993). All disruption cassettes and selective markers were amplified from the plasmids listed in Table 2 and the primers used are in Table 3. For labeling of *RAD52*, plasmid pYM27 was used to provide PCR templates for C-terminal EGFP tagging and the primers Rad52-gfp-Fw/Rev were used. For disruption of *ADA2* with the HphMX4 cassette, plasmid pAG32 and primers Ada2-Fw/Rev were used. The PCR products and plasmids were transferred to yeast cells using the lithium-acetate transformation protocol (Gietz et al., 1995). Yeast was grown at 30°C in yeast extract-peptone-dextrose (YPD) medium supplemented, when required, with the appropriate antibiotic G418 (200 µg/mL) or hygromycin (200 µg/mL) or in selective minimal yeast (MY)

medium supplemented with appropriate nutrients. 5-fluoroanthranilic acid (FAA, 500 µg/mL) and 5-fluoroorotic acid (FOA, 500 µg/mL) were used for the counter-selection of tryptophan autotrophs and uracil autotrophs, respectively.

For spot plate assays cultures were adjusted to an OD_{620} of 0.1 after growth to saturation in liquid YPD or MY medium supplemented with appropriate nutrients. Then, ten-fold serial dilutions were made and aliquots of 5 µl were spotted. Dilution assays were incubated at 30°C, except where noted. HU sensitivity was analyzed with serial dilutions from 5 mg/ml. MMS sensitivity was analyzed with serial dilutions from 0.005% to 0.02% MMS and eventually 0.012% MMS was used to screen for genes which can rescue the lethality of *ada2Δ* deletion mutants.

Plasmid construction

The plasmids used in this chapter are listed in Table 2 and the primers are presented in Table 3. To visualize T-DNA via homologous recombination, initially *TRP1* flanking sequences (575 bp sequence at 5'end and 482 bp sequence at 3'end) were amplified from yeast genomic DNA using primers Trp1-up-Fw/Rev and Trp1-down-Fw/Rev and cloned into the *Sac*I/*Bam*HI and *Pst*I/*Hind*III sites, respectively, of the *Agrobacterium* binary vector pCAMBIA2300 to generate the binary vector pCAMBIA2300[*trp1*]. Subsequently, the *lacO* repeats were introduced into T-DNA by cloning the *Bam*HI-*Sal*I fragment with the tandem array of 256 copies of the Lac operator from pAFS59 into vector pCAMBIA2300[*trp1*] to yield pCAMBIA2300[*trp1-lacO*]. Similarly, the *Bam*HI-*Sal*I *lacO* repeats was cloned into vector pCAMBIA2300 to generate pCAMBIA2300[*lacO*]. The test vector pRS425[*lacO*] was constructed by cloning the tandem array of *lacO* repeats digested with *Xho*I and *Sal*I restriction enzymes from pAFS59 into the *Sal*I site of pRS425. All plasmids containing *lacO* repeats were propagated in the *E.coli* STBL2 strain.

To express GFP-Lac repressor fusions in yeast, the segment containing GFP-LacI-NLS (nuclear localization signal) was amplified from the plasmid pAFS152 by PCR using primers LacI-Fw/Rev and cloned into pRS305 digested with *Spe*I and *Sal*I restriction enzymes yielding pRS305-LacI-GFP. Subsequently, the plasmid was linearized by *Nhe*I digestion and integrated into the *LEU2* locus of yeast strain BY4741. To distinguish integrated and extrachromosomal T-DNA, the tandem array of *lacO* repeats were cloned into pSDM8001 between the *PDA1* downstream flank and the right border as well. The tandem array of *lacO* repeats was digested with *Xho*I and *Sal*I restriction enzymes from pAFS59 and then cloned into *Xho*I site of pSDM8001 to generate pSDM8001-*lacO*.

In order to investigate the AMT efficiency of *ada2Δ* deletion mutant, a T-DNA containing *TRP1* flanking sequences was constructed. In brief, the *Xho*I/*Xba*I fragment containing the KanMX marker cassette was digested from plasmid pUG6 and then inserted between the *TRP1* flanking sequences with *Sal*I/*Xba*I sites of pCAMBIA2300[*trp1*] resulting in plasmid pCAMBIA2300[*trp1*-kanMX]. Once T-DNA is integrated into the *TRP1* locus by a double crossover, the *TRP1* will be disrupted and the transformants can be selected on the FAA plates.

Isolation of multi-copy suppressors

The *ada2Δ* strain was transformed with a yeast genomic DNA library in the multi-copy plasmid YEp24 (Carlson and Botstein, 1982) using the standard LiAc method with a few modifications for improving transformation efficiency, such as extending incubation time to 30 or 40 minutes at 42°C. A total of 8,000 transformants were selected on selective minimal medium lacking

uracil (MY-Ura) plates after incubation at 30°C for 2-3 days and then replica-plated onto MY-Ura plates with or without 0.012% MMS. The screen was repeated three independent times. After that, all colonies that grew on both plates were restreaked and candidates were retested at least twice. Subsequently, plasmids were isolated from the remaining 33 transformants using a Plasmid DNA Miniprep Kit (Thermo Fisher Scientific) with addition of lyticase to the resuspension buffer to disrupt cell walls. Then plasmids were propagated in *E. coli* DH10B, isolated and sequenced using general primers YEp24-Fw/Rev from the flanking sequences of the *BamH*I site. The corresponding genomic regions were identified by alignment with the *S. cerevisiae* genome DNA sequence (SGD, www.yeastgenome.org). Each candidate gene was amplified from genomic DNA of BY4741 by PCR using a set of primers 500 bp upstream and 500 bp downstream of the target ORF (Table 3) and the PCR products were cloned in YEp24. The constructed plasmids as well as the parental plasmid YEp24 were introduced into the *ada2Δ* deletion mutant and wild type, respectively, to test the sensitivity towards MMS as described above.

Agrobacterium mediated transformation efficiency test

AMT efficiency was determined as described by Bundock et al. (1995) with some modifications. Firstly, *S. cerevisiae* strains and *Agrobacterium* were cultured overnight at 30°C and 28°C, respectively, under continuous agitation and with the appropriate nutrition or antibiotic selection. The following day, the *Agrobacterium* cells were washed and re-suspended to an OD_{600} of 0.25 in induction medium (IM) with added glucose (10 mM), acetosyringone (AS, 0.2 mM) and the appropriate antibiotics, and incubated for another 6 hours at 28°C. Meanwhile, yeast cultures were diluted to an OD_{620} of 0.1 and incubated in either liquid YPD or MY (when the yeast contains a plasmid) medium. After 6 hours the yeast cells were washed and re-suspended in 0.5 ml of IM, to a final OD_{620} of 0.4 - 0.6 and mixed with an equal volume of *Agrobacterium* cells and vigorously vortexed. Subsequently, 100 μl of the mixture were pipetted onto sterile nitrocellulose filters laid on IM plates supplemented with histidine, leucine and methionine. Once filters were dry, plates were incubated at 21 °C for 6-7 days. After co-cultivation, the cells on each filter were resuspended and then spread onto solid medium containing cefotaxime (200 μg/mL) with or without G418 (200 μg/ml). Finally, after a 3-day incubation at 30 °C, colonies were counted. Yeast AMT efficiency was calculated by dividing the number of colonies on the selective plate by the number of colonies on the non-selective plates. For GFP signal visualization, yeast cells were co-cultivated 2-3 days with *Agrobacterium* cells following the AMT protocol as mentioned above, then washed and re-suspended in MilliQ water or MY medium.

Determination of growth rates

Growth curves were generated by growing cells overnight in MY at 30°C and then diluting them to an OD_{620} of around 0.1. The new cultures were grown at 30°C and the OD_{620} was measured at the indicated time points. All measurements were repeated at least three times in independent experiments.

Assay for cell cycle arrest

The cells used are derivatives of W303a carrying a deletion of the α-factor-inactivating protease, *BAR1*, rendering cells hypersensitive to pheromone and allowing for prolonged cell cycle arrest.

The *ada2Δ* deletion mutant in this genetic background was constructed as described above. The cells were diluted at 0.1 OD_{620} after culturing overnight. The α-factor (Sigma-Aldrich) was added at 1 µg/ml and yeast cells incubated with shaking at 30°C for 3 hrs. Then microscopy was used to check whether more than 90% of the cells were arrested at G1/S phase. To release cells from α-factor arrest, cells were spun down, washed four times and then re-suspended in fresh medium. Cells were collected every half an hour and analyzed by microscopy.

Confocal microscopy

For microscopy a Zeiss Imager M1 confocal microscope equipped with a LSM5 Exciter with a 63x oil objective was used. GFP was detected using an argon laser of 488 nm and a band-pass emission filter of 505-600 nm. Images were processed with ImageJ (ImageJ National Institute of Health) (Schindelin et al., 2012).

Statistical analysis

All data shown are representative of at least three independent experiments of which each contains at least three biological replicates and represented as mean of the performed experiments with standard deviation. Statistic test were done with two-tailed Student's t-test. Statistical analyses were performed using a function equipped in Microsoft Excel. An asterisk indicates the significant differences with the p-values <0.05.

Results

Increased *Agrobacterium*-mediated transformation efficiency of the yeast *ada2Δ* deletion mutant

Before studying the role of *ADA2* in AMT, we compared the transformation efficiency of three different yeast strains, W303a (haploid), BY4741 (haploid) and BY4743 (similar as BY4741, but diploid). To this end, these strains were transformed with the *Agrobacterium* strain LBA1100 carrying pRAL7100 allowing integration of *URA3* into the *PDA1* locus by homologous recombination (Figure 1A) (Bundock et al., 1995). As illustrated in Table S1A, the transformation efficiency of the diploid strain BY4743 is highest (1.2 ± 0.2 ×10^{-4} n=3) and the haploid strain BY4741 has a higher transformation efficiency than the haploid strain W303a at frequencies of 9.7 ± 0.3 ×10^{-5}(n=3) and 1.4 ± 0.04×10^{-5} (n=3), respectively. Compared to the frequencies of W303a, it was 6.9-times higher in BY4741 and 8.6-times higher in BY4743. Therefore, and because of the high level of auto-fluorescence of W303a cells, we used strain BY4741 and derived mutants for our further research. In addition, we prefer a haploid strain over a diploid strain because genetic manipulations are easier in haploid strains.

Previously we showed that the *ada2Δ* deletion mutant in the diploid BY4743 background has an increased transformation efficiency (Soltani et al., 2009). To investigate whether this is also the case in the haploid BY4741 background, we constructed an *ada2Δ* deletion mutant in BY4741 by replacing the coding sequences by an hygromycin resistance marker. As shown in Figure 1C and Table S1B, an *ada2Δ* deletion mutant in the haploid BY4741 background has a 4-fold increased transformation efficiency compared to the parental strain at frequencies of 1.1 ± 0.6 ×10^{-4} and 4.4 ± 0.8 ×10^{-4}, respectively. The effect of the *ada2Δ* deletion on T-DNA integration may be different for different chromosomal loci. To further investigate this possibility, we constructed a new binary vector containing T-DNA in which the

KAN.MX marker is flanked by sequences allowing integration by homologous recombination into the *TRP1* locus instead of into the *PDA1* locus (Figure 1B). The *TRP1* locus is physically close to the centromere on chromosome IV, whereas the *PDA1* locus is located at the end of chromosome V. When using *TRP1* as target, also counter-selection is possible allowing screening of transformants for loss of the *TRP1* gene due to the integration of the T-DNA (Toyn et al., 2000). Therefore, after co-cultivation, the yeast cells were forward-selected using the KAN.MX marker. Subsequently, transformants were replicated to plates supplemented with or without 5-fluoroanthranilic acid (FAA), which is used for the counterselection of *TRP1*. Using this T-DNA the transformation frequency of the *ada2Δ* strain was around 2-fold higher than that of the BY4741 strain ($1.2 \pm 0.7 \times 10^{-4}$ vs $0.6 \pm 0.3 \times 10^{-4}$) (Figure 1D, Table S1B). In the wild type strain, only few transformants (2.4 ± 0.5 %) could not survive on the plates supplemented with FAA suggesting that in most transformants the T-DNA was integrated into the *TRP1* locus disrupting the synthesis of tryptophan (Table S1C). In the *ada2Δ* background, around one third of transformants (27 ± 5 %) were not able to grow on plates containing FAA, indicating that in *ada2Δ* in a large number of the transformants the *TRP1* locus is still intact. This was confirmed by PCR for 7 of these colonies that were randomly chosen. This suggests that in the *ada2Δ* background AMT gives much more transformants (>10 fold) in which the T-DNA has not integrated at the *TRP1* locus in the chromosomal DNA.

Figure 1. Increased *Agrobacterium*-mediated transformation efficiency of the yeast *ada2Δ* deletion mutant by HR. Yeast strain BY4741*ada2Δ* and its parent strain BY4741 were co-cultivated with *Agrobacterium* strain LBA1100 harboring pRAL7100 or pCAMBIA2300[*trp1*-kanMx]. The schematic diagrams present the structure of the T-DNA of pRAL7100 (A) and of pCAMBIA2300[*trp1*-kanMx] (B). AMT efficiency for *Agrobacterium* strains carrying pRAL7100 and pCAMBIA2300[*trp1*-kanMx] are shown in panels C and D, respectively. Error bars indicate the standard deviations of triplicate independent assays.

Foreign DNA can be introduced into yeast cells by other methods which don't employ *Agrobacterium* such as LiAc transformation or electroporation. In a number of experiments we transformed BY4741 and the *ada2Δ* deletion mutant with linear DNA fragments like constructs allowing integration of GFP at the *RAD52* locus or with plasmids like the overexpression plasmid YEp24 using the LiAc transformation protocol. Although we did not study the effect of deletion of *ADA2* on the efficiency of the LiAc transformation in a systematic way, we observed a decreased rather than an increased transformation efficiency for the *ada2Δ* deletion mutant. In a transformation experiment aimed to tag *RAD52* with GFP (see below), around one hundred transformants were obtained for the wild type strain BY4741, whereas less than 20 transformants were obtained for the *ada2Δ* deletion mutant. The number of transformants of the *ada2Δ* deletion mutant after transformation with plasmid YEp24 was much lower than the number of transformants of BY4741 (approx. 20 *vs* approx. 100). Thus, the increased transformation efficiency of *ada2Δ* deletion mutant is specific for AMT in accordance with previous results from our group (Soltani, 2009)

Integration of T-DNA by NHEJ

Although homologous recombination is the predominant mechanism of T-DNA integration in yeast, integration via non-homologous end-joining is possible as well (Bundock and Hooykaas, 1996; van Attikum et al, 2001). In order to investigate the effect of the *ada2Δ* deletion on T-DNA integration via NHEJ, we exploited *Agrobacterium* strain LBA1100 harboring plasmid pRAL7102. This plasmid contains T-DNA with the *URA3* marker but has almost no homology with the BY4741 genome (see below) and no yeast replication origin (Figure 2A). As illustrated in Figure 2A, *Agrobacterium* carrying pRAL7102 is able to transform BY4741, but at an extremely low frequency of $2.4 \pm 1.4 \times 10^{-6}$. Compared with the wild type strain, the transformation efficiency ($1.0 \pm 0.2 \times 10^{-4}$) was much higher for the *ada2Δ* deletion mutant, indicating that also AMT via NHEJ is affected by the *ADA2* deletion. To prevent homology between T-DNA and recipient's genome, we used BY4741 derived strains as recipients in which the *URA3* locus had been removed completely (Figure 2B). However, a short sequence of 47 bp was found on T-DNA close to the RB that has homology to a region of yeast chromosome V adjacent to *URA3*. This may imply a small chance for T-DNA integration by HR through a single cross-over rather than a double cross-over. To investigate integration into the *URA3* locus, we analyzed the *URA3* locus in five randomly selected transformants by PCR. Using primer pair URA3-C / URA3-D a fragment of 150 bp is expected when the T-DNA is not integrated into the *URA3* locus, whereas a larger fragment is expected when the T-DNA is integrated at this locus (Figure 2C, left part). For all of the selected transformants, the expected 150 bp fragment was obtained, indicating that the chromosomal *URA3* locus was not altered, excluding integration of the T-DNA at the *URA3* locus. On the other hand, using primer combination URA3-A / URA3-B which has no homology to the BY4741 genome, but can amplify the *URA3* originating from the T-DNA, a band of 565 bp was found for all five transformants indicating the presence of the T-DNA (Figure 2C, right part).

In order to further investigate whether the T-DNA was integrated into one of the chromosomes or whether it was present as extrachromosomal structures, some of the transformants were streaked on plates containing 5-fluoro-orotic acid (5-FOA). It is expected that transformants with the T-DNA integrated into one of the chromosomes will not be able to
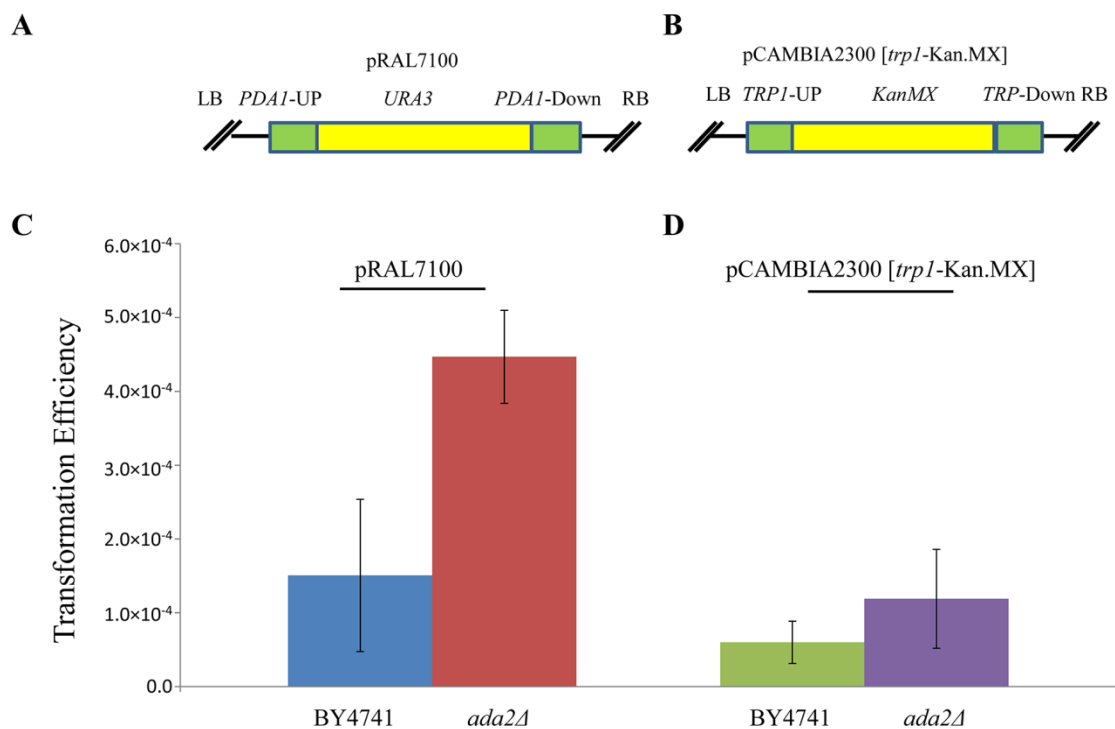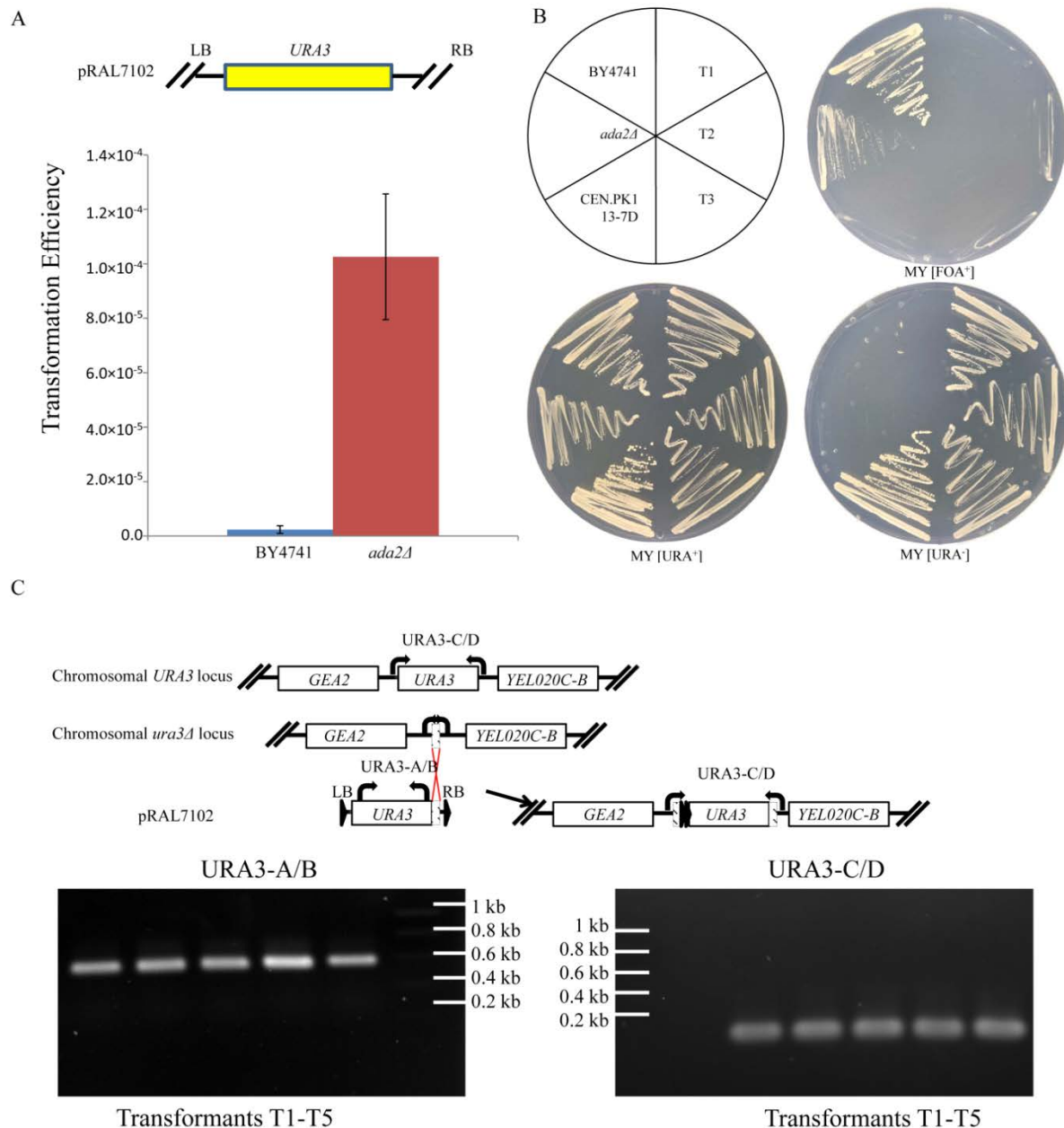
Figure 2. Increased *Agrobacterium*-mediated transformation efficiency of the yeast *ada2Δ* deletion mutant by NHEJ. Yeast strain BY4741*ada2Δ* and its parent strain BY4741 were co-cultivated with *Agrobacterium* strain LBA1100 harboring pRAL7102. (A) The structure of the T-DNA used and the effect of the *ADA2* deletion on AMT efficiency via NHEJ. Error bars indicate the standard deviations of triplicate assays. (B) The transformants generated by non-homologous T-DNA were used for further study. Growth on plates either with 5-FOA or without 5-FOA was used to detect whether the *URA3* marker was stably integrated in the genome or could be lost. The wildtype BY4741, CEN.PK113-7D and *ada2Δ* deletion mutant were included as controls. (C) The presence of the selective *URA3* marker was determined with the primer pair URA3-A/B via PCR. Due to the short homologous sequence of 47 bp found in T-DNA mentioned above, T-DNA may integrate at the *URA3* locus by HR through a single cross-over. The primer pair URA3-C/D was used to check for integration at the *URA3* locus. All DNA templates were isolated from individual transformants.

grow in the presence of 5-FOA, whereas transformants with the T-DNA on extrachromosomal structures are able to grow slowly on media with 5-FOA as extrachromosal DNA structures are usually unstable on non-selective media. In total, 16 transformants of BY4741 and 104 transformants of the *ada2Δ* deletion mutant from three independent experiments were analyzed. All analyzed transformants were able to grow on plates lacking uracil, but not on plates with 5-FOA indicating that the *URA3* selection marker of the T-DNA was stably maintained and expressed. Further growth analysis showed that one of six tested transformants of BY4741 did grow slowly on 5-FOA, but none of the eighteen selected transformants of the *ada2Δ* deletion mutant were able to grow in the presence of 5-FOA, indicating that these transformants cannot easily loose the T-DNA. The effect of 5-FOA on three selected representative transformants and on an uracil prototrophic (CEN.PK113-7D) and auxotrophic (BY4741) control strain is shown in Figure 2B.

The *ada2Δ* deletion mutant has a slow growth phenotype

It has been reported that *ada2* mutants have a decreased growth compared to the corresponding parental strains (Berger et al., 1992). As the increased transformation efficiency of the *ada2Δ* mutant may be related to the decreased growth rate of this mutant we compared the growth of the *ada2Δ* deletion mutant in BY4741 with that of the parental strain. As illustrated in Figure 3A, the *ada2Δ* deletion mutant indeed grew slower than BY4741 (growth rates: 0.184 ± 0.007 h^{-1} and 0.150 ± 0.011 h^{-1}; n= 4, P=0.0019 for *ada2Δ* and BY4741, respectively). Also the *ada2Δ* deletion mutant formed slightly smaller colonies than the BY4741 strain, both on rich (YPD) and minimal (MY) medium (Figure 3B). In another experiment we showed that also the *ada2Δ* deletion mutant harboring plasmid YEp24 grows significantly (P=0.048) slower than BY4741 carrying this plasmid (0.176 ± 0.027 h^{-1} *vs* 0.247 ± 0.034 h^{-1}, respectively) (see also Figure 8).
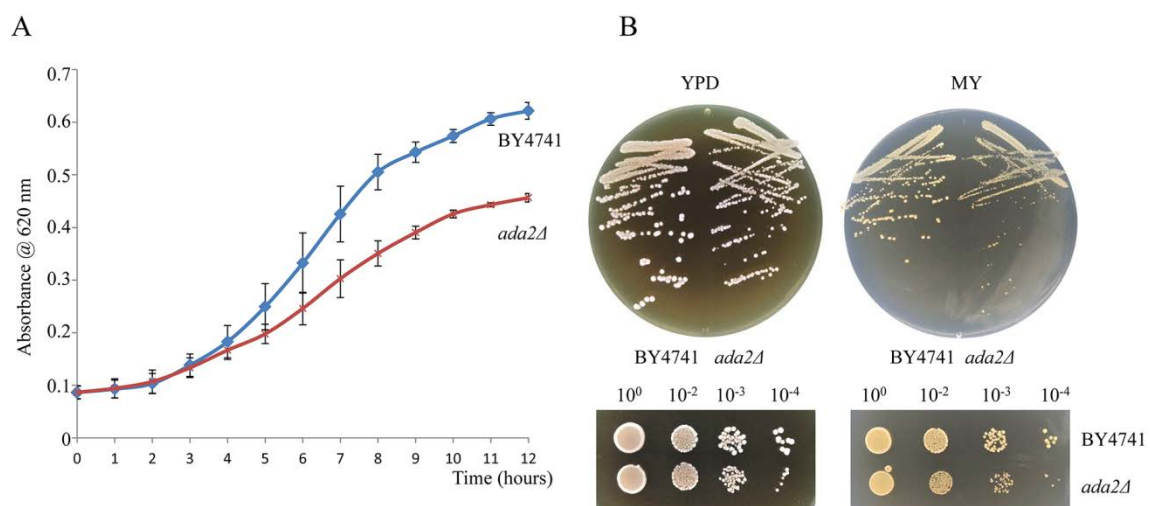
Figure 3. The *ada2Δ* deletion mutant grows more slowly than the wild type BY4741. (A) Analysis of the effect of the *ADA2* deletion on growth in liquid culture. Dilutions of yeast cells were inoculated into MY media after overnight culture in rich medium YPD and absorbance measurements were performed at the indicated time points. Error bars indicate the standard deviations of three independent replicates. (B) Analysis of the effect of the *ADA2* deletion on growth on plates. Both wild type BY4741 and *ada2Δ*

deletion mutant were tested on rich YPD and minimal MY plates. Yeast cells were serially diluted and spotted onto the plates. The photos were taken after three days and the experiment was performed three independent times. Results of representative plates are shown.

A

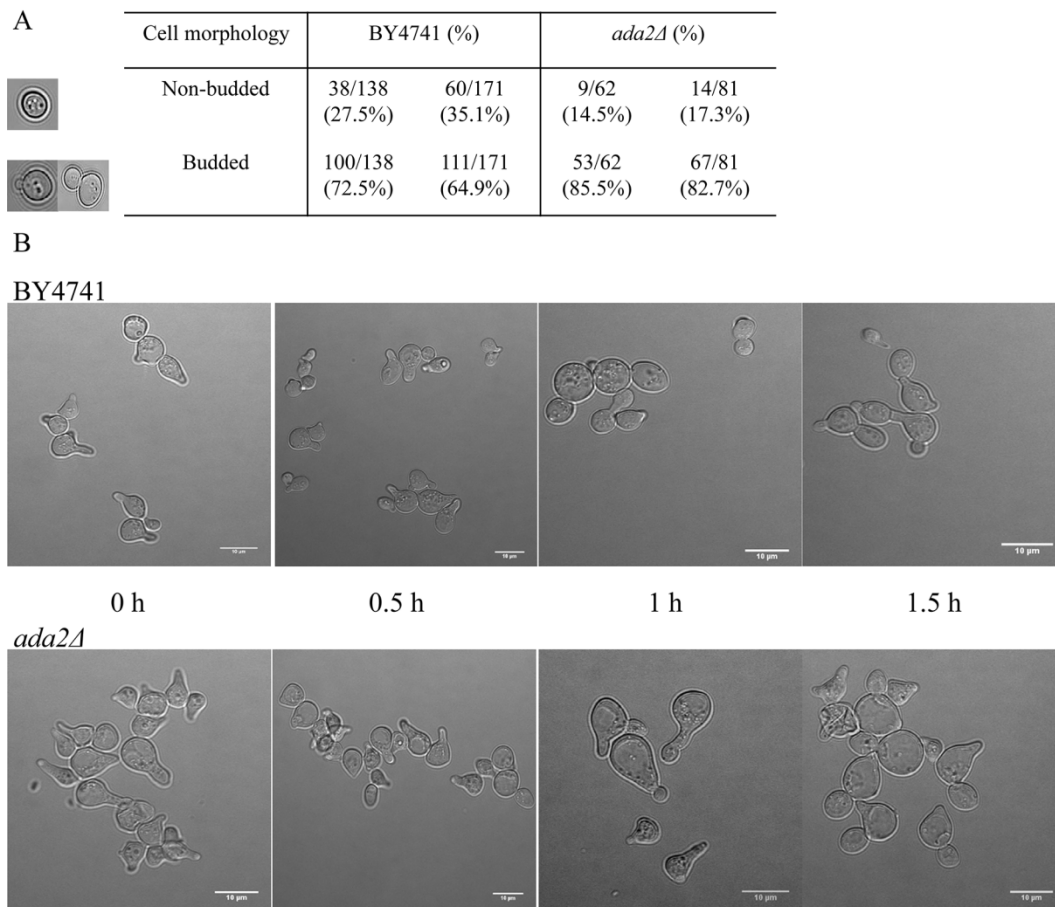| Cell morphology	BY4741 (%)		*ada2Δ* (%)	
Non-budded	38/138 (27.5%)	60/171 (35.1%)	9/62 (14.5%)	14/81 (17.3%)
Budded	100/138 (72.5%)	111/171 (64.9%)	53/62 (85.5%)	67/81 (82.7%)

B

BY4741

0 h   0.5 h   1 h   1.5 h

ada2Δ

Figure 4. The effect of the *ada2Δ* deletion on cell cycle progression. (A) Analysis of budding by microscopy. Yeast cells were cultured in minimal liguid medium until log phase. Then the images of cell populations were taken and the numbers of budding and non-budding yeast cells were manually counted. (B) Analysis of cell cycle progression after release from α-factor arrest. α-factor was used to arrest wild type and *ada2Δ* deletion mutant cells at G1/S phase. Cells were collected at the indicated times after release from α-factor arrest. The wildtype W303 *bar1Δ* recovered rapidly and continued to grow (upper panel). However, the *ada2Δ* deletion mutant presented significant growth deficiency (lower panel).

It is well known that Ada2 is a general transcriptional regulator and the expression of 2.5% of the genes is affected at least 2-fold by the *ADA2* deletion (Hoke et al., 2008). Combined with the lower growth of its deletion mutant, it is conceivable that some genes regulated by Ada2 can affect the cell cycle of yeast cells. To get more insight in the effect of the *ADA2* deletion on the cell cycle, the number of budded cells present in exponentially growing cultures of BY4741 and *ada2Δ* cells were counted. As shown in Figure 4A, 72.5% / 64.9% of the BY4741 cells and 85.5% / 82.7% of the *ada2Δ* cells were budded, suggesting a slight difference of their cell cycle. To further investigate whether the *ADA2* deletion affected a specific moment in the cell cycle, we made use of the *bar1Δ* deletion mutant in the W303 background allowing arrest at

the G1-S boundary by the α-factor and constructed an *ada2Δ* deletion in this strain. More than ninety percent of both wild type and *ada2Δ* deletion mutant cells have small buds when arrested by the α-factor (Figure 4B). In the first 1.5 hours after release from the α-factor arrest the bud size strongly increased in wild type cells, but not in *ada2Δ* cells (Figure 4). Only after 2 hours the bud size of these cells started to increase. These experiments may indicate that *ada2Δ* cells are delayed in the S-phase of the cell cycle. However, additional experiments have to be performed to show whether or not *ADA2* specifically regulates this phase of the cell cycle.
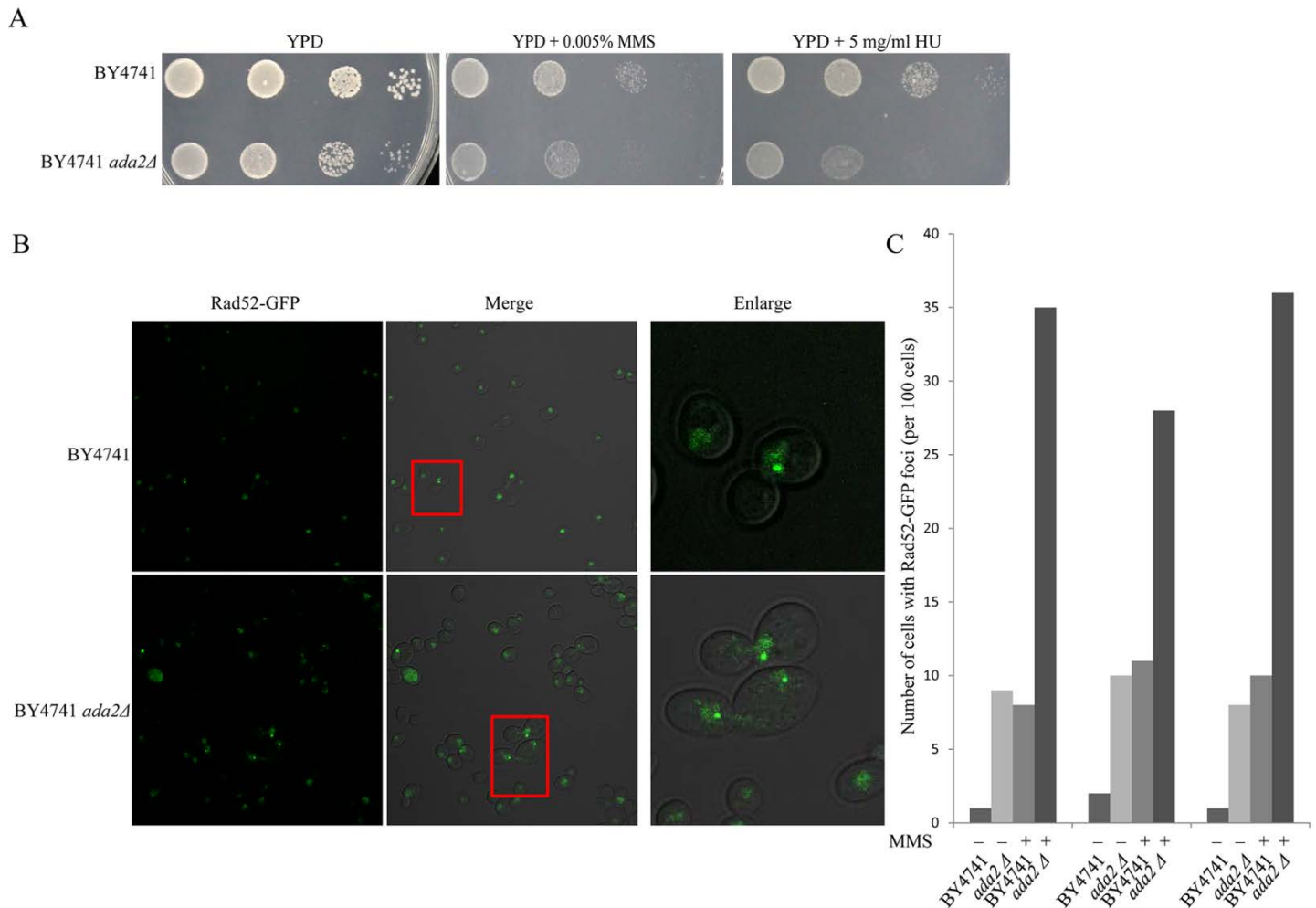
Figure 5. The *ada2Δ* deletion mutant is more sensitive to the DNA damaging agents MMS and HU and has an increased number of double strand breaks. (A) Both wild type BY4741 and *ada2Δ* deletion mutant were tested on YPD plates with two commonly used DNA damaging agents, MMS and HU. Yeast cells were serially diluted and spotted onto the plates. The photos were taken after three days and representative results of three independent experiments are shown. (B) The Rad52 protein was marked by GFP to visualize DSBs in yeast cells. The DNA repair foci were observed using confocal microscopy. (C) The number of DNA repair foci is increased in the *ada2Δ* deletion mutant both with or without the addition of MMS. The data from three independent experiments are shown.

The *ada2Δ* deletion mutant is more sensitive to DNA damaging agents

Ada2-dependent histone acetylation may be involved in double strand break (DSB) repair (Muñoz-Galván et al., 2013). As DSBs may promote AMT we investigated the sensitivity of the *ada2Δ* deletion mutant for DNA damaging agents. The DNA alkylating agent methyl methanesulfonate (MMS) has been used for many years to induce double-stranded breaks during replication and hydroxyurea (HU) is a potent inhibitor of the enzyme ribonucleotide reductase (RNR) in S-phase and leads to stalling of DNA replication. Survival viability was estimated by plating serial dilutions of cultures of wild type and *ada2Δ* deletion mutant cells on YPD plates containing MMS or HU. Then all plates were incubated at 30°C for 2 days. As demonstrated in Figure 5A, the deletion of *ADA2* enhanced the sensitivity to the DNA damaging agents MMS and HU. To obtain more direct evidence for the presence of chromosomal DNA damage accidents in *ada2Δ* mutants, we analyzed *Rad52* foci formation in the absence and presence of MMS. *Rad52* is a master regulator protein of DNA repair via homologous recombination and *Rad52* is recruited at DSBs which can be seen as foci when using GFP-tagged Rad52 (Lisby et al., 2001). Such foci were also formed in the *ada2Δ* deletion mutant (Figure 5B). As shown in Figure 5C, the number of *Rad52* foci in the *ada2Δ* mutant seems higher than that in the wild type. However, upon treatment with MMS, the number of *Rad52* foci is around three fold higher in the *ada2Δ* mutant compared to the wild type strain. In other words, in the *ada2Δ* mutant there is more DNA damage or repair is slower and this could lead to chromosomal DNA instability which may contribute to the integration of T-DNA.

Overexpression of *SFP1* can rescue the MMS sensitivity of the *ada2Δ* deletion mutant

To gain more insight into why yeast cells lacking *ADA2* are hypersensitive to DNA damaging agents, we performed a synthetic rescue screen (Figure 6A). To this end, a YEp24-based *S. cerevisiae* genomic DNA overexpression library (Carlson and Botstein, 1982) was transformed into the *ada2Δ* deletion mutant. First Ura$^+$ transformants were selected and subsequently all transformants were copied to plates supplemented with MMS to screen for MMS resistant transformants. From approximately 80,000 transformants, 32 of them could grow on MY plates containing 0.01 % MMS. Plasmids, isolated from candidate yeast transformants, were transformed into *E. coli* to be sequenced. Sequence analysis indicated that 6 distinct genetic loci were represented among the original yeast transformants (Figure 6B). Of these 32 candidates, most (23 candidates) contained the *ADA2* gene. To identify which gene on the other plasmids is responsible for the viability of the original transformants, each hypothetical gene was individually inserted into YEp24 and introduced back into the original recipient *ada2Δ* deletion mutant. As a control, the empty vector YEp24 was used. To investigate whether the effects of overexpression of the candidate genes are specific for the *ada2Δ* deletion mutant, also the wild type strain BY4741 was transformed with the empty vector YEp24 or with YEp24 carrying the candidate genes. After selection for resistance to MMS only two candidate genes (*HEM2* and *SFP1*) remained (Figure 6C). However, BY4741 cells carrying YEp24 with *HEM2* were also more resistant to MMS compared to the BY4741 cells carrying the empty YEp24 vector. In contrast, the effect of *SFP1* was specific for the *ada2Δ* deletion mutant (Figure 6C). Cells of the *ada2Δ* deletion mutant carrying YEp24 with *HEM2* or *SFP1* were not more resistant to HU compared to cells carrying the empty vector, indicating that the effects are specific for the
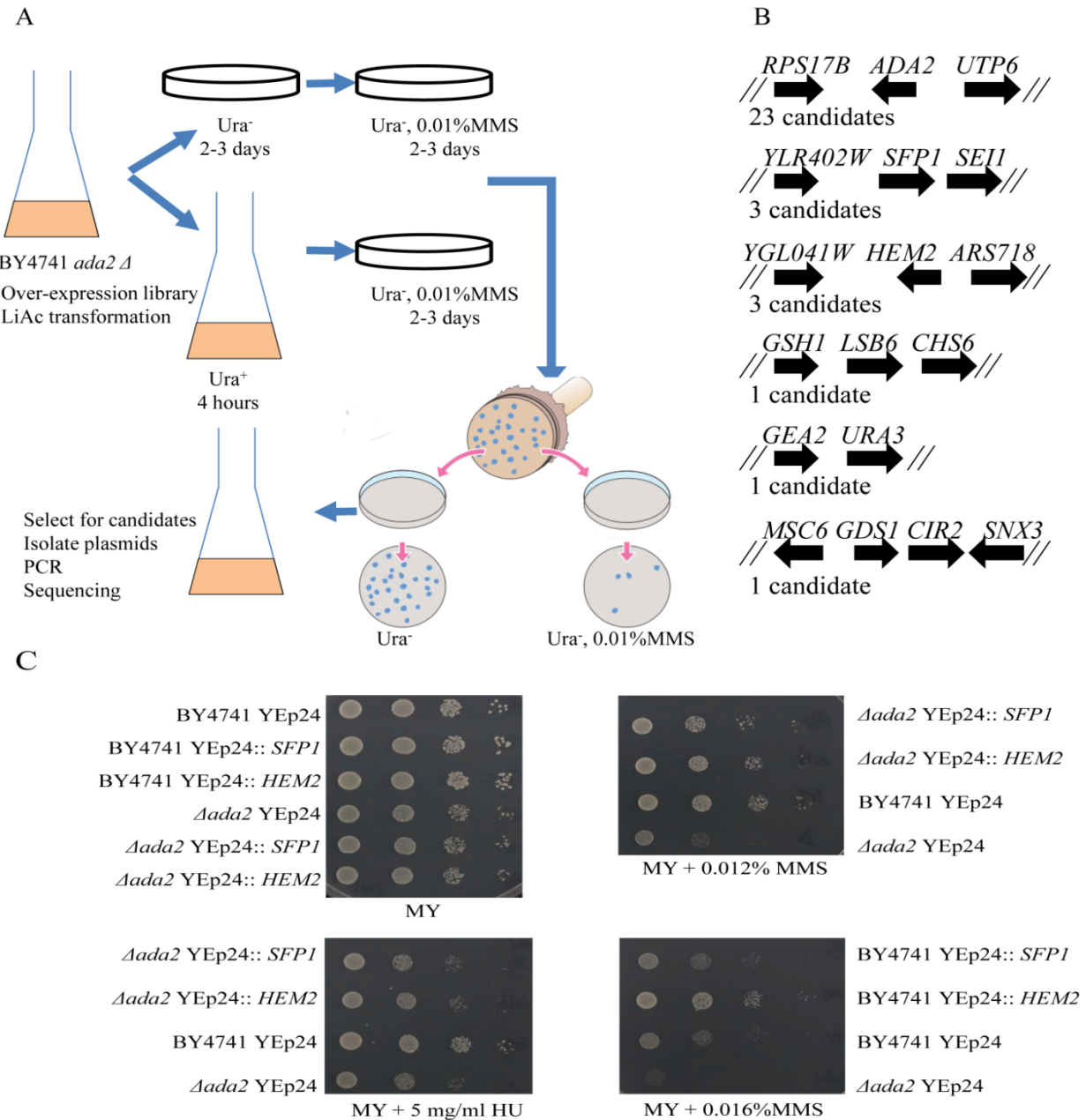
Figure 6. A genome-wide overexpression screen to identify genes that can suppress the MMS sensitivity of the *ada2Δ* deletion mutant. (A) schematic diagram of the synthetic rescue experiment. In brief, the whole genome-wide library was transformed into *ada2Δ* deletion mutant and plated on supplemented MY medium lacking uracil. Subsequently, all transformants were copied and spotted onto supplemented MY medium lacking uracil with MMS to test their sensitivity. (B) schematic representation of the yeast chromosomal regions contained in the YEp24 genomic DNA clones obtained by the synthetic rescue experiments. All candidates were from the combination of three independent experiments. After that, each single candidate including its promoter region was cloned into YEp24 vector and introduced in *ada2Δ* deletion mutant again to test for sensitivity to MMS. (C) MMS resistance in *ada2Δ* deletion mutant can be rescued by overexpression of *SFP1* and *HEM2*. The overexpression vector YEp24 containing *SFP1* or *HEM2* were introduced into *ada2Δ* deletion mutant and wild type BY4741. The empty plasmid YEp24 was introduced as control. All mutants were serially diluted and spotted onto MY medium with appropriate concentrations of MMS. After 3 days, the photos were taken and the spot assays were performed at least three independent times.

sensitivity to MMS. *HEM2* is an essential gene, encoding a homo-octameric enzyme involved in heme biosynthesis and *SFP1* is a non-essential gene which is involved in the activation of transcription of the genes encoding ribosomal proteins (Saccharomyces Genome Database, SGD, www.yeastgenome.org)(Marion et al., 2004). These experiments indicate that there is a genetic interaction between *SFP1* and *ADA2* in the response to the DNA damaging agent MMS.

Overexpression of *SFP1* affects the efficiency of AMT in the *ada2Δ* deletion mutant

To investigate whether overexpression of *SFP1* in the *ada2Δ* deletion mutant can affect AMT efficiency, we co-cultivated the *ada2Δ* deletion mutant and the parental strain BY4741 carrying YEp24[*SFP1*] or the empty vector with *Agrobacterium* strain LBA1100 (pRAL7100). As shown in Figure 7, the increased transformation efficiency of the *ada2Δ* deletion mutant is abolished after introduction of YEp24[*SFP1*]. Such an effect was not seen for the BY4741 control strain.

Both Ada2 and Gcn5 are components not only of the ADA, but also of the SAGA histone acetyltransferase complex (Balasubramanian et al., 2002; Sun et al., 2018). In addition, both of them are required for repair of DNA damage (Mckinney et al., 2013). To determine whether overexpression of *SFP1* also affects the *gcn5Δ* deletion mutant, we generated a *gcn5Δ* deletion mutant and introduced YEp24[*SFP1*] or the empty vector. As shown in Figure 7, there was no significant difference in AMT efficiency between the *gcn5Δ* deletion mutant harboring the overexpression vector with *SFP1* or the empty vector. No effect of the *gcn5Δ* deletion on AMT was found, which is in contrast with previous observations in the diploid BY4743 background (Soltani et al., 2009).
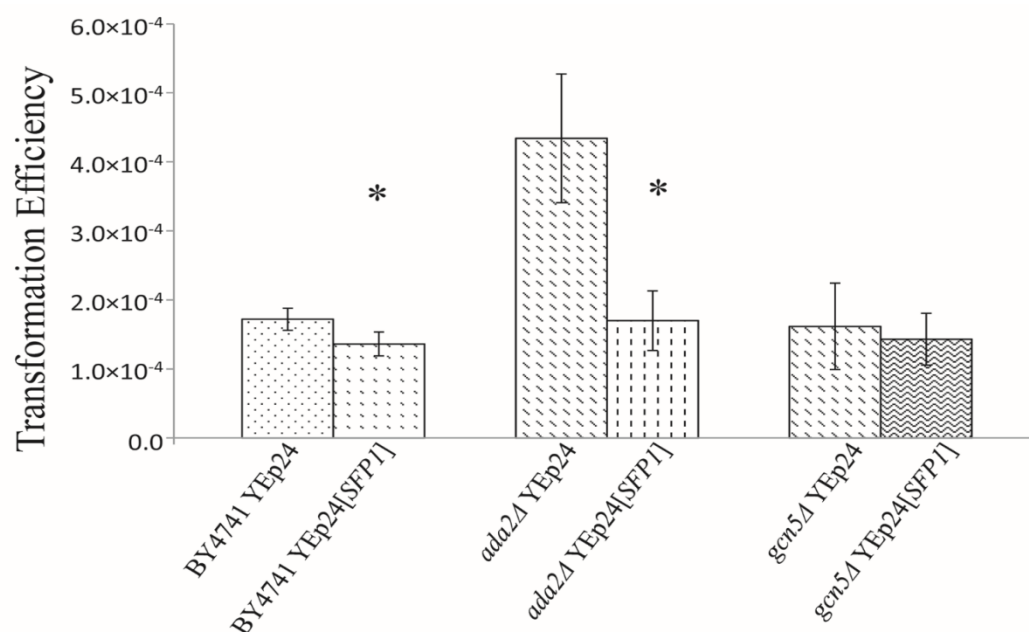
Figure 7. Effect of overexpression of *SFP1* on AMT efficiency. Yeast strains BY4741, BY4741*ada2Δ* and BY4741*gcn5Δ* containing YEp24 or YEp24[SFP1] were co-cultivated with *Agrobacterium* strain LBA1100 harboring pSDM8001 and AMT efficiency was determined. Error bars indicate the standard deviations of triplicate independent assays. Differences were statistically significant comparing overexpression of *SFP1* with its corresponding control by Student's test. *: P < 0.05.

The effect of overexpression of *SFP1* on *ada2Δ* is probably not related to the role of *SFP1* in regulation of the TORC1 complex nor to an effect on the growth rate

The proposed dominant cellular role of *SFP1* is regulation of transcription of ribosomal protein and biogenesis genes (Jorgensen et al., 2004). It has been shown that Sfp1 interacts directly with the Target of Rapamycin Complex 1 (TORC1) in a rapamycin-regulated manner, and that Sfp1 is phosphorylated by this kinase complex. On the other hand, Sfp1 negatively regulates TORC1 suggesting feedback mechanisms controlling the activity of these proteins in response to different nutrient and stress conditions (Lempiainen et al., 2009). To investigate a possible role of *ADA2* in the TORC pathway 27 genes involved in the this pathway or involved in regulation of TORC1 signaling were selected (Table S2) and overexpression vectors harboring these genes were obtained from the Yeast Genomic Tiling Collection. These vectors were introduced into the *ada2Δ* deletion mutant and the sensitivity to MMS was tested. After the first screen (supplementary Figure S1A) 10 potential candidates (*MRS6, LST8, LST7, TCO89, SEA4, VAM6, GTR1, GTR2, LST4* and *TOR1*) were further tested. As each vector from this collection contains more than one gene, candidate genes were amplified from yeast genomic DNA, cloned into the multi-copy vector pRS425 and introduced into the *ada2Δ* deletion mutant. As controls, both the vector pRS425 containing *SFP1* and the empty vector were introduced into the wild type strain BY4741. As demonstrated in Figure S1B, unlike *SFP1*, none of these genes on a multi-copy plasmid could rescue the sensitivity towards MMS of the *ada2Δ* deletion mutant. Thus, these observations didn't provide evidence for involvement of the TORC pathway in sensitivity of the *ada2Δ* deletion mutant towards MMS, but did not exclude such a role.

More recently, Ohmine et al. showed that deletion of *ERG28* resulted in a decreased AMT efficiency and that growth of the *erg28Δ* strain was unaffected by the presence of donor bacterial cells, while the growth of the wild type and other mutant yeast strains was suppressed by their presence (Ohmine et al., 2016). This may imply that the growth rate affects AMT efficiency. It has been reported that *SFP1* may also be involved in the regulation of cell division and the cell cycle (Jorgensen et al., 2002; Albert et al., 2019). Hence, we determined the effect of overexpression of *SFP1* on the growth rate of wild type, *ada2Δ*, *gcn5Δ* and *ada2Δgcn5Δ* strains cultured in liquid medium. As shown in Figure 8, the growth rate of the tested deletion mutants was lower compared with that of the wild type strain. However, their growth rates were not restored by overexpression of *SFP1*, suggesting that the effect of overexpression of *SFP1* on AMT efficiency is not related to an effect on the growth rate.
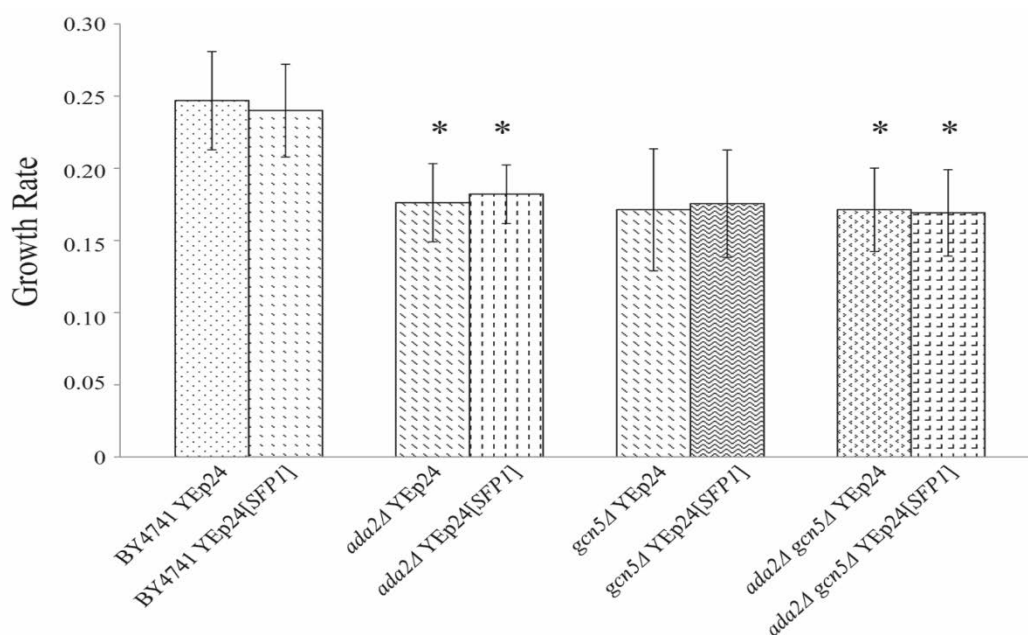
Figure 8. Effect of overexpression of *SFP1* on the growth rates of different yeast strains. The yeast cells were cultured overnight, then diluted to a final OD_{620} of 0.1 and inoculated into fresh medium on the next morning. The OD_{620} was systematically measured every hour until a total of 8 time points. Values were normalized and three growth curves were obtained for each mutant. Analysis of the curves by fitting to an exponential trend line yielded the growth rate. Afterwards, average growth rate for each mutant was calculated together with their SD (Standard deviation). Error bars indicate the standard deviations of triplicate assays and each contains three independent replicates. Differences were statistically significant compared to strain BY4741 YEp24 by Student's test. *: $P < 0.05$.

Figure 9. Visualization of the subcellular localization of GFP-Sfp1 in BY4741 and *ada2Δ* deletion mutant cells. BY4741 and BY4741*ada2Δ* were transformed with plasmid pUG34[SFP1] encoding the GFP-Sfp1 fusion protein. Yeast cells were cultured overnight and then analyzed by confocal microscopy. The experiment was repeated three independent times and similar results were obtained.

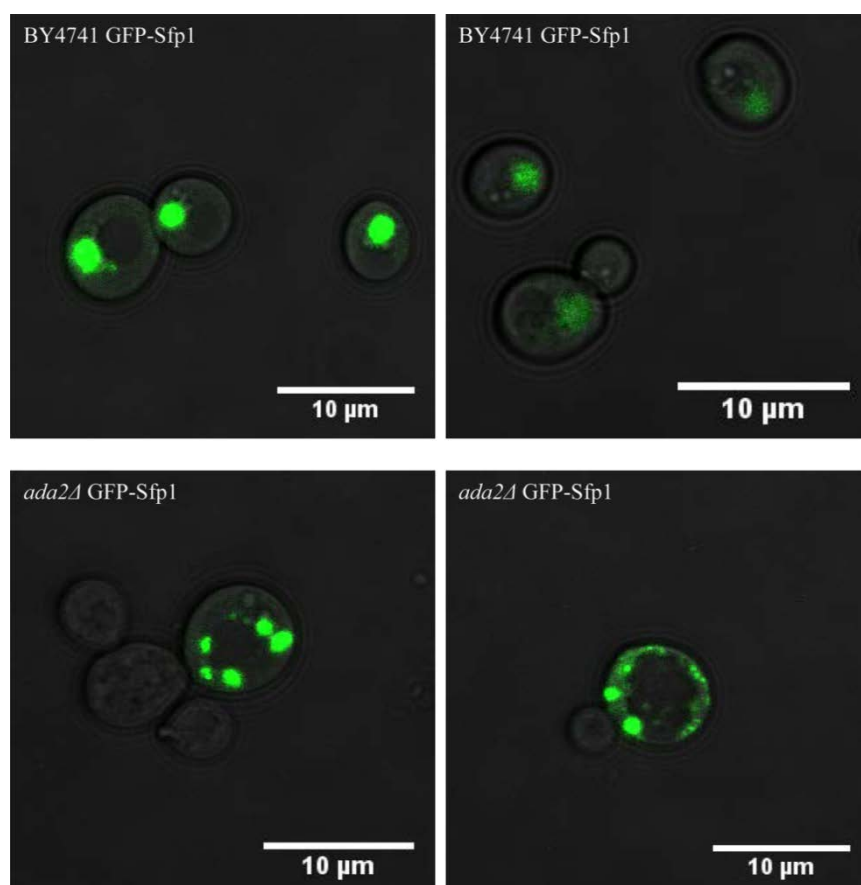Sfp1 is regulated, at least in part, by its subcellular localization (Lempiainen et al., 2009). In optimally growing cells, Sfp1 is predominantly nuclear and localizes to a number of Ribosomal Protein gene promoters. By contrast, unfavorable nutrient conditions or stress lead to a rapid re-localization of Sfp1 from the nucleus to the cytoplasm. Therefore, we investigated the effect of the *ada2Δ* deletion on the localization of Sfp1. To this end, we used the centromeric plasmid pUG34 to express a GFP-Sfp1 fusion protein. After 24 hours of culture, GFP-Sfp1 is still concentrated in the nucleus in the wild type strain BY4741, while in the *ada2Δ* deletion mutant most of GFP-Sfp1 was present in multiple unknown structures (Figure 9). This observation suggests that deletion of *ADA2* makes yeast cells more sensitive to environmental stress or to other unfavorable factors. In other words, in the absence of Ada2, Sfp1 failed to localize in the nucleus, indicative of stress caused by persistent DNA damage in the *ada2Δ* deletion mutant.

Visualization of T-DNA in *ada2Δ* and BY4741 cells during co-cultivation with *Agrobacterium*

To compare the fate of the T-DNA translocated from *Agrobacterium* into *ada2Δ* cells with that of T-DNA translocated into BY4741 cells, we attempted to visualize the T-DNA during AMT. To this end a series of *Agrobacterium* binary vectors harboring *lacO* repeats were constructed. The plasmid pCAMBIA2300[*trp1-lacO*] contains an array of 256 *lacO* repeats on its T-DNA flanked with sequences of the yeast *TRP1* locus. Upon translocation of this T-DNA from *Agrobacterium* into yeast cells expressing the Lac repressor tagged with GFP, the T-DNA may become visible (Figure 10A). This system has been used successfully before for example to visualize chromosome duplication (Straight et al., 1996).

To order to test this system, we introduced the 256 tandem *lacO* repeats in the high-copy vector pRS425 and then transferred the plasmid into the yeast strain expressing the fusion protein LacI-GFP. As can be seen in the Figure 10B, there are multiple dots representing the *lacO* array in the nuclei of this strain. Upon co-cultivation of BY4741 or *ada2Δ* expressing LacI-GFP with *Agrobacterium* carrying pCAMBIA2300[*trp1-lacO*] only a few cells containing fluorescent structures could be detected. AMT efficiency of *ada2Δ* was around 1.0×10^{-4} and that of BY4741 was half that; therefore, it is not surprising that the number of cells with fluorescent signal was low even we observed more than 10,000 cells in every independent experiment. As shown in Figure 10C a single fluorescent dot was found in one of the cells of the wild type strain and two dots were found in one of the *ada2Δ* cells expressing LacI-GFP. Unfortunately, only a low number of fluorescent cells were found preventing statistical analysis.

In order to monitor unintegrated T-DNA, a new T-DNA construct was generated with the KanMX selection marker flanked by sequences of the *PDA1* gene and with the 256 tandem repeats of the *lacO* located outside the homologous flanking sequences (pSDM8001-*lacO*; Figure 10D). Upon co-cultivation of *Agrobacterium* carrying this vector with BY4741 or *ada2Δ* expressing LacI-GFP again only very few cells with fluorescent structures were seen. In one of the *ada2Δ* cells a fluorescent structure residing in the nucleus and in another cell in addition a single spot was observed (Figure 10D). We did not observe fluorescent structures in the BY4741 strain after checking a similar number of cells as for the *ada2Δ* mutant.

Figure 10. 256 *lacO*-LacI assay to visualize the fate of T-DNA after it has been delivered into recipient yeast cells.
(A) Schematic of the 256 *lacO*-LacI assay. pRS305-LacI-GFP was integrated into the *LEU2* locus of yeast by homologous recombination so that LacI-GFP fusions can be expressed in the nucleus continuously. The T-DNA construct carries the 256 *lacO* repeat array and once it enters into the nucleus and has become double stranded, the *lacO* repeat array can recruit GFP–LacI fusions to form foci which can be monitored.
(B) The high copy number plasmid pRS425 harboring the 256 *lacO* repeat array was introduced into the recipient yeast cells to test the stability of the 256 *lacO*-LacI assay. DAPI was used to visualize the nucleus of yeast cells. (C) The structure of T-DNA of plasmid pCAMBIA2300[trp1-lacO] is presented. *Agrobacterium* harboring this T-DNA was co-cultivated with wild type and *ada2Δ* deletion mutant cells expressing LacI-GFP. After 3 days, the mixture of *Agrobacterium* and yeast cells were washed three times with MilliQ water and were visualized using confocal microscopy. (D) A T-DNA construct harboring 256 *lacO* repeat array outside the flanking homologous sequences was used to visualize the fate of T-DNA which was not integrated into the desired locus.
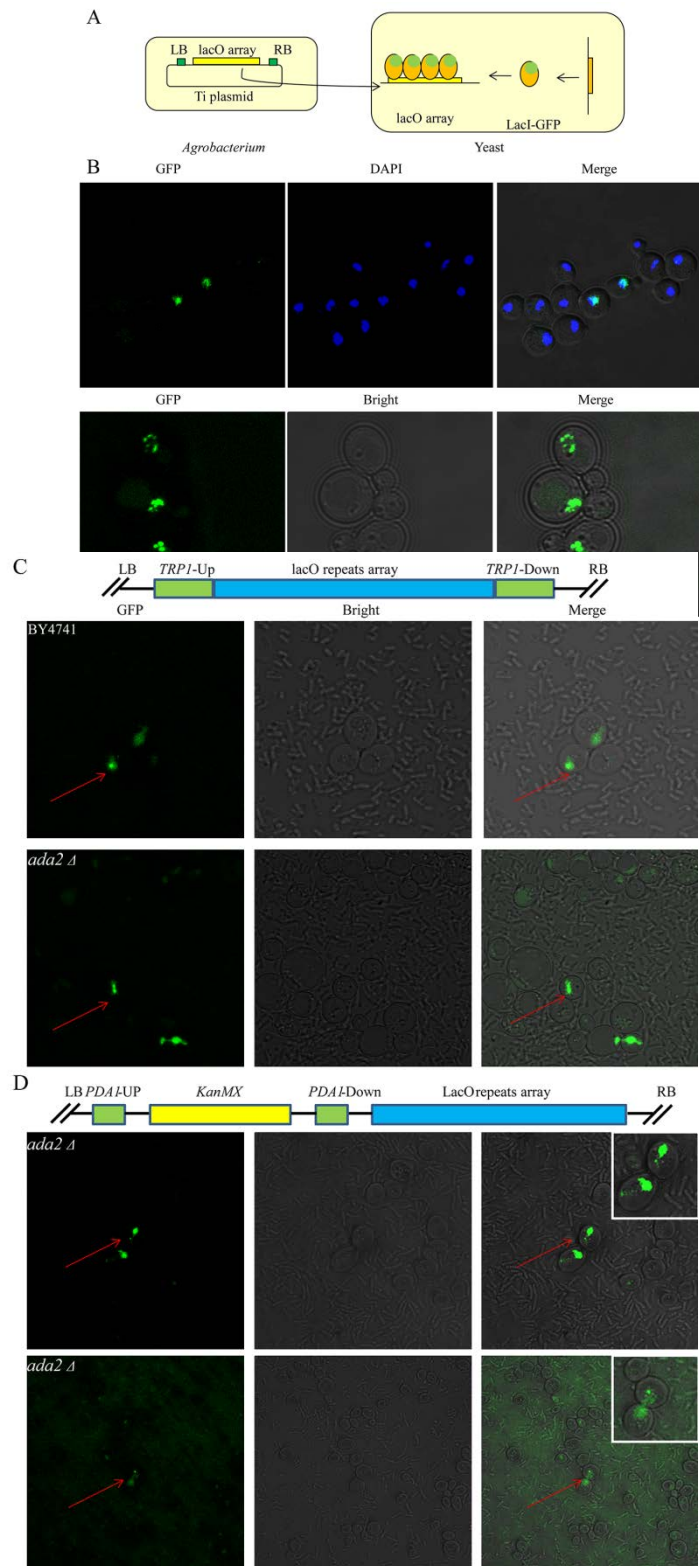
Discussion

The bacterial side of AMT has been investigated rather well but about the host factors involved in this process less is known. The yeast *S. cerevisiae* has emerged as an excellent model host to

study the mechanism of AMT (Bundock et al, 1995; van Attikum et al., 2001, 2003; Zhang et al., 2017). Several large scale genome-wide screens have been used to find yeast host genes involved in this process (Soltani et al., 2009; Ohmine et al., 2016; Hooykaas et al., 2018). Previous research in our group revealed that in the *ada2Δ* deletion mutant the efficiency of AMT with a vector integrating via HR is higher than that in the wild type. In this chapter, we exploited the non-homologous T-DNA vector pRAL7102 to investigate AMT with a vector integrating by NHEJ in the *ada2Δ* deletion mutant. As shown in Figure 2 the AMT efficiency in the *ada2Δ* deletion mutant is highly increased not only when the T-DNA is integrated by HR but also when it is integrated by NHEJ.

It is unknown why deletion of *ADA2* results in an increased AMT efficiency. In the commonly used experimental procedures to check the efficiency of AMT, the initial ratio of donor *Agrobacterium* cells and recipient yeast cells is important for a high AMT efficiency. Ohmine et al., 2016 reported an observation that when recipient wild type yeast cells were overloaded with *Agrobacterium* or when a yeast mutant was used which could continue to grow in the presence of *Agrobacterium* to produce more cells than the wild type during the co-cultivation, the final AMT efficiency was attenuated. As the *ada2Δ* deletion mutant exhibits a slower growth and delayed cell cycle compared to the wild type, the ratio of *Agrobacterium* to yeast may be suboptimal. In our experiments, the initial amount of *ada2Δ* deletion mutant cells was almost identical with that of wild type cells. After co-cultivation the final output number of *ada2Δ* cells was lower (around half of that of wild type), thus contributing to an increased AMT efficiency.

An alternative way to explain the enhanced AMT efficiency may be that the observed effect is the result of the expression level of the selection marker. It has been reported that there may be a significant difference between the number of transformants obtained after selection and obtained without selection (Shilo et al., 2017). T-DNA integration does not necessarily lead to expression of the selection marker and it is possible that observed differences in AMT efficiency are caused by differences in the expression of the selection marker. In the AMT protocol used in this chapter stable expression of the selection marker was used to determine the efficiency of T-DNA integration. Thus, the increased AMT efficiency may be caused by an altered expression of the selection gene in the *ada2Δ* mutant when the T-DNA is inserted into a region of the chromosome that is silenced in the wild type and not in the *ada2Δ* mutant. In line with this hypothesis, the absence of *ADA2* was reported to facilitate the spread of heterochromatic proteins and to cause de-repression of a normally silenced telomere gene (Jacobson and Pillus, 2009). Moreover, T-DNA was reported to preferentially integrate at (sub)telomere regions in *rad50*, *mre11* and *xrs2* mutants (van Attikum et al., 2001). On the other hand, we have no evidence that silencing of the selection markers indeed plays a role in the AMT experiments shown in this chapter.

Simultaneously, the affected AMT efficiency could be the consequence of altered expression of certain genes directly or indirectly regulated by *ADA2*. The expression of approximately 2.5% of all yeast genes was found to be affected at least 2-fold by the *ada2Δ* deletion (Hoke et al., 2008). It can be speculated that the increased AMT efficiency of the *ada2Δ* deletion mutant is the consequence an indirect effect of the altered regulation of certain genes involved in the transformation process. Taking together, the increased AMT efficiency of *ada2Δ* deletion mutant can also be caused by an altered expression of genes other than *ADA2*.

It is well known that chromatin modifications play an crucial role in DNA repair mechanisms which are exploited to facilitate T-DNA integration. Several observations were described and reviewed (Magori and Citovsky, 2011), indicating that the histone acetylation balance is important for T-DNA integration even though its molecular basis remains unclear. *ADA2* is a component of histone acetyltransferase (HAT) complexes related to chromatin modifications and T-DNA integration was reported to be accompanied with chromosomal instability in *rad50, mre11, xrs2* and *lig4* mutants (van Attikum et al., 2001). Hence, another plausible hypothesis could be that genomic DNA surrounding DSBs would be more "open" and accessible thus serving as "hot spots" to be recognized for T-DNA integration during AMT and the increased AMT efficiency could be due to more emergent DNA damage events in the *ada2Δ* deletion mutant.

Deletion of *ADA2* renders yeast cells sensitive to various DNA damaging agents such as the drug methyl methanesulfonate (MMS) and the replication inhibitor hydroxyurea (HU). In order to gain more insight in the role of *ADA2* in the response to DNA damaging agents, we performed a synthetic rescue screen by using a YEp24-based genomic overexpression *S. cerevisiae* library. In this way, we identified *SFP1* and *HEM2* to rescue the *ada2Δ* deletion mutant in presence of MMS. However, the effect of overexpression of *HEM2* was not specific for the *ada2Δ* mutant as it has a similar effect on the wild type strain BY4741. Moreover *HEM2* was reported to be a potential target of MMS (Lum et al, 2004); therefore, we focused on the interaction between *SFP1* and *ADA2*. A physical interaction between Sfp1 and Ada2 has not yet been reported. Likewise, we were not able to show a physical interaction between Ada2 and Sfp1 in an yeast two-hybrid analysis in our lab. In addition, *sfp1Δ* deletion mutant grows very slowly and unusually small colonies were obtained (Jorgensen et al., 2002). We were unable to construct an *ada2Δsfp1Δ* double deletion mutant and this phenotype may be caused by synthetic lethality of *ada2Δsfp1Δ* double deletion, although this synthetic lethality has not been reported before (SGD, www.yeastgenome.org).

SFP1 overexpression was reported to influence the expression of over 30% of RNAPII-transcribed genes and Sfp1 can bind at G1/S gene promoters as a negative regulator (Albert et al., 2019). Similarly, Ada2 is a subunit of a transcription activator complex and thought to be involved in activation of RNA polymerase II transcription. In our studies, the *ada2Δ* deletion mutant grows slower and has a smaller cell size than the corresponding wild type (Figure 3) and the cell cycle of *ada2Δ* deletion mutant may be impaired (Figure 4). Considering the overlapping effects caused by manipulating *SFP1* and *ADA2* expression and the genetic interaction between these genes found in this study, it is very likely they participate in the same process. However, the nature of this process is so far unknown.

Recently, *SFP1* was identified to be involved in the regulation of Tra1 expression in response to polyQ proteotoxicity (Jiang et al., 2019). An interaction between *SFP1* and *TRA1* has been reported before and *SFP1* was postulated to be a negative and indirect regulator of *TRA1*, potentially through its regulation of the TORC1 complex (Lempiainen et al., 2009). Moreover, the expression of *TRA1* was upregulated by polyQ expansions in the absence of *ADA2*. Like Ada2, Tra1 is a key component of the SAGA histone acetyltransferase complex and Tra1 physically interacts with Ada2 (Saleh et al., 1998) and a synthetic lethal interaction between *ADA2* and *TRA1* has been found (Berg et al., 2018). In our study, overexpressed *TRA1* could not rescue the growth in the *ada2Δ* background in the presence of MMS (Figure S1). The

results suggest that the interaction between *SFP*1 and *ADA2* has no relevance to the role of *SFP1* in its regulation of TORC1 complex. It is consistent with the growth deficiency of *ada2Δ* deletion mutant no matter whether *SFP1* is overexpressed or not.

There is more evidence supporting a role of Sfp1 in the DSB repair pathway. For instance, *SFP1* was identified in a screen using diploid yeast deletion strain libraries for sensitivity to various DSB agents, such as MMS, bleomycin and *EcoR*I (McKinney et al., 2013). However, which role *SFP1* may have in DSB repair has not been elucidated. Meanwhile, in a proteome-wide study, a physical interaction has been reported between Sfp1 and Asf1 (Krogan et al., 2006). Asf1 is a nucleosome assembly factor, involved in chromatin assembly and disassembly and required for recovery after DSB repair, although the functional significance of the interaction with Sfp1 has not been established. In this chapter, we demonstrated that the *ada2Δ* mutant is sensitive to DNA damaging agents such as MMS and HU and that overexpression of *SFP1* in the *ada2Δ* background specifically suppressed MMS sensitivity. These findings define a genetic interaction between *SFP1* and *ADA2* in the DNA damage response which is separate from the function of *SFP1* in the regulation of ribosomal protein and biogenesis genes.

Besides such indirect roles of *ADA2* in gene expression or chromatin structure, we need also to consider that there is more DNA damage in the *ada2Δ* deletion mutant as revealed by an increased number of *RAD52* foci. Due to the absence of *ADA2*, there are either more DNA damaging events in the cell or this damage is repaired less efficiently, thus providing more available sites for T-DNA integration. The increased DNA damage events could be expected to occur at random chromosomal loci and a small portion of them may be at or near the desired locus for T-DNA HR integration. Correspondingly, as described above the AMT efficiency by NHEJ increased 50 fold, while that by HR was only 2 fold higher in the absence of *ADA2* compared to the wildtype strain (Table S1B). Overexpression of *SFP1* in the *ada2Δ* deletion background not only reduces MMS sensitivity and the occurrence of DNA damage (the number of *RAD52* foci is less than 5 per 100 cells in the absence of MMS), but also leads to a reduction in AMT to wildtype levels. Thus, the increased AMT efficiency of the *ada2Δ* mutants together with the genetic interaction between *ADA2* and *SFP1* in response to DNA damaging agents suggest an important role of DNA damage events in T-DNA integration in the *ada2Δ* mutant.

Acknowledgements

Table 1. Yeast strains used in this study.

Yeast strain	Genotype	Source/Reference
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Brachmann et al., 1998
BY4741 *ada2Δ*	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ada2*::HphMX4	This study
BY4741 *gcn5Δ*	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcn5*::HphMX4	This study
BY4741-Rad52-GFP	BY4741 Rad52-GFP (KanMX4)	This study
BY4741 *ada2Δ*-Rad52-GFP	BY4741 *ada2*::HphMX4 Rad52-GFP	This study
BY4741 YEp24	BY4741 YEp24	This study
BY4741 YEp24::SFP1	BY4741 YEp24::SFP1	This study
BY4741 *ada2Δ* YEp24	BY4741 *ada2*::HphMX4 YEp24	This study
BY4741 *ada2Δ* YEp24::SFP1	BY4741 *ada2*::HphMX4 YEp24::SFP1	This study
BY4741 *gcn5Δ* YEp24	BY4741 *gcn5*::HphMX4	This study
BY4741 *gcn5Δ* YEp24::SFP1	BY4741 *gcn5*::HphMX4 YEp24::SFP1	This study
BY4741-LacI-GFP	BY4741 *leu2*::LacI-GFP	This study
BY4741 *ada2Δ*-LacI-GFP	BY4741 *ada2*::HphMX4 *leu2*::LacI-GFP	This study
BY4741 pUG34-SFP1	BY4741 pUG34::SFP1	This study
BY4741 *ada2Δ* pUG34-SFP1	BY4741 *ada2*::HphMX4 pUG34::SFP1	This study
W303a	MATa *leu2-3,112 trp1-1 can1-100 ura3-1 ade2-1 his3-11,15*	Thomas and Rothstein, 1989
W303a-lacO-LacI	W303a *ura3-1*::LacI-GFP pRS425::*lacO*256	This study
BY4741-lacO-LacI	BY4741 *ura3Δ0*::LacI-GFP pRS425::*lacO*256	This study
W303a *bar1Δ*	W303a *bar1*:: KanMX4	Zhang et al., 2017
W303a *bar1Δ ada2Δ*	W303a *bar1*:: KanMX4 *ada2*::HphMX4	This study

Table 2. Plasmids used in this study.

Plasmids	Specifications	Source/Reference
pAG32	contains the *hph* gene encoding hygromycin B phosphotransferase	Goldstein and McCusker, 1999
pUG6	contains the kanMX gene encoding kanamycin cassette	Güldener et al, 1996
pYM27	PCR template for C-terminal EGFP tagging	Janke et al., 2004
YEp24	Yeast episomal cloning vector with a *URA3* marker	Carlson and Botstein,1982
YEp24-SFP1	YEp24 with the coding sequence of *SFP1*	This study
pAFS59	Yeast integrative vector harboring 256 *lacO* repeats	Straight et al., 1996
pAFS152	Yeast integrative vector harboring LacI-GFP fusions	Straight et al., 1996
pRS305	Yeast integrative vector with a *LEU2* marker	Sikorski and Hieter, 1989
pRS305-LacI-GFP	pRS305 with LacI-GFP fusion protein	This study
pRS425	Yeast episomal expression vector with a *LEU2* marker	Sikorski and Hieter, 1989
pRS425[*lacO*]	pRS425 with *lacO* repeat array	This study

pRS425[SFP1]	pRS425 with the coding sequence of *SFP1*	This study
pCAMBIA2300	*Agrobacterium* binary vector for transformation, Km	Hajdukiewicz et al., 1994
pCAMBIA2300[*trp1*]		
pCAMBIA2300-*lacO*	pCAMBIA2300 with *lacO* repeat array	This study
pCAMBIA2300[*trp1-lacO*]	pCAMBIA2300 with *lacO* repeat array and *TRP1* flanking sequence	This study
pCAMBIA2300[*trp1*-kanMX]	pCAMBIA2300 with KanMX marker and *TRP1* flanking sequence	This study
pRAL7100	*Agrobacterium* binary vector with *URA3* selectable marker and PDA1 flanking sequence	Bundock et al., 1995
pRAL7102	*Agrobacterium* binary vector with *URA3* selectable marker	Bundock and Hooykaas, 1996
pSDM8001	*Agrobacterium* binary vector with KanMX selectable marker and *PDA1* flanking sequence	van Attikum and Hooykaas, 2003
pSDM8001-*lacO*	pSDM8001 containing *lacO* repeat array between PDA1 and right border	This study
pUG34	Centromeric plasmid with a *HIS3* marker for N-terminal GFP fusions under control of the *MET25* promoter	Guldener and Hegemann, unpublished
pUG34-*SFP1*	pUG34 with the coding sequence of *SFP1*	This study
Yeast Genomic Tiling Collection	complete overlapping collection of entire *S. cerevisiae* genome, screen for overexpression phenotypes	Dharmacon

Table 3. Primers used in this study.

primer name	sequence (5'-3')
Ada2-Fw	TAAAATATCAGCGTAGTCTGAAAATATATACATTAAGCAAAAAGACAGCTGAAGCTTCGTACGC
Ada2-Rev	ATAATAACTAGTGACAATTGTAGTTACTTTTCAATTTTTTTTTTGCCGCGGCCGCATAGGCCAC
Ada2-Ctrol-Fw	ACGACCTCTGAGAAAACGA
Ada2-Ctrol-Rev	GGTCCCTTTATGACTTGGC
Ada2-gfp-Fw	GCGAATAGAATATACGATTTTTTCCAGAGCCAGAATTGGATGCGTACGCTGCAGGTCGAC
Ada2-gfp-Rev	TAGTGACAATTGTAGTTACTTTTCAATTTTTTTTTTTGTTAACTATAGGGAGACCGGCAGATCCGC
Rad52-gfp-Fw	AGAGAAGTTGGAAGACCAAAGATCAATCCCCTGCATGCACGCAAGCCTACTCGTACGCTGCAGGTC
Rad52-gfp-Rev	AGTAATAAATAATGATGCAAATTTTTTATTTGTTTCGGCCAGGAAGCGTTTCAATCGATGAATTCGAG
Trp1-up-Fw	AA<u>GAGCTC</u>AAAACGGAAGAGGAGTAGGG
Trp1-up-Rev	AA<u>GGATCC</u>TGCTTAATCACGTATACTCACG
Trp1-down-Fw	AA<u>CTGCAG</u>AGGCCTTTTGAAAAGCAAGCA
Trp1-down-Rev	AA<u>AAGCTT</u>TACTTGACGACTTGAGGCTG
LacI-Fw	A<u>CTAGT</u>GTTTTCCCAGTCACGACGTT
LacI-Rev	<u>GTCGAC</u>ACGCCAAGCTCGGAATTAAC
SFP1-Fw/YEp24	CTA<u>GCTAGC</u>TTACCTCCCAGGCCCTATCT
LSB6-Fw	CTA<u>GCTAGC</u>AAGGTTTCCAGCCACAGTTG
GDS1-Fw	CTA<u>GCTAGC</u>AACCCATGCGATGAGAAGTC
CIR2-Fw	CATG<u>CATGC</u>TCCGACATCAATTCCAGTCA

HEM2-Fw	CAT<u>GCATGC</u>AGAAATGACTGCCTGGTGCT
SNX3-Fw	CAT<u>GCATGC</u>TCCTGTACCATCCGGAACTC
SFP1-Rev/YEp24	ACGC<u>GTCGAC</u>AGAAAAACGATCCGACAACG
LSB6-Rev	ACGC<u>GTCGAC</u>GGGACCTAGTCCAACATTGC
GDS1-Rev	ACGC<u>GTCGAC</u>TCCCAGAGTGGATTTTCCTG
CIR2-Rev	ACGC<u>GTCGAC</u>TGCTGTCAAAAGGACATGGA
HEM2-Rev	ACGC<u>GTCGAC</u>ATCTTGGGACGACAGACAGC
SNX3-Rev	ACGC<u>GTCGAC</u>CGGGAAACTTGGTGGATATG
SFP1-pUg34-Fw	TCC<u>CCCGGG</u>GATTTTACAACAATGACTATGGC
SHP1-pUg34-Rev	ACGC<u>GTCGAC</u>TTAGTGAGTGGAGTGGCCCCTGTG
MRS6-Fw	TCC<u>CCCGGG</u>TGTGCTTATCGGGGGTCTAC
MRS6-Rev	ACGC<u>GTCGAC</u>AGACCATCGTGAACCAAAGC
LST8-Fw	CGC<u>GGATCC</u>AGCAGCCTCGAGACCGTAT
LST8-Rev	ACGC<u>GTCGAC</u>GAAGAGCGTTTGAAGGCAAG
TCO89-Fw	CGC<u>GGATCC</u>CCTACGACCATCGAAAGAGC
TCO89-Rev	TCC<u>CCCGGG</u>GGGCTTTAGCGACAGAAACA
SEA4-Fw	AA<u>CTGCAG</u>CGCTTACATGCAACGTTTTG
SEA4-Rev	ACGC<u>GTCGAC</u>ATCTTCGCTGCCATTTTCAC
VAM6-Fw	CGC<u>GGATCC</u>GTTGGCGCCATGTGTGTATT
VAM6-Rev	ACGC<u>GTCGAC</u>GAACACCAGCCGGTATTAGC
LST7-Fw	AA<u>CTGCAG</u>TACAAAGTCATCGCCAGCAG
LST7-Rev	ACGC<u>GTCGAC</u>CGGTGATAACGATGGGAAAG
GTR2-Fw	CGC<u>GGATCC</u>TTTCAGAAGGGACGCTCCTC
GTR2-Rev	ACGC<u>GTCGAC</u>TTCTTCGGGTGTTGTCTTCC
TOR1-Fw	AA<u>CTGCAG</u>CCATGTGATCCCAATTTTCC
TOR1-Rev	ACGC<u>GTCGAC</u>GCAAGAGGGGGTACTTGGAC
LST4-Fw	CGC<u>GGATCC</u>TGTGGGCAATTGGGTGTACT
LST4-Rev	ACGC<u>GTCGAC</u>GATAACGGCCAAATGAAACG
GTR1-Fw	AA<u>CTGCAG</u>TTTTCAGCCGGGCAACATTT
GTR1-Rev	ACGC<u>GTCGAC</u>CCAAGTAGCGCACCAAGAGG
URA3-A	TGCACGAAAAGCAAACAAAC
URA3-B	AATGCGTCTCCCTTGTCATC
URA3-C	GAAGGTTAATGTGGCTGTGG
URA3-D	TTGGTTCTGGCGAGGTATTG

The sequences of restriction enezymes are underlined.

Table S1.

A. Frequencies of T-DNA integration by homologous recombination in yeast cells.

Yeast	*Agrobacterium*	Frequency of Ura$^+$ colonies per output recipient (mean $\pm$ SD, n=3)
W303a		$1.4 \pm 0.04 \times 10^{-5}$
BY4743	LBA1100 pRAL7100	$1.2 \pm 0.2 \times 10^{-4}$
BY4741		$9.7 \pm 0.3 \times 10^{-5}$

B. Frequencies of T-DNA integration by homologous recombination and non-homologous end-joining in yeast cells.

Yeast	*Agrobacterium*	Frequency of positive colonies per output recipient (mean $\pm$ SD, n=3)
BY4741		$1.1 \pm 0.6 \times 10^{-4}$
BY4741 *ada2ΔΔ*	LBA1100 pRAL7100	$4.4 \pm 0.8 \times 10^{-4}$
BY4741	LBA1100 pCAMBIA2300[*trp1*-kanMx]	$0.6 \pm 0.3 \times 10^{-4}$
BY4741 *ada2Δ*		$1.2 \pm 0.7 \times 10^{-4}$
BY4741	LBA1100 pRAL7102	$2.4 \pm 1.4 \times 10^{-6}$
BY4741 *ada2Δ*		$1.0 \pm 0.2 \times 10^{-4}$

C. Effect of the absence of *ADA2* on the fraction of transformants surviving on plates supplemented with FAA[a].

	Experiment			Average percentage $\pm$	P value
BY4741	6/198 (3.03)[b]	8/386 (2.07)	9/404 (2.22)	2.4 ± 0.5	
ada2Δ	45/209 (21.53)	156/482 (32.37)	44/157 (28.03)	27.3 ± 5.4	0.00141

a. The transformants were first selected for G418 resistance and were subsequently streaked on plates with or without FAA. Transformants with the T-DNA integrated in the *TRP1* locus are expected to grow on media with FAA.

b. The percentage of surviving transformants is given in parentheses. All yeast cells were transformed by the derivatives of *Agrobacterium* strain LBA1100 with corresponding T-DNA vectors listed in the table. The frequencies were calculated as the number of positive colonies divided by the whole output number of yeast cells and the average value of at least three independent experiments is shown.

Table S2. Selected genes involved in the TORC1 pathway for synthetic rescue studies.

Gene	Description of encoded protein
MRS6	rab escort protein
SCH9	required for TORC1-mediated regulation
KOG1	subunit of TORC1
TRA1	subunit of SAGA complex
TOR1	subunit of TORC1
LST8	component of the TOR signaling pathway
RPC82	subunit of RNA polymerase III
MOT1	protein involved in regulation of transcription
TCO89	subunit of TORC1
RAP1	DNA-binding transcription regulator
IFH1	coactivator of ribosomal protein genes
FHL1	Regulator of ribosomal protein transcription
GCR2	transcriptional activator of genes involved in glycolysis
GCR1	transcriptional activator of genes involved in glycolysis
YMR111C	enriches ubiquitin on chromatin
EGO2	subunit of the EGO/GSE complex
NPR3	nitrogen permease regulator
SEA4	subunit of the SEA complex
VAM6	vacuole membrane protein
MTC5	maintenance of telomere capping
SEH1	subunit of the SEA complex
LST7	subunit of the Lst4p-Lst7p GTPase complex for Gtr2p
GTR2	subunit of a TORC1-stimulating GTPase
LST4	subunit of the Lst4p-Lst7p GTPase complex for Gtr2p
SEC13	subunit of the SEA complex
GTR1	subunit of a TORC1-stimulating GTPase
RTC1	subunit of the SEA complex
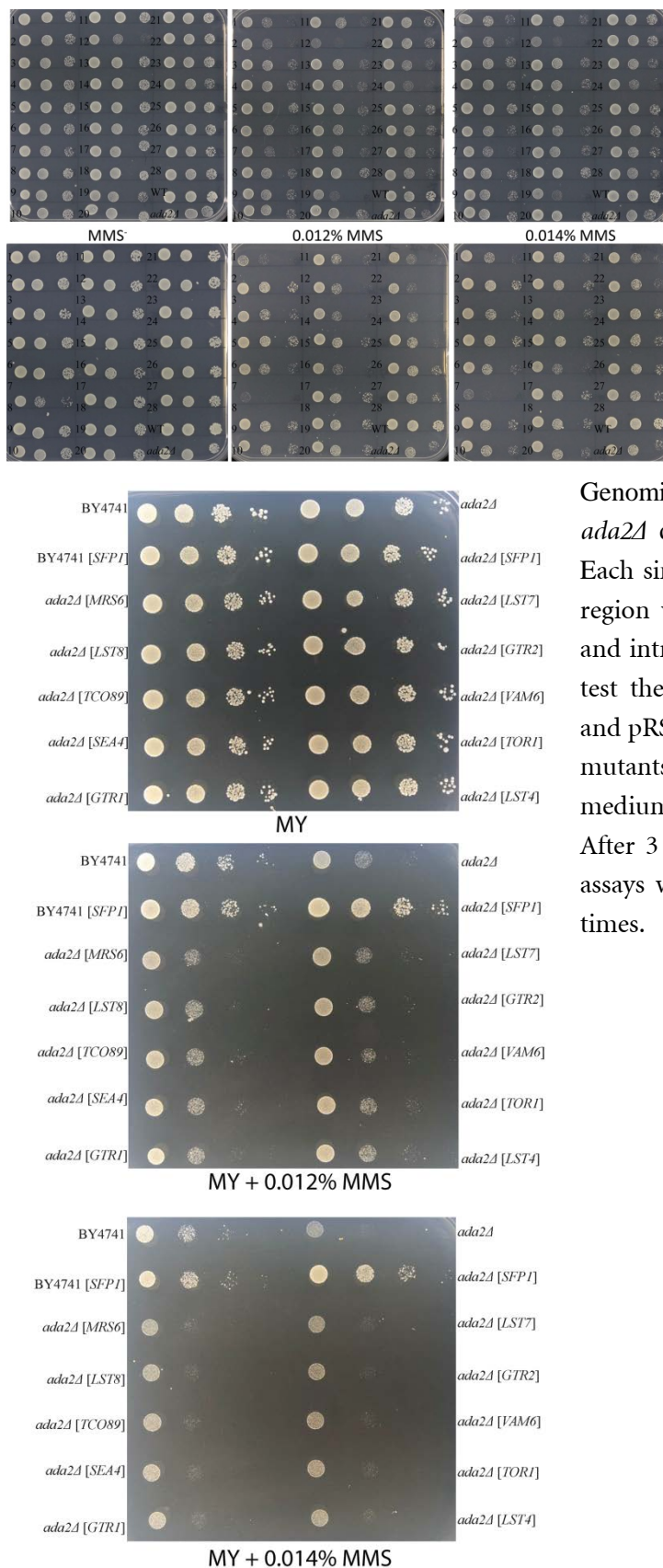
Figure S1. MMS sensitivity of *ada2Δ* deletion mutant could not be rescued by overexpression of genes related to the TORC pathway. (A) 27 candidate genes were selected to test whether overexpression of them can enhance or reduce sensitivity to MMS in the *ada2Δ* deletion mutant. These genes are involved in the TORC pathway or the regulation of TORC1 signaling. Overexpression vectors harboring these genes were obtained from the Yeast Genomic Tiling Collection and introduced into the *ada2Δ* deletion mutant via the LiAc method. (B) Each single candidate gene including its promoter region was cloned into high copy vector pRS425 and introduced into the *ada2Δ* deletion mutant to test the sensitivity to MMS. The vectors pRS425 and pRS425[SFP1] were introduced as controls. All mutants were serially diluted and spotted onto MY medium with the indicated concentrations of MMS. After 3 days, the photos were taken and the spot assays were performed at least three independent times.

References

Albert, B., Tomassetti, S., Gloor, Y., Dilg, D., Mattarocci, S., Kubik, S., & Shore, D. (2019). Sfp1 regulates transcriptional networks driving cell growth and division through multiple promoter-binding modes. *Genes & development*, 33, 288-293.

Anand, A., Vaghchhipawala, Z., Ryu, C. M., Kang, L., Wang, K., del-Pozo, Olga., Martin, G. B., & Mysore, K. S. (2007). Identification and characterization of plant genes involved in *Agrobacterium*-mediated plant transformation by virus-induced gene silencing. *Molecular plant-microbe interactions*, 20, 41–52.

Balasubramanian, R., Pray-Grant, M. G., Selleck, W., Grant, P. A., & Tan, S. (2002). Role of the Ada2 and Ada3 transcriptional coactivators in histone acetylation. *Journal of Biological Chemistry*, 277, 7989-7995.

Baudin, A., Ozier-Kalogeropoulos, O., Denouel, A., Lacroute, F., & Cullin, C. (1993). A simple and efficient method for direct gene deletion in *Saccharomyces cerevisiae. Nucleic acids research*, 21, 3329.

Beijersbergen, A., Den Dulk-Ras, A., Schilperoort, R. A., & Hooykaas, P. J. J. (1992). Conjugative transfer by the virulence system of *Agrobacterium tumefaciens. Science*, 256, 1324-1327.

Berger, S. L., Piña, B., Silverman, N., Marcus, G. A., Agapite, J., Regier, J. L., Triezenberg, S. J., & Guarente, L. (1992) Genetic isolation of ADA2: A potential transcriptional adaptor required for function of certain acidic activation domains. *Cell*, 70, 251-265

Berg, M. D., Genereaux, J., Karagiannis, J., & Brandl, C. J. (2018). The pseudokinase domain of *Saccharomyces cerevisiae* Tra1 is required for nuclear localization and incorporation into the SAGA and NuA4 complexes. *G3: Genes, Genomes, Genetics*, 8, 1943-1957.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14, 115-132.

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

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

Carlson, M., & Botstein, D. (1982). Two differentially regulated mRNAs with different 5′ ends encode secreted and intracellular forms of yeast invertase. *Cell*, 28, 145-154.

Christie, P. J., & Gordon, J. E. (2014). The *Agrobacterium* Ti plasmids. *Microbiology spectrum*, 2, 1-18.

Crane, Y. M., & Gelvin, S. B. (2007). RNAi-mediated gene silencing reveals involvement of *Arabidopsis* chromatin-related genes in *Agrobacterium*-mediated root transformation. *Proceedings of the National Academy of Sciences*, 104, 15156-15161.

De Groot, M. J., Bundock, P., Hooykaas, P. J. J., & Beijersbergen, A. G. (1998). *Agrobacterium* tumefaciens-mediated transformation of filamentous fungi. *Nature biotechnology*, 16, 839.

den Dulk-Ras, A., & Hooykaas, P. J. J. (1995). Electroporation of *Agrobacterium tumefaciens*. In *Plant Cell Electroporation and Electrofusion Protocols*, 63-72. Springer, Totowa, NJ.

Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, 11, 355-360.

Grant, P. A., Duggan, L., Côté, J., Roberts, S. M., Brownell, J. E., Candau, R., Ohba, R., Owen-Hughes, T., Allis, C. D., Winston, F., Berger, S. L., & Workman J. L. (1997). Yeast Gcn5 functions in two multisubunit complexes to acetylate nucleosomal histones: characterization of an Ada complex and the SAGA (Spt/Ada) complex. *Genes & development*, 11, 1640-1650.

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and molecular biology reviews*, 67, 16-37.

Gelvin, S. B. (2017). Integration of *Agrobacterium* T-DNA into the plant genome. *Annual review of genetics*, 51, 195-217.

Goldstein, A. L., & McCusker, J. H. (1999). Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast*, 15, 1541-1553.

Güldener, U., Heck, S., Fiedler, T., Beinhauer, J., & Hegemann, J. H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic acids research*, 24, 2519-2524.

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

Hark, A. T., Vlachonasios, K. E., Pavangadkar, K. A., Rao, S., Gordon, H., Adamakis, I. D., Kaldis, A., Thomashow, M. F., & Triezenberg, S. J. (2009). Two *Arabidopsis* orthologs of the transcriptional coactivator ADA2 have distinct biological functions. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1789, 117-124.

Hoke, S. M., Genereaux, J., Liang, G., & Brandl, C. J. (2008). A conserved central region of yeast Ada2 regulates the histone acetyltransferase activity of Gcn5 and interacts with phospholipids. *Journal of molecular biology*, 384, 743-755.

Hooykaas, P. J. J., van Heusden, G. P. H., Niu, X., Roushan, M. R., Soltani, J., Zhang, X., & van der Zaal, B. J. (2018). Agrobacterium-mediated transformation of yeast and fungi. *Current Topics in Microbiology and Immunology*, 418, 349-374.

Jacobson, S., & Pillus, L. (2009). The SAGA subunit Ada2 functions in transcriptional silencing. *Molecular and cellular biology*, 29, 6033-6045.Janke, C., Magiera, M. M., Rathfelder, N., Taxis, C., Reber, S., Maekawa, H., Moreno‐Borchart, A., Doenges, G., Schwob, E., Schiebel, E., & Knop, M. (2004). A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast*, 21, 947-962.

Jiang, Y., Berg, M. D., Genereaux, J., Ahmed, K., Duennwald, M. L., Brandl, C. J., & Lajoie, P. (2019). Sfp1 links TORC1 and cell growth regulation to the yeast SAGA-complex component Tra1 in response to polyQ proteotoxicity. *Traffic*, 20, 267-283.

Jorgensen, P., Nishikawa, J. L., Breitkreutz, B. J., & Tyers, M. (2002). Systematic identification of pathways that couple cell growth and division in yeast. *Science*, 297, 395-400.

Jorgensen, P., Rupeš, I., Sharom, J. R., Schneper, L., Broach, J. R., & Tyers, M. (2004). A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes & development*, 18, 2491-2505.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., & Punna, T. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*, 440, 637.

Lai, J., Jiang, J., Wu, Q., Mao, N., Han, D., Hu, H., & Yang, C. (2018). The transcriptional coactivator ADA2b recruits a structural maintenance protein to double-strand breaks during DNA repair in plants. *Plant physiology*, 176, 2613-2622.

Lempiäinen, H., Uotila, A., Urban, J., Dohnal, I., Ammerer, G., Loewith, R., & Shore, D. (2009). Sfp1 interaction with TORC1 and Mrs6 reveals feedback regulation on TOR signaling. *Molecular cell*, 33, 704-716.

Leung, G. P., Lee, L., Schmidt, T. I., Shirahige, K., & Kobor, M. S. (2011). Rtt107 is required for recruitment of the SMC5/6 complex to DNA double-strand breaks. J*ournal of Biological Chemistry*, 286, 26250-26257.

Lisby, M., Rothstein, R., & Mortensen, U. H. (2001). Rad52 forms DNA repair and recombination centers during S phase. *Proceedings of the National Academy of Sciences*, 98, 8276-8282.

Lum, P. Y., Armour, C. D., Stepaniants, S. B., Cavet, G., Wolf, M. K., Butler, J. S., Hinshaw, J. C., Garnier, P., Prestwich, G. D., Leonardson, A., Garrett-Engele, P., Rush, C. M., Bard, M., Schimmack, G., Phillips, J. W., Roberts, C. J., & Shoemaker, D. D. (2004). Discovering modes of action for therapeutic compounds using a genome-wide screen of yeast heterozygotes. *Cell*, 116, 121-137.

Luo, Y., Chen, Z., Zhu, D., Tu, H., & Pan, S. Q. (2015). Yeast actin-related protein ARP6 negatively regulates *Agrobacterium*-mediated transformation of yeast cell. *BioMed research international*, 2015, 275092.

MacKay, V. L., Welch, S. K., Insley, M. Y., Manney, T. R., Holly, J., Saari, G. C., & Parker, M. L. (1988). The Saccharomyces cerevisiae BAR1 gene encodes an exported protein with homology to pepsin. *Proceedings of the National Academy of Sciences*, 85, 55-59.

Magori, S., & Citovsky, V. (2011). Epigenetic control of *Agrobacterium* T-DNA integration. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1809, 388-394.

Marion, R. M., Regev, A., Segal, E., Barash, Y., Koller, D., Friedman, N., & O'Shea, E. K. (2004). Sfp1 is a stress-and nutrient-sensitive regulator of ribosomal protein gene expression. *Proceedings of the National Academy of Sciences*, 101, 14315-14322.

McKinney, J. S., Sethi, S., Tripp, J. D., Nguyen, T. N., Sanderson, B. A., Westmoreland, J. W., Resnick. M. A., & Lewis, L. K. (2013). A multistep genomic screen identifies new genes required for repair of DNA double-strand breaks in *Saccharomyces cerevisiae*. *BMC genomics*, 14, 251.

Muñoz-Galván S, Jimeno S, Rothstein R, Aguilera A (2013) Histone H3K56 Acetylation, Rad52, and Non-DNA Repair Factors Control Double-Strand Break Repair Choice with the Sister Chromatid. *PLoS Genetics*, 9, e1003237

Muratoglu, S., Georgieva, S., Pápai, G., Scheer, E., Enünlü, I., Komonyi, O., Cserpan, I., Lebedeva, L., Nabirochjina, E., Udvardy, A., Tora, L., & Boros, I. (2003). Two different *Drosophila* ADA2 homologues are present in distinct GCN5 histone acetyltransferase-containing complexes. *Molecular and cellular biology*, 23, 306-321.

Mysore, K. S., Nam, J., & Gelvin, S. B. (2000). An Arabidopsis histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *Proceedings of the National Academy of Sciences*, 97, 948-953.

Nester, E. W., Gordon, M. P., Amasino, R. M., & Yanofsky, M. F. (1984). Crown gall: a molecular and physiological analysis. *Annual Review of Plant Physiology*, 35, 387-413.

Niu, X., Zhou, M., Henkel, C. V., van Heusden, G. P. H., & Hooykaas, P. J. (2015). The *Agrobacterium tumefaciens* virulence protein VirE3 is a transcriptional activator of the F-box gene VBF. *The Plant Journal*, 84, 914-924.

Offringa, R., De Groot, M. J., Haagsman, H. J., Does, M. P., Van Den Elzen, P. J., & Hooykaas, P. J. J. (1990). Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *The EMBO journal*, 9, 3077-3084.

Ohmine, Y., Satoh, Y., Kiyokawa, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2016). DNA repair genes RAD52 and SRS2, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene ERG28 are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. *BMC microbiology*, 16, 58.

Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, C., Bellini, C. (2011). *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. *Physiological and molecular plant pathology*, 76, 76-81.

Rolloos, M., Dohmen, M. H., Hooykaas, P. J. J., & van der Zaal, B. J. (2014). Involvement of Rad52 in T-DNA circle formation during *Agrobacterium* tumefaciens-mediated transformation of *Saccharomyces cerevisiae*. *Molecular microbiology*, 91, 1240-1251.

Saleh, A., Schieltz, D., Ting, N., McMahon, S. B., Litchfield, D. W., Yates, J. R., Lees-Miller, S. P., Cole, M.D., & Brandl, C. J. (1998). Tra1p is a component of the yeast Ada Spt transcriptional regulatory complexes. *Journal of Biological Chemistry*, 273, 26559-26565.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J. Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., & Cardona, A. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9, 676.

Schrammeijer, B., Risseeuw, E., Pansegrau, W., Regensburg-Tuïnk, T. J., Crosby, W. L., & Hooykaas, P. J. (2001). Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. *Current Biology*, 11, 258-262.

Shilo, S., Tripathi, P., Melamed-Bessudo, C., Tzfadia, O., Muth, T. R., & Levy, A. A. (2017). T-DNA-genome junctions form early after infection and are influenced by the chromatin state of the host genome. *PLoS genetics*, 13, e1006875.

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

Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. (2012). Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant physiology*, 160, 511-522.

Soltani, J. (2009). Host genes involved in *Agrobacterium*-mediated transformation. Ph.D. thesis, Leiden University, Leiden, The Netherlands.

Soltani, J., Van Heusden, G. P. H., Hooykaas, P. J. J. (2009). Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*. *FEMS microbiology letters*, 298, 228-233.

Sterner, D. E., Wang, X., Bloom, M. H., Simon, G. M., & Berger, S. L. (2002). The SANT domain of Ada2 is required for normal acetylation of histones by the yeast SAGA complex. *Journal of Biological Chemistry*, 277, 8178-8186.

Straight, A. F., Belmont, A. S., Robinett, C. C., & Murray, A. W. (1996). GFP tagging of budding yeast chromosomes reveals that protein–protein interactions can mediate sister chromatid cohesion. *Current Biology*, 6, 1599-1608.

Sun, J., Paduch, M., Kim, S. A., Kramer, R. M., Barrios, A. F., Lu, V., Luke, J., Usatyuk, S., Kossiakoff, A. A., & Tan, S. (2018). Structural basis for activation of SAGA histone

acetyltransferase Gcn5 by partner subunit Ada2. *Proceedings of the National Academy of Sciences*, 115, 10010-10015.

Thomas, B. J., & Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell*, 56, 619-630.

Toyn, J. H., Gunyuzlu, P. L., Hunter White, W., Thompson, L. A., & Hollis, G. F. (2000). A counterselection for the tryptophan pathway in yeast: 5-fluoroanthranilic acid resistance. *Yeast*, 16, 553-560.

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. *The EMBO Journal*, 20, 3596-3607.

Tzfira, T., & Citovsky, V. (2006). *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. *Current opinion in biotechnology*, 17, 147-154.

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

Van Attikum, H., & Hooykaas, P. J. J. (2003). Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in Saccharomyces cerevisiae. *Nucleic acids research*, 31, 826-832.

Van Kregten, M., de Pater, S., Romeijn, R., van Schendel, R., Hooykaas, P. J. J., & Tijsterman, M. (2016). T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nature plants*, 2, 16164.

Winans, S. C. (1992). Two-way chemical signaling in *Agrobacterium*-plant interactions. *Microbiological reviews*, 56, 12-31.

Ward, E. R., & Barnes, W. M. (1988). VirD2 protein of *Agrobacterium tumefaciens* very tightly linked to the 5'end of T-strand DNA. *Science*, 242, 927.

Yadav, N. S., Vanderleyden, J., Bennett, D. R., Barnes, W. M., & Chilton, M. D. (1982). Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proceedings of the National Academy of Sciences*, 79, 6322-6326.

Zhang, X., van Heusden, G. P. H., & Hooykaas, P. J. J. (2017). Virulence protein VirD5 of *Agrobacterium tumefaciens* binds to kinetochores in host cells via an interaction with Spt4. *Proceedings of the National Academy of Sciences*, 114, 10238-10243.

Zhu, J., Oger, P. M., Schrammeijer, B., Hooykaas, P. J.J., Farrand, S. K., & Winans, S. C. (2000). The bases of crown gall tumorigenesis. *Journal of Bacteriology*, 182, 3885-3895.

Chapter 3
Detection of extrachromosomal circular T-DNA formed during *Agrobacterium*-mediated transformation

Shuai Shao, Daniela M. d'Empaire Altimari, G. Paul. H. van Heusden, Paul J. J. Hooykaas

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Abstract

Agrobacterium tumefaciens is able to transfer a segment of its Ti-plasmid, called transferred DNA (T-DNA), to plant cells and under laboratory conditions also to other eukaryotes such as yeast and fungi. Inside the host cell the T-DNA integrates into the chromosomal DNA. Alternatively, the T-DNA can be circularized and form (sometimes complex) extrachromosomal structures. In this chapter, we adopted an approach developed for plants to recover circularized extrachromosomal T-DNA structures from transformed yeast and plants. With this approach we isolated circularized T-DNA from wild type and *ada2Δ* mutant yeast transformants and from the plant *Nicotiana benthamiana*. Sequence analysis showed that the isolated T-circles have very different structures, some containing a precise fusion of the left and right border repeat sequences, whereas others contain filler DNA originating either from the binary vector or from the host genome. Our data also suggest that the frequency of T-circle formation is higher in the *ada2Δ* deletion mutant than in wild type yeast cells.

Introduction

Agrobacterium tumefaciens can induce tumor formation in plants by transferring a segment of its tumor-inducing plasmid (Ti-plasmid) to the plant cell. Tumorigenic DNA within the transferred DNA (T-DNA) can be substituted by foreign DNA without perturbing its ability to integrate into the plant genome. Under laboratory conditions, other eukaryotes such as yeast and fungi can also be transformed. Therefore, *Agrobacterium*-mediated transformation (AMT) is widely used to create transgenic plants and fungi (for review see: Nester et al., 1984; Gelvin, 2003; Tzfira and Citovsky, 2006; Păcurar et al., 2011; Christie and Gordon, 2014; Gelvin, 2017; Hooykaas et al., 2018).

The process of T-DNA integration has been widely studied and several alternative mechanisms have been proposed including T-strand invasion and double-strand T-DNA ligation into the host genome. Generally the integration is not precise and clean, leading to complex junction patterns between T-DNA and the plant genome. Deletions of host DNA at the integration sites are common and insertion is often accompanied by the insertion of "filler DNA" from neighbor sequences or from distant areas (reviewed in Gelvin, 2017). At the integration sites often multiple copies are present, commonly in an inverted repeat structure. Besides, double-stranded non-integrated T-DNA molecules were occasionally found in transformed plants (Offringa et al, 1990; Narasimhulu et al., 1996). Some of these might serve as intermediates for integration and eventually result in the formation of various T-DNA structures. Alternatively, they are only temporarily present in a linear (concatemeric) form or a circular form (T-circle).

T-circles were first recovered from bacteria after activation of the virulence genes. They were formed by precise DNA recombination of nicked border repeats and rescued by using λ *in vitro* packaging followed by plasmid rescue (Koukolilova-Nicola et al., 1985). At the time, before the discovery of the T-strand, they were thought to be transfer intermediates. Subsequently, a virus-based system was used to recover T-circles after transfer by replacing the T-DNA genes by a virus genome (Bakkeren et al., 1989). Rescued virus had a structure compatible with non-precise end-joining of a T-strand, with loss of few nucleotides from the LB and insertion of 1-14 nt filler from unknown origin. Based on these ideas, Singer et al. developed a plasmid-recue approach to recover extrachromosomal T-circles from *Agrobacterium* infected plants (Singer et al., 2012). By sequence analysis, it was shown that the junctions between the ends of the T-DNA included T-DNA border fusions, T-DNA truncations, binary plasmid sequences, and filler DNA sequences, suggesting that they were formed by a process resembling T-DNA integration.

In yeast extrachromosomal T-circles could be stably maintained, when they contained a replication unit (Bundock et al., 1995; Soltani, 2009; Rolloos et al, 2014; Ohmine et al., 2016). For instance, T-DNAs containing the 2μ replication origin exhibited high transformation efficiencies and were present as plasmids generated by the fusion of left and right border (Bundock et al., 1995). To date there is no evidence to support the idea that T-circles can further integrate into the host genome in nature even though T-circle formation may share a similar mechanism as T-DNA integration. Moreover, T-DNAs containing two directly repeated *lox* sites were reported to circularize and integrate through Cre/*lox*-mediated single cross over in the plant genome at a genomic *lox* site, but no evidence for similar Cre/*lox*-mediated

integration was found when using a T-DNA with a single *lox* site, whereby circularization would be dependent on border fusion (Vergunst and Hooykaas, 1998).

The formation of T-circles could be beneficial for *Agrobacterium* because the transformed plant cells would already start generating opines and tumor producing hormones before or even without the need of genomic integration of the T-DNA. Although probably less prone to integration, circular T-DNA molecules are probably more stable than linear molecules and may thus be important for transient expression and intermediates for further T-DNA integration. In this chapter, we adapted the approach developed by Singer et al. to recover T-circles from yeast and plants. To this end, we constructed a new T-DNA vector containing a selection marker and replication unit for *E.coli*, but not for yeast and plant cells. We compared the recovery of T-circles from wildtype yeast and *ada2Δ* mutant, and determined the structures of them in comparison to those from plants.

Methods

Bacterial cells and growth conditions

Agrobacterium tumefaciens strain LBA1100 and its derivatives used in this chapter are listed in Table 1. All the plasmids were transferred into *Agrobacterium* by electroporation (den Dulk-Ras and Hooykaas, 1995). *A. tumefaciens* strains were grown and cultured at 29°C in LC medium (10 g/L tryptone, 5 g/L yeast extract and 8 g/L NaCl) containing the appropriate antibiotics such as carbenicillin (100 μg/mL), kanamycin (100 μg/mL) or rifampicin (20 μg/mL) when required. *Escherichia coli* DH10B was used for plasmid amplification and grown on LC plates with carbenicillin (100 μg/mL) and kanamycin (100 μg/mL) as required.

Yeast strains and growth conditions

S. cerevisiae strains used in this study can be found in Table 1. Yeast was grown at 30°C in yeast extract-peptone-dextrose (YPD) medium supplemented when required with the appropriate antibiotic hygromycin (200 μg/mL).

Plasmid constructions

The plasmids used in this chapter are listed in Table 2. All newly constructed plasmids were checked by restriction digestion and sequence analysis. All primes used are presented in Table 3.

For detection of T-circles, the binary vector pSDM4708 was constructed as follows. A T-DNA fragment containing the KAN.MX cassette was amplified by PCR using pSDM8000 as a template and pSDM8000-Fw/Rev as primers and then cloned into plasmid pBBR6 after digestion with *Not*I and *Apa*I to form plasmid pSDM4710. Then pUC9 was linearized with *Sal*I and ligated into *Sal*I digested pSDM4710, generating plasmid pSDM4711. The linearized pUC9 with ampicillin resistance gene replaced the KAN.MX cassette in the T-DNA. Subsequently, the bacterial kanamycin resistance gene (Km^R) was amplified by PCR on the vector pCAMBIA1390 using primers Km-Fw/Rev thereby introducing *Not*I and *Apa*I sites at the ends and inserted into plasmid pSDM4711 after digestion with *Not*I and *Apa*I to remove the pBBR6 vector part, producing plasmid pSDM4712. Finally, a fragment containing the pVS1 replication origin of replication needed for maintenance in *Agrobacterium* was amplified from vector pCAMBIA1390 using the primers pVS1-Fw/Rev and cloned into the *Apa*I site of pSDM4712, resulting in plasmid pSDM4708 (Figure 1A).

Agroinfiltration

Seeds of *Nicotiana benthamiana* were germinated and grown in controlled climate chambers at 24°C with a 16h light/8h dark photoperiod with 75% humidity for three weeks before infiltration. *A. tumefaciens* strains were cultured overnight at 29°C. After dilution to A_{600} of around 0.8 in 10 ml of induction medium (IM) containing 200 μM acetosyringone (AS), the cultures were grown for three hours at 28°C. Subsequently, the cultures were transferred into a blunt-tipped plastic 10 ml syringe (Nissho NIPRO Europe N.V.) and injected into the leaves of the *N. benthamiana* line. This was done by applying the tip of the syringe to the lower surface of the leaves and injecting with gentle pressure. Young leaves of three weeks old plants were used for agroinfiltration. After 2 days, the lower side of the injected leaf was harvested for DNA isolation (Wroblewski et al., 2005).

***Agrobacterium* mediated transformation of yeast**

AMT efficiency was determined as described by Bundock et al. (1995) with some modifications. Firstly, *S. cerevisiae* strains and *Agrobacterium* strains were grown overnight at 30°C and 28°C respectively, under continuous shaking and with the appropriate antibiotic selection. The following day, the *Agrobacterium* culture was centrifuged, the cells were washed with induction medium (IM) and re-suspended to an A_{620} of 0.25 in IM with added glucose (10 mM), acetosyringone (AS, 0.2 mM) and appropriate antibiotics, and incubated for another 6 hours at 28°C. Meanwhile, yeast cultures were diluted to an A_{620} of 0.1 and incubated again in liquid YPD for 6 hours. Then, yeast cultures were centrifuged and cells were washed with IM and re-suspended in 0.5 ml of IM to a final A_{620} of 0.4 - 0.6 and mixed vigorously with an equal volume of Agrobacterium cells. Subsequently, 100 μl of the mixture were pipetted onto sterile nitrocellulose filters laid on IM plates supplemented with histidine, leucine and methionine. Once the filters were dry, plates were incubated at 21 °C for 3 days, after which DNA was isolated from the mixture

DNA isolation, recovery of T-circle, sequencing and analysis

For DNA isolation from plants, leaves were harvested and disrupted to powders under liquid N_2 in a Tissuelyser (Retch, Haan, Germany). Total DNA was extracted according to the CTAB method (De Pater et al., 2009). DNA isolation from yeast was performed following standard rapid isolation protocol (Sambrook and Russell, 2006). Isolated DNA was used for the transformation of *E.coli* strain DH10B with selection for carbenicillin and kanamycin resistance. Plasmid DNA was isolated from Cb^R/Km^S colonies.

For analysis of the extrachromosomal structures the junctions between the LB and RB sequences were amplified via PCR with the primers pA1/pA2 or pA3/pA4. Then, PCR products were purified by agarose gel electrophoresis and sequenced by Macrogen (Amsterdam, the Netherlands). Primers were ordered from Sigma-Aldrich.

Results

A plasmid-rescue method to capture T-circles from plant and yeast cells

We were interested in capturing T-circles for which we found genetic evidence in Chapter 2 in the yeast *ada2Δ* mutant. As no T-circles had previously been recovered from transformed yeast without prior selection of transformants, we first tested the recovery of T-circles from plant cells. In order to study circularization of T-DNA we modified the plasmid-rescue technique described previously (Singer et al., 2009; Rolloos et al., 2014). To this end, we constructed plasmid pSDM4708 containing T-DNA with the ampicillin resistance marker and the MP1 origin of replication (Figure 1A). Upon circularization of the T-DNA in the host cell a plasmid is formed with an origin of replication for *E. coli* and an ampicillin resistance marker allowing amplification and selection in *E. coli*. This construct was introduced in *Agrobacterium* strain LBA1100 (lacking T-DNA). The T-DNA construct was introduced into plant leaves through agroinfiltration and into yeast cells by co-cultivation using the AMT protocols mentioned in Chapter 2. At 2 days post agroinfiltration or co-cultivation, total DNA was isolated from the mixture of *Agrobacterium* and infected host cells, including the T-DNA binary vectors from *A. tumefaciens* itself. To remove the original T-DNA binary vector pSDM4708, the isolated DNA was digested with *EcoRV* which cuts the plasmid in the backbone, but not in the T-DNA (Figure 1B). The digested DNA was used to transform *E. coli* and bacteria containing circularized T-DNA (T-circles) were selected on plates containing ampicillin.

In our experiments using the wild type yeast strain BY4741 we isolated only a very low number of ampicillin resistant, but kanamycin sensitive (backbone marker) *E. coli* colonies (less than 10 per experiment). As the *ada2Δ* mutant gives a higher transformation frequency (Chapter 2) we also isolated T-circles after AMT of this mutant. This resulted in the isolation of more than 30 ampicillin resistant *E. coli* colonies per experiment, whereby the proportion of kanamycin sensitive was 10 out of 80. In addition, the same approach was applied to transform the plant *N. benthamiana*, which was used successfully in previous research (Singer et al., 2012). Also from the plant we were able to obtain several ampicillin resistant kanamycin sensitive *E. coli* colonies for further analysis. The number of colonies obtained from DNA isolated from plants was much higher than the number of colonies obtained from DNA isolated from yeast. Among those resistant to ampicillin, kanamycin sensitive colonies occupied around 17% (16 out of 96). As there is an alternative possibility that T-circles are generated inside *A. tumefaciens* instead of in the host cell, we also used *A. tumefaciens* strains LBA1143 and LBA1144 carrying pSDM4708, in which *virB4* and *virB7*, respectively, are mutated and thus fail to form a T4SS apparatus. After AMT of the *ada2Δ* mutant with either of these *virBΔ* mutant strains, we did not obtain any ampicillin resistant kanamycin sensitive *E. coli* colonies, suggesting that T-circle formation occurs inside the host cell after transfer rather than in *A. tumefaciens* prior to transformation, in line with previous results showing that T-DNA transfer is required for T-circle formation (Singer et al., 2012).

Figure 1. Constructs and experimental procedure for the isolation of T-circles. A, schematic diagram of the T-DNA construct pSDM4708 used throughout the experiments. B, illustration of the experimental procedure: 1, *A. tumefaciens* carrying the binary plasmid pSDM4708 was injected into the leaves of *N. benthamiana* or co-culvitated with yeast cells; 2, extraction of DNA from the agroinfiltrated leaves or from cells of the Agrobacterium-yeast co-cultivation mixture (3 days after infiltration/cocultivation); 3, digestion with restriction enzymes *NotI* and *EcoRV* to remove the original pSDM4708; 4, transformation of *E. coli* competent cells with digested DNA; 5, selection of *E. coli* transformants resistant to ampicillin; 6, further selection of *E. coli* transformants sensitive to kanamycin; 7, isolation of plasmid-like DNA from each independent colony and analysis by DNA sequencing. Amp, ampicillin; Km, kanamycin.
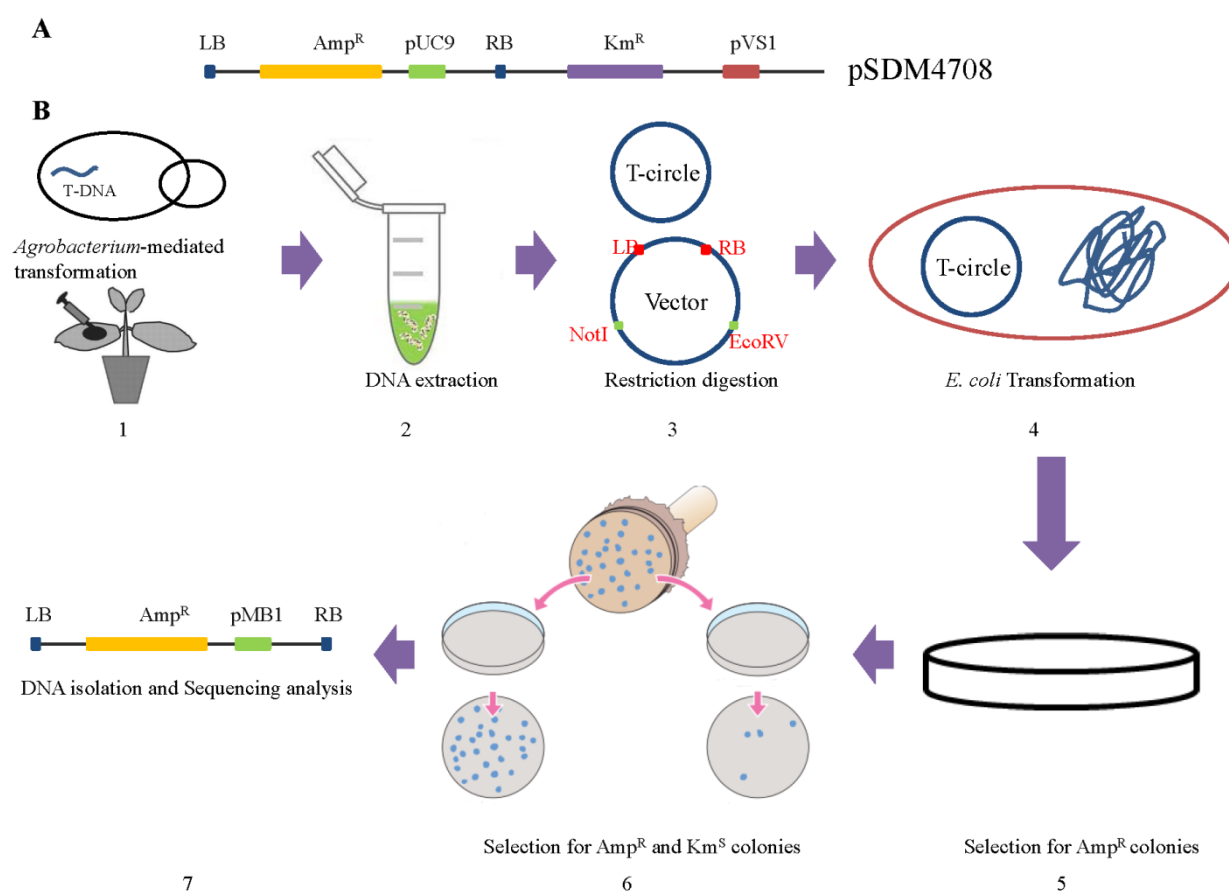
Details of T-DNA junctions recovered from host cells

To characterize the T-circles, the LB-RB junctions were analyzed by DNA sequencing of either PCR products obtained by using two pairs of primers (pA1/A2 or pA3/A4) or by direct sequencing the isolated T-circles. In total, sixteen T-circles isolated from plant *N. benthamiana* and ten T-circles isolated from the yeast *ada2Δ* mutant were analyzed. The structures of six representatives (T-1 to T-6) are shown in Figure 2. The majority of the T-circles isolated from yeast (7 out of 10) had the same structure as T-1. The other three T-circles have a different structure (T5 and two times T6). On the other hand, of the T-circles isolated from plants, only

1 out of 16 had the same structure as T1. The majority (13 out of 16) of plant derived T-circles had a structure like T-3 and two had a different structure (T-2 and T-4).
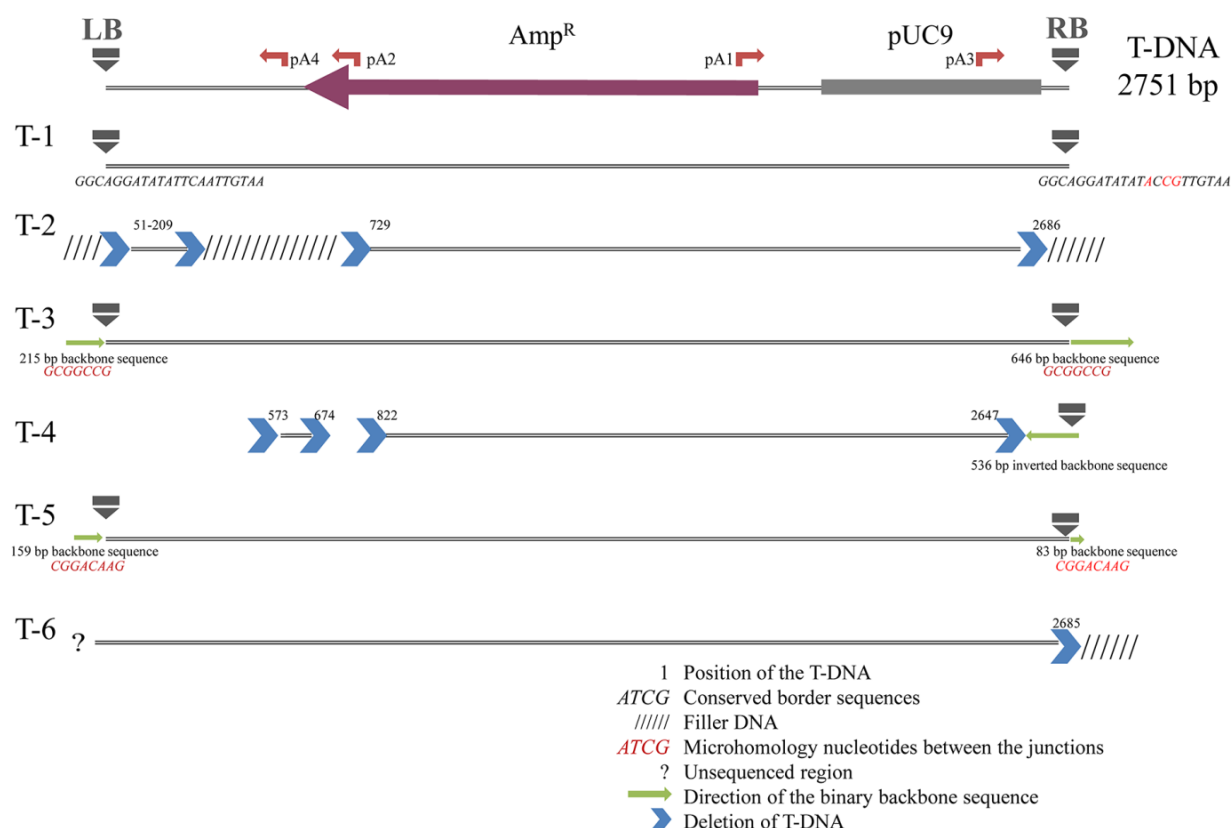

Figure 2. Schematic illustration of the T-circles T-1 to T-6 based on DNA sequencing. T-1 to T-6 are shown in a linear form, opened at the position of the junction, and aligned to the complete construct-DNA of plasmid pSDM4708. The annotations represent different events identified by sequence analysis.

In T-1, a T-circle isolated from yeast, a precise border fusion was found, compatible with a reversion of the nicking reaction of relaxase VirD2. This could immediately lead to circularization (monomeric reaction) or first lead to concatemerization (dimer or multimer formation), after which HR is necessary for conversion in a T-circle as suggested by Rolloos et al., 2014. A similar structure of T-circle was also identified in infected plant cells at a much lower frequency (1 out of 16).

The structure of T-2, isolated from plant, is complex. Instead of a RB-LB fusion two pieces of filler DNA were found with T-DNA sequences in between. It contained a patchwork of filler sequences ("scrambled") interrupted by short T-DNA sequences (159 bp). Both border sequences were lost. At the LB side two fragments (50 bp and 520 bp) were missing and from the RB side of T-DNA, a 66-bp fragment including right border sequence disappeared. All these filler sequences start with short sequences homologous with parts of the associated *vir*-helper plasmid in *Agrobacterium* (79 bp, 49 bp and 63 bp). Previous studies have already shown that transfer of *Agrobacterium* chromosomal DNA fragments is possible (Ulker et al., 2008). By

tracing the sources of filler sequence found at T-DNA integration sites in plants, around 1.1% of them turned out to be homologous with *Agrobacterium* genome sequences (Kleinboelting et al., 2015). Taken together, we hypothesize that some DNA fragments could be translocated into plant cells independent of T-DNA and these fragments may further serve as templates for microhomology-mediated end-joining.

In T-3, isolated from plant, two stretches of backbone sequences adjacent to either side of the T-DNA were present at the junction. The length of the stretch of backbone sequences at the LB side is 215 bp and that of the RB side is 646 bp. They may have been fused due to 7 bp of microhomologous sequences (GCGGCCGC) at the end of them. Such "read-through" binary backbone sequence fragments were reported before and probably caused by improper "border skipping" of the VirD2 endonuclease (Kononov et al., 1997). They are most likely translocated into host cells as single strand DNA, then converted to double strand molecules and finally ligated with microhomology end-joining. A similar structure was present in 13 out of 16 T-circles isolated from *N. benthamiana.*

In T-4, isolated from plant, at the LB side two large fragments including the border sequence are missing. The larger deletion (572 bp) is at the LB and the other deletion corresponds to nucleotides 675-821 of the original T-DNA. The RB side is incomplete and is replaced by an 536 bp inverted backbone fragment from inside to outside of T-DNA. The RB is still intact, while, the orientation is reverse, thus indicating a possible RB-RB junction formed between two T-DNA repeats. This distinct structure may be formed when two T-strands were delivered into plant cells independently, then converted to dsDNA and eventually linked to form a T-circle. Similarly, a repeat feature is frequently found at the T-DNA integration sites of plant cells after AMT (Krizkova and Hrouda, 1998) and also supported by the co-integration of different T-DNAs from different binary vectors at the same site following co-transformation (De Neve et al., 1997).

With respect to the T-circles recovered from yeast, most of them were generated by the "perfect" fusions of LB and RB (T-1). An example of an imprecise T-circle is represented as T-5. Two short binary read-through backbone sequences adjacent to either repeat border (159 bp at the LB side and 83 bp at the RB side) were linked together, possibly because of the microhomology at the end of them as in the T-3 structure.

In T-6, isolated from yeast, small fragments close to both the LB and RB are missing and are replaced by large filler sequences with strong homology with yeast chromosomal DNA. However, unfortunately, we were not able to sequence the internal region of the fragment, possibly due to the formation of secondary structures. Unlike filler sequences in T-2, they were identified to share homology with host genomic DNA, such as with the *BUD3* locus in chromosome III and the *ECM38* locus in chromosome XII from yeast in two junctions. The relationship between them has not been identified and seems to be obtained by random integration. These patterns of filler sequence are often seen at the T-DNA integration sites in *Agrobacterium*-infected plants as well and may indicate that a random DNA template can potentially be used utilizing microhomology by DNA polymerase during end-joining or integration.

Discussion

It has been shown that the AMT efficiency of yeast is higher, when the T-DNA contains an autonomous replication origin, than when the T-DNA lacks such a replication origin and needs to be integrated into one of the chromosomes for stable maintenance (Bundock et al., 1995; Soltani, 2009). However, maintenance of T-DNAs with an autonomous replication origin requires circularization. Circles were formed by precise joining of the RB and LB ends, possibly mediated by VirD2 as a reversal of the nicking reaction. This reaction could immediately lead to circularization (monomeric reaction) or lead to concatemer formation first (dimeric and multimeric reaction), which can be resolved by the homologous recombination pathway (Soltani, 2009; Rolloos et al., 2014; Ohmine et al., 2016). Non-integrated T-DNA "free" molecules (T-circles) were detected in transformed plants as well in early stages after transformation (Singer et al., 2012). These seem to be formed by non-homologous end joining of the ends. To our knowledge, the formation and (temporary) presence of extrachromosomal T-circles without host replication origin in yeast has not been reported as yet. Here we obtained evidence for the presence of such T-circles without any replication origin after AMT of yeast cells.

The complex extrachromosomal T-DNA structures identified in this research contain various formats including perfect fusions of LB and RB, binary plasmid sequences, filler host DNA sequences and T-DNA truncations (Figure 2). These observations are consistent with previous results from similar experiments with plants (Singer et al., 2012). The ligation patterns present in T-circles mentioned above are thought to resemble the parts of repaired DNA double-strand breaks sites, in other words, it is likely that the formation of extrachromosomal T-circles share the same DNA repair mechanisms with the integration of T-DNA into host genome. Recently, the successful transfer of a T-DNA plasmid to *E.coli* has been reported (Ohmine et al., 2018). They co-cultivated *Agrobacterium* containing a modified T-DNA plasmid harboring a RB but not a LB with *E.coli* or yeast. The modified plasmid would be nicked by VirD2, subsequently guided into the recipient cells as single-strand and eventually circularized and maintained as plasmid. In bacteria, the AMT efficiency was affected by RecA which plays a crucial role in DNA HR repair pathway. It is in line with the observation that T-DNA circularization efficiency relies on Rad52 and Rad51 in yeast (Rolloos et al., 2014; Ohmine et al., 2016). Compared to wild type yeast cells, the efficiency was reduced around 70-75% when using *rad52Δ* deletion mutant as recipient cells, while only a slight difference was seen in the y*ku70Δ* yeast strains. In yeast, Rad51 is a homologous protein of RecA and works together with Rad52 in the HR pathway, whereas KU70 is responsible for NHEJ. Results indicated that the process of T-DNA circularization largely depends on the HR pathway in yeast and bacteria, which is the most important DSB repair pathway in these microorganisms. In contrast, end-joining is the dominant DSB repair pathway in plants and the identified patterns of T-DNA junctions resemble those of end-joining events. It would be interesting to investigate the frequency of T-circle formation in plant end-joining mutants. Recently, the DNA polymerase theta (*Pol θ*) was unraveled to mediate random T-DNA integration in plants, whose role is indispensable and essential (van Kregten et al., 2016). It mediates the 3' end capture of T-DNA at genomic breaks for which only one or few bps of micro-homology are required, thus also explaining the frequent presence of filler DNA sequences at the integration sites. As the junctions present in T-circles obtained from plants also contain microhomology and filler

sequences, their formation may have been mediated by *Pol θ*. The analysis of T-circles may thus lead to new insights in the mechanism of T-DNA integration in host cells.

VirD2 may be involved in the formation of T-circles because of its ability to catalyze a reverse transesterification reaction between the free ends of the T-strand and reverse the nicking reaction (Pansegrau et al., 1993). This could explain the precise fusion of the border repeat sequences seen in yeast T-circles and in exceptional plant T-circles. However, if the formation of T-circles is mediated by VirD2 without any help from host factors, T-circles could be generated in *Agrobacterium* cells as well. This hypothesis is not consistent with our finding that no T-circles were detected from donor *Agrobacterium* strains LBA1143 and LBA1144, lacking the T4SS. It confirms the notion that the circularization of T-circles occurs in the host cells and not in *Agrobacterium* and that T-DNA transfer to the host is required (Singer et al., 2012). Considering the high frequency of precise fusions of the two borders occurring in the host (yeast), further research on the involvement of VirD2 in T-DNA integration is needed.

As described in Chapter 2, a relatively large number of T-DNAs were maintained extrachromosomally, thus not integrated into the yeast genome, especially in the *ada2Δ* deletion mutant. Analogously, the frequency with which T-circles were obtained, varied in different plant species. In tobacco *N. tabacum*, the frequency of T-circles among the rescued plasmids ranged from 5% to 10%, while it increased to more than 30% in *N. benthamiana* (Singer et al., 2012). It can be proposed that host factors play an important role in the circularization of T-DNAs. In yeast, *ADA2* may be involved in either the fusion of the border repeats or the maintenance of T-circles. Further research is needed to find out how *ADA2* is involved.

Table 1. Yeast and Agrobacterium strains used in this study

Yeast strain	Genotype	Source/Reference
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Brachmann et al., 1998
BY4741 *ada2Δ*	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ada2::hphMX4*	Chapter 2

Agrobacterium strain	Description	Source/Reference
LBA1100	*Agrobacterium* strain C58 containing pTiB6Δ(ΔT-DNA, Δocc, Δtra), Rif, Spc	Beijersbergen et al., 1992
LBA1143	LBA1100 *virB4Δ*, Rif, Spc, Cb	Beijersbergen et al., 1992
LBA1144	LBA1100 *virB7Δ*, Rif, Spc, Cb	Beijersbergen et al., 1992
LBA1100-pSDM4708	LBA1100 with plasmid pSDM4708	This study
LBA1143-pSDM4708	LBA1143 with plasmid pSDM4708	This study
LBA1144-pSDM4708	LBA1144 with plasmid pSDM4708	This study

Table 2. All the plasmids involved in this study

Plasmid	Specifications	Source/Reference
pUC9	Cloning vector with ampicillin-resistance gene	Vieiria and Messing, 1982
pCAMBIA1390	*Agrobacterium* binary vector with kanamycin-resistance gene	Cambia labs
pSDM8000	*Agrobacterium* binary vector with KanMX selectable marker	van Attikum et al., 2001
pSDM4708	pSDM4712 with pVS1 replication origin	This study
pSDM4710	pBBR6 with T-DNA harboring kanamycin-resistance gene	This study
pSDM4711	T-DNA harboring ampicillin-resistance gene and pUC9 replication origin	This study
pSDM4712	pSDM4711 with kanamycin-resistance gene out of T-DNA region	This study

Table 3. Primers used in this study

Primer name	Sequence (5'-3')	Restriction enzyme
pSDM8000-Fw	AAAGCGGCCGCATCATGAGCGCCAAGCTCAA	*NotI*
pSDM8000-Rev	AAAGGGCCCGGCGCCAGATCTGAGCTTT	*ApaI*
Km-Fw	AAAGCGGCCGCTTGATCCGGCAAACAAACCAC	*NotI*
Km-Rev	AAAGGGCCCCCGCGGTTTCAAAATCGGCT	*ApaI*
pVS1-Fw	AAAGGGCCCTGCTATAGTGCAGTCGGCTT	*ApaI*
pVS1-Rev	AAAGGGCCCACAAGCTGTGACCGTCTCCG	*ApaI*
pA1	ATCGCTGAGATAGGTGCCTCA	
pA2	GATGCTGAAGATCAGTTGGG	
pA3	GAGCGCAACGCAATTAATGTGAGTTA	
pA4	ATATGGTGCACTCTCAGTACAATC	

Restriction sites are underlined.

References

Bakkeren, G., Koukolikova-Nicola, Z., Grimsley, N., & Hohn, B. (1989). Recovery of *Agrobacterium tumefaciens* T-DNA molecules from whole plants early after transfer. *Cell*, 57, 847-857.

Beijersbergen, A., Den Dulk-Ras, A., Schilperoort, R. A., & Hooykaas, P. J. J. (1992). Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science*, 256, 1324-1327.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14, 115-132.

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

Christie, P. J., & Gordon, J. E. (2014). The *Agrobacterium* ti plasmids. *Microbiology spectrum*, 2.

Den Dulk-Ras, A., & Hooykaas, P. J. J. (1995). Electroporation of *Agrobacterium tumefaciens*. In *Plant Cell Electroporation and Electrofusion Protocols*, 63-72. Springer, Totowa, NJ.

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. *The Plant Journal*, 11, 15-29.

De Pater, S., Neuteboom, L. W., Pinas, J. E., Hooykaas, P. J. J., & Van Der Zaal, B. J. (2009). ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. *Plant biotechnology journal*, 7, 821-835.

Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, 11, 355-360.

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and molecular biology reviews*, 67, 16-37.

Gelvin, S. B. (2017). Integration of *Agrobacterium* T-DNA into the plant genome. *Annual review of genetics*, 51, 195-217.

Kononov, M. E., Bassuner, B., & Gelvin, S. B. (1997). Integration of T-DNA binary vector 'backbone'sequences into the tobacco genome: evidence for multiple complex patterns of integration. *The Plant Journal*, 11, 945-957.

Koukolikova-Nicola, Z., Shillito, R. D., Hohn, B., Wang, K., Van Montagu, M., & Zembryski, P. (1985). Involvement of circular intermediates in the transfer of T-DNA from *Agrobacterium tumefaciens* to plant cells. *Nature*, 313, 191-196.

Krizkova, L., & Hrouda, M. (1998). Direct repeats of T-DNA integrated in tobacco chromosome: characterization of junction regions. *The Plant Journal*, 16, 673-680.

Narasimhulu, S. B., Deng, X. B., Sarria, R., & Gelvin, S. B. (1996). Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *The Plant Cell*, 8, 873-886.

Nester, E. W., Gordon, M. P., Amasino, R. M., & Yanofsky, M. F. (1984). Crown gall: a molecular and physiological analysis. *Annual Review of Plant Physiology*, 35, 387-413.

Offringa, R., De Groot, M. J., Haagsman, H. J., Does, M. P., Van Den Elzen, P. J., & Hooykaas, P. J. J. (1990). Extrachromosomal homologous recombination and gene targeting in plant cells after *Agrobacterium* mediated transformation. *The EMBO journal*, 9, 3077-3084.

Ohmine, Y., Satoh, Y., Kiyokawa, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2016). DNA repair genes RAD52 and SRS2, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene ERG28 are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. *BMC microbiology*, 16: 58.

Ohmine, Y., Kiyokawa, K., Yunoki, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2018). Successful Transfer of a Model T-DNA Plasmid to *E. coli* Revealed Its Dependence on Recipient RecA and the Preference of VirD2 Relaxase for Eukaryotes Rather Than Bacteria as Recipients. *Frontiers in Microbiology*, 9, 895.

Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, C., & Bellini, C. (2011). *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. *Physiological and molecular plant pathology*, 76, 76-81.

Rolloos, M., Dohmen, M. H., Hooykaas, P. J. J., & van der Zaal, B. J. (2014). Involvement of Rad 52 in T-DNA circle formation during *Agrobacterium tumefaciens*-mediated transformation of *Saccharomyces cerevisiae. Molecular microbiology*, 91, 1240-1251.

Sambrook, J., & Russell, D. W. (2006). Rapid isolation of yeast DNA. *Cold Spring Harbor Protocols*, 2006, pdb-prot4039.

Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. (2012). Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant physiology*, pp-112.

Soltani, J. (2009) Host genes involved in *Agrobacterium*-mediated transformation. PhD thesis, Institute of Biology, Leiden University, The Netherlands.

Soltani, J., Van Heusden, G. P. H., & Hooykaas, P. J. J. (2009). Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae. FEMS microbiology letters*, 298, 228-233.

Tzfira, T., & Citovsky, V. (2006). *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. *Current opinion in biotechnology*, 17, 147-154.

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

van Kregten, M., de Pater, S., Romeijn, R., van Schendel, R., Hooykaas, P. J. J., & Tijsterman, M. (2016). T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nature plants*, 2, 16164.

Vergunst, A. C., & Hooykaas, P. J. J. (1998). Cre/*lox*-mediated site-specific integration of *Agrobacterium* T-DNA in *Arabidopsis thaliana* by transient expression of *cre. Plant Molecular Biology*, 38, 393-406.

Vieira, J., & Messing, J. (1982). The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic universal primers. *Gene*, 19, 259-268.

Wroblewski, T., Tomczak, A., & Michelmore, R. (2005). Optimization of *Agrobacterium*-mediated transient assays of gene expression in lettuce, tomato and *Arabidopsis. Plant Biotechnology Journal*, 3, 259-273

Chapter 4
Use of the Auxin-induced degron system to study the role of virulence protein VirD2 in the integration of T-DNA into the plant and yeast genome

Shuai Shao, Xiaorong Zhang, G. Paul. H. van Heusden, Paul J. J. Hooykaas

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Abstract

Agrobacterium tumefaciens is the causative agent of crown gall, an important plant disease. It induces tumor formation by transferring a segment of its tumor-inducing plasmid (Ti-plasmid), the T-DNA, to plant cells. The virulence protein VirD2 plays a crucial role in many stages of the transfer process. It is not only essential for the formation of single-stranded T-DNA (T-strand), but also for the translocation of the T-strand through the Type IV secretion system (T4SS) into host cells. Inside host cells it supports the targeting of the T-complex into the nucleus and within the nucleus it may be involved in the integration of the T-DNA into the chromosomal DNA. However, whether VirD2 plays any role within the nucleus is still unclear. In this chapter we exploited the auxin-inducible degron system to investigate the importance of VirD2 within the nuclei of host cells for transformation. The results demonstrate that AMT efficiency is affected by the degradation of VirD2 within the nuclei of both plant and yeast cells. Stable transformation was strongly affected, but transient expression relying on the delivery into the nucleus was much less affected suggesting a role of VirD2 in the targeting to chromatin or in T-DNA integration.

Introduction

Agrobacterium tumefaciens, a gram-negative plant pathogen belonging to the family *Rhizobiaceae*, is the causative agent of crown gall disease, which causes severe damage to worldwide agriculture. It induces tumor formation in plants by transferring a segment of its tumor-inducing plasmid (Ti-plasmid) to plant cells. This transferred DNA (T-DNA) contains genes encoding enzymes involved in the synthesis of auxin, cytokinin and opines resulting in uncontrolled cell proliferation and production of opines. Tumorigenic genes within the T-DNA segment can be deleted, to disarm the plasmid, and be replaced with foreign DNA which then instead becomes integrated into the plant genome as part of the infection process. Under laboratory conditions *A. tumefaciens* is also able to transform other eukaryotes such as yeast and fungi (Bundock et al. 1995; Lacroix et al. 2006). Hence, *A. tumefaciens* has been extensively used as a vector not only to create transgenic plants, but also for fungal transformation and over the past decades *Agrobacterium*-mediated transformation (AMT) has become the preferred method of transformation of many of these organisms (reviewed by Nester et al., 1984; Gelvin, 2003; Tzfira and Citovsky, 2006; Păcurar et al., 2011; Christie and Gordon, 2014; Gelvin, 2017).

AMT can be depicted as a process with multiple stages (reviewed by Gelvin, 2017). First of all, several plant chemicals, such as phenolic compounds along with neutral and acidic sugars, which are produced at the wound sites of plant tissues, trigger a two-component sensory-response system. This system is composed of the *Agrobacterium* VirA and VirG proteins and upon stimulation it will lead to the expression of virulence genes. The virulence protein VirD2, together with VirD1, nicks at 25-bp direct repeat sequences surrounding the T-DNA called left border (LB) and right border (RB) to generate a single stranded copy of this DNA segment (T-strand). VirD2 remains covalently bound to the 5' end of the T-strand at the RB. Simultaneously, a variety of virulence proteins are expressed. The VirB1-11 and VirD4 proteins form a dedicated secretion channel (Type IV secretion system, T4SS) to translocate the T-strand-VirD2 complex along with other effector proteins, including VirE2, VirE3, VirD5 and VirF, to the host cell. Inside the host cell, VirE2 is thought to coat the T-strand to prevent digestion by host nucleases and may target together with VirD2 the T-strand into the nucleus. The VirD5 protein can cause chromosome instability, while VirF may contribute to the removal of VirE2 proteins bound to the T-strand and VirE3 acts as a transcription factor in plant cells. Eventually, the T-DNA will integrate into the host genome mainly through host specific DNA repair mechanisms(reviewed by Tzfira and Citovsky, 2006; Păcurar et al., 2011; Gordon and Christie, 2014; Gelvin, 2017) .

VirD2 plays a crucial role in several steps of the AMT process. The N-terminal region of VirD2 has endonuclease activity, nicking at the borders of the T-DNA to release the T-strand (Yanofsky et al., 1986; Young and Nester, 1988). The T-strand remains covalently bound to VirD2 and VirD2 helps in the translocation of the T-strand through the T4SS into the host cell (Vergunst et al., 2005; van Kregten et al., 2009). The nuclear localization signal sequence (NLS) present in the C-terminal part of VirD2 interacts with importin α and thus contributes to targeting the T-complex to the host nucleus (Howard et al., 1992). It may even be involved in the integration of the T-DNA into the chromosomal DNA. However, the role of VirD2 in the integration of the T-DNA into the host chromosomal DNA remains obscure and in dispute. An aberrant integration pattern caused by mutated VirD2 protein was revealed, thus indicating an

important role of VirD2 in the precise integration of T-DNA into plant genome (Tinland et al., 1995). Later, by a ligation-integration assay *in vitro*, VirD2 was found not to possess general ligase activity, which argued against a function of VirD2 as an integrase and ligase in T-DNA integration (Ziemienowicz et al., 2000). Moreover, plant enzymes were found to mediate T-DNA ligation *in vitro*. VirD2 interacts in plant and yeast cells with proteins important for DNA and histone modification. The S-adenosyl-L-homocysteine hydrolase (involved in DNA methylation) and a MYST-like histone acetyltransferase 2 were reported to interact with VirD2 (Lee et al., 2012). Also it was found that VirD2 can bind to histone proteins in the yeast *Saccharomyces cerevisiae* (Wolterink-van Loo et al., 2015). In addition, VirD2 can mediate ligation of T-DNA border sequences to themselves by a strand transferase reaction, effectively reversing the original nicking reaction (Pansegrau et al., 1993) and thus may contribute to T-DNA circularization in host cells.

To gain more insight into the role of VirD2 in host cells during AMT, conditional inactivation or depletion of VirD2 may be a powerful method. The described auxin-inducible degron (AID) system, using a plant hormone-induced degradation signal, has successfully been used to control protein levels in yeast (Nishimura et al., 2009). In its natural context, auxin (indole-3-acetic acid; IAA) induces degradation of the IAA proteins, a family of short-lived transcriptional repressors, by mediating the interaction of a degron domain in the target protein with the substrate recognition domain of an F-box protein, TIR1, which forms part of a SCF-type ubiquitin ligase (E3). Interaction in the presence of auxin leads to ubiquitylation of the target protein and proteasomal degradation. The SCF complex, consisting of a cullin subunit, a catalytic RING finger protein (RBX1), the adaptor SKP1 and an F-box protein as a substrate recognition subunit, is highly conserved among eukaryotes. Therefore, the plant F-box protein TIR1 is able to form an active E3 complex with the remaining SCF components from other organisms. Hence, constitutive expression of TIR1 allows a reconstitution of the AID system in yeast (Morawska and Ulrich, 2013) or mammalian cells (Holland et al., 2012). Proteins of interest are fused to an auxin-dependent degron sequence derived from IAA17. This AID tag can in principle be placed at the N- or C-terminus of the target protein, thus making the system more flexible than the N-end rule degron. Degradation is reversible and quick, active in minutes rather than in hours. Furthermore, due to the lack of an auxin-responsive system in animals or yeast, the hormone as well as the F-box protein are otherwise biologically silent and cause no measurable physiological changes in the absence of a target, thus minimizing possible side-effects of the treatment (Nishimura et al., 2009).

The TIR1 protein is localized to the nucleus in plant cells (Dharmasiri et al., 2005). This allowed us to exploit the system to specifically target VirD2 for degradation when it entered the nucleus and to study whether this degradation would affect AMT. The results presented here support the idea that VirD2 participates in processes inside the host nucleus that are important for transformation.

Methods

Bacterial cells and growth conditions
Agrobacterium tumefaciens strains used in this chapter are listed in Table 1. Plasmids were transferred into *Agrobacterium* by electroporation (den Dulk-Ras and Hooykaas, 1995). *A. tumefaciens* strains were grown and maintained at 29°C in LC medium (10 g/L tryptone, 5 g/L

yeast extract and 8 g/L NaCl) containing appropriate antibiotics such as carbenicillin (100 µg/mL), kanamycin (100 µg/mL) or rifampicin (20 µg/mL) when required. *Escherichia coli* DH10B was used for plasmid amplification.

Yeast strains and growth conditions

S. cerevisiae strains used in this study can be found in Table 1. DF5-TIR1 *ada2Δ* was constructed using a standard PCR-mediated one-step gene disruption protocol (Kaiser et al., 1994). The disruption cassette was amplified with the primers ADA2-Fw/Rev from plasmid pAG32 (as mentioned in Chapter 2). Yeast was grown at 30°C in yeast extract-peptone-dextrose (YPD) medium supplemented, when required, with the antibiotic G418 (200 µg/mL) or hygromycin (200 µg/mL) or in selective minimal yeast (MY) (Zonneveld, 1986) medium supplemented with appropriate nutrients. Plasmids were transferred to yeast cells using the lithium-acetate transformation protocol (Gietz et al., 1995)

Plasmid constructions

All plasmids and primers used in this chapter are listed in Table 2 and Table 3, respectively. The newly constructed plasmids were checked by restriction digestion and sequence analysis.

Plasmids used in the auxin-induced degron system were constructed as follows. Initially, a fragment with the C-terminal part (AID^{71-114}) of IAA17 was obtained by PCR with the primers AID-Fw/Rev using pHyg-AID^{1-114}-8myc as template. This fragment was digested with *KpnI* and *XmaI* and cloned into pHyg-AID^{1-114}-8myc digested with the same enzymes to replace IAA^{1-114} forming plasmid pSDM4694. Subsequently, the C terminal T4SS translocation signal of protein VirF (VirFCT) was amplified from vector pGPINTAM (Friml et al., 2004) with the primers VirFCT-Fw/Rev and this VirFCT fragment was cloned into the *XmaI* site of pSDM4694 and pSDM4697, resulting in pSDM4694(VirFCT) and pSDM4697(VirFCT), respectively. Finally, two fragments, IAA^{1-114}-VirFCT and AID^{71-114}-VirFCT, were generated by PCR with the primers IAA-Fw/Rev for cloning (see below).

The *virD2* gene, including its promoter region, but lacking its last 3 codons and the stop codon, was amplified by PCR using pSDM3149 as template with the primers VirD2-Fw/Rev and then cloned as *XbaI/EcoRI* fragment into plasmid pSDM3149 to form plasmid pSDM4698 (VirD2-no stop codon). The fragment containing IAA^{1-114}-VirFCT or AID^{71-114}-VirFCT signals were cloned into the *EcoRI* site of pSDM4698, resulting in the plasmids pSDM4699 ($VirD2^{AID}$) and pSDM4700 ($VirD2^{IAA}$), respectively. Simultaneously, in order to monitor the degradation of VirD2 by using microscopy, the same fragments were cloned into the *EcoRI* site of plasmid pUG34-virD2 (pRUL1146) (Soltani, 2009), to form plasmids pSDM4696 ($VirD2^{AID}$) and pSDM4695 ($VirD2^{IAA}$), respectively.

Tumor assay

Four-week old *Nicotiana glauca* and *Kalanchoe tubiflora* were wounded with a sterile syringe at three series sites on the stem. *A. tumefaciens* cells were cultured overnight at 28°C with the appropriate antibiotics and then washed three times and resuspended in 0.9% (w/v) NaCl to an A_{600} of 1.0. Subsequently, *N. glauca* and *K. tubiflora* were injected with 20 µl of *A. tumefaciens* cells at the wound sites and tumors were photographed after 3-4 weeks.

Root transformation and GUS assays

Root transformation was performed as described (Vergunst et al., 2000). Root segments from *Arabidopsis thaliana* were infected with *A. tumefaciens* LBA1100 harboring the binary vector pCambia2301. After co-cultivation on callus induction medium containing 100 μM acetosyingone for 48 hours, root segments were washed, dried and incubated on plates containing shoot induction medium with 30 μg/ml phosphinothricin, 500 μg/ml carbenicillin and 100 μg/ml vancomycin. After 3-4 weeks, plates were photographed and callus formation was scored to calculate transformation efficiencies.

For transient GUS activity assays (Jefferson et al., 1987), *Nicotiana benthamiana* leaves and *A. thaliana* roots were co-cultivated for 3 days and then washed and placed in a 6-well microtiter plate. The wells were sealed with tinfoil. The leaf and root segments were incubated with X-Gluc overnight at 37°C and destained by washing three times with 70% ethanol. After that the GUS staining was examined under a Leica MZ 12 microscope (Leica microsystems). The leaves were photographed with a Leica DC 500 digital camera (Leica microsystems) and the blue spots on the roots were counted.

***Agrobacterium* mediated transformation of yeast**

AMT efficiency was determined as described by Bundock et al. (1995) with some modifications. Firstly, *S. cerevisiae* and *Agrobacterium* strains were grown overnight at 30°C and 28°C, respectively, under continuous shaking and with the appropriate antibiotic selection. The following day, the *Agrobacterium* culture was centrifuged, the cells were washed with induction medium (IM) and re-suspended to an A_{620} of 0.25 in IM with added glucose (10 mM), acetosyringone (AS, 0.2 mM) and appropriate antibiotics, and incubated for another 6 hours at 28°C. Meanwhile, yeast overnight cultures were diluted to an A_{620} of 0.1 and incubated again in either liquid YPD or MY medium for 6 hours. Then, yeast cultures were centrifuged and cells were washed with IM and re-suspended in 0.5 ml of IM to a final A_{620} of 0.4 - 0.6 and mixed by vigorous vortexing with an equal volume of *Agrobacterium* cells. Subsequently, 100 μl of the mixture were pipetted onto sterile nitrocellulose filters laid on IM plates supplemented with histidine, leucine, uracil and methionine. Once the filters were dry, plates were incubated at 21 °C for 6-7 days. After co-cultivation, the cell mixture was washed off the filters and then spread onto YPD plates containing cefotaxime (200 μg/mL) with or without G418 (200 μg/ml). Finally, after a 3-days incubation at 30 °C, colonies were counted. Yeast AMT efficiency was calculated by dividing the number of colonies on the selective plates by the number of colonies on the non-selective plates.

Confocal microscopy and Flow Cytometry

Yeast cells were co-cultivated for 2-3 days with *Agrobacterium*, then washed off the filters with MY medium and an aliquot was analyzed by confocal microscopy using a 63x oil objective on the Zeiss Imager M1 confocal microscope equipped with a LSM5 Exciter. GFP was detected using an argon laser of 488 nm and a band-pass emission filter of 505-600 nm. Images were processed with ImageJ (ImageJ National Institute of Health) (Schindelin et al., 2012).

For flow cytometry all yeast strains were grown in MY medium supplemented with the appropriate nutrients. Cultures were diluted 10-fold before flow cytometry and measured by a Guava easyCyte™ system (Merck Millipore). Data were analyzed with CytoSoft software and in each experiment, 5000 cells were analyzed.

Statistical analysis

All data shown are representative of at least 3 independent experiments of which each contains at least 3 biological replicates and represented as mean of the performed experiments with standard deviation. Statistic tests were done with two-tailed Student's t-test. Statistical analyses were performed using Microsoft Excel. An asterisk indicates a significant differences with a P-value <0.05.

Results

AID degron system allows VirD2 protein degradation in yeast

In order to get a better understanding of the role of VirD2 in the integration process during AMT, we exploited the auxin-inducible degron system to manipulate the level of the VirD2 protein (Figure 1).
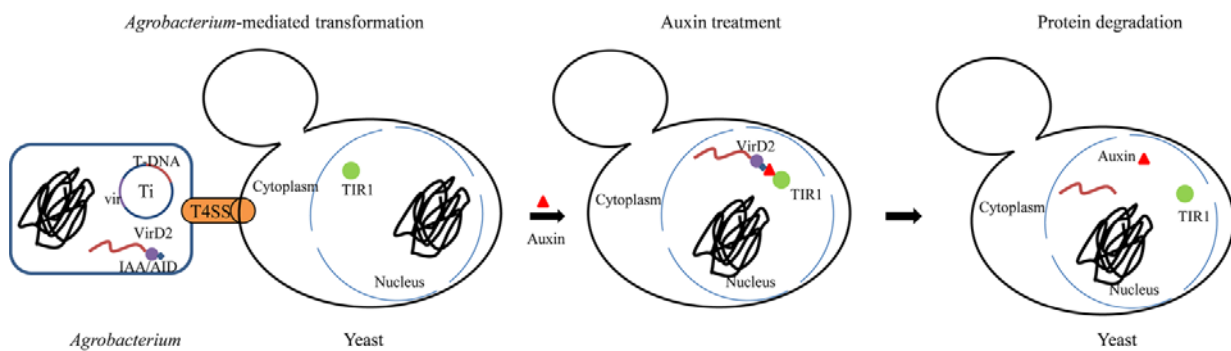

Figure 1. Schematic illustration of the AID system used to study the role of VirD2 inside yeast in AMT. From left to right: the modified VirD2 is bound to the T-strand at its RB and guides the T-strand through the T4SS into the yeast cell; upon translocation of the T-complex into the nucleus addition of auxin promotes the interaction between TIR1 and the AID degron signal (IAA/AID) of the modified VirD2 protein; finally, the modified VirD2 is subject to proteasomal degradation. Purple dot, VirD2; blue diamond, IAA/AID sequence; green dot, TIR1.

This system relies on the F-box protein Tir1, which is located in the nucleus of plant cells (Dharmasiri et al., 2005; Bian et al., 2012). To this end we tagged the VirD2 protein with the IAA17 protein resulting in VirD2^{IAA} containing amino acids 1-114 of IAA17 or with an N-terminal truncation of IAA17 resulting in VirD2^{AID} containing amino acids 71-114 of IAA17 as described (Morawska and Ulrich, 2013). To restore the ability of VirD2 to translocate through the T4SS, we added the T4SS recognition sequence of VirF to the C-terminal ends. For our initial analysis, we tagged both modified VirD2 proteins with GFP (constructs shown at the bottom of Figure 2B and 2C). The fusion proteins were expressed in yeast strain DF5-TIR1 expressing TIR1-9Myc. As in plant cells TIR1 is localized in the nucleus, we assumed that TIR1-9Myc is also located in the nuclei of yeast cells. As demonstrated in Figure 2 (B and C), both GFP-VirD2^{IAA} and GFP-VirD2^{AID} accumulated in structures resembling the nucleus. In control strains expressing free GFP, fluorescence was found all over the cell (Figure 2A). These results are in agreement with previous research showing that ectopically expressed GFP-VirD2 is localized in the nucleus (Wolterink et al, 2015). Upon treatment with Auxin (IAA) the control GFP signals were unaffected (Figure 2A), while both GFP-VirD2^{IAA} and GFP-VirD2^{AID} became

Figure 2 (A and B). Degradation of modified VirD2 protein in yeast can be achieved with the auxin-induced degron system. Yeast cells DF5 stably expressing TIR1-9Myc were transformed with plasmid pUG34 encoding either free GFP (A) and GFP-VirD2^{IAA} (B) then they were treated with (+) or without (−) 1 mM Auxin (Indole-3-acetic acid) in an 1h incubation after cultivation overnight.

A.
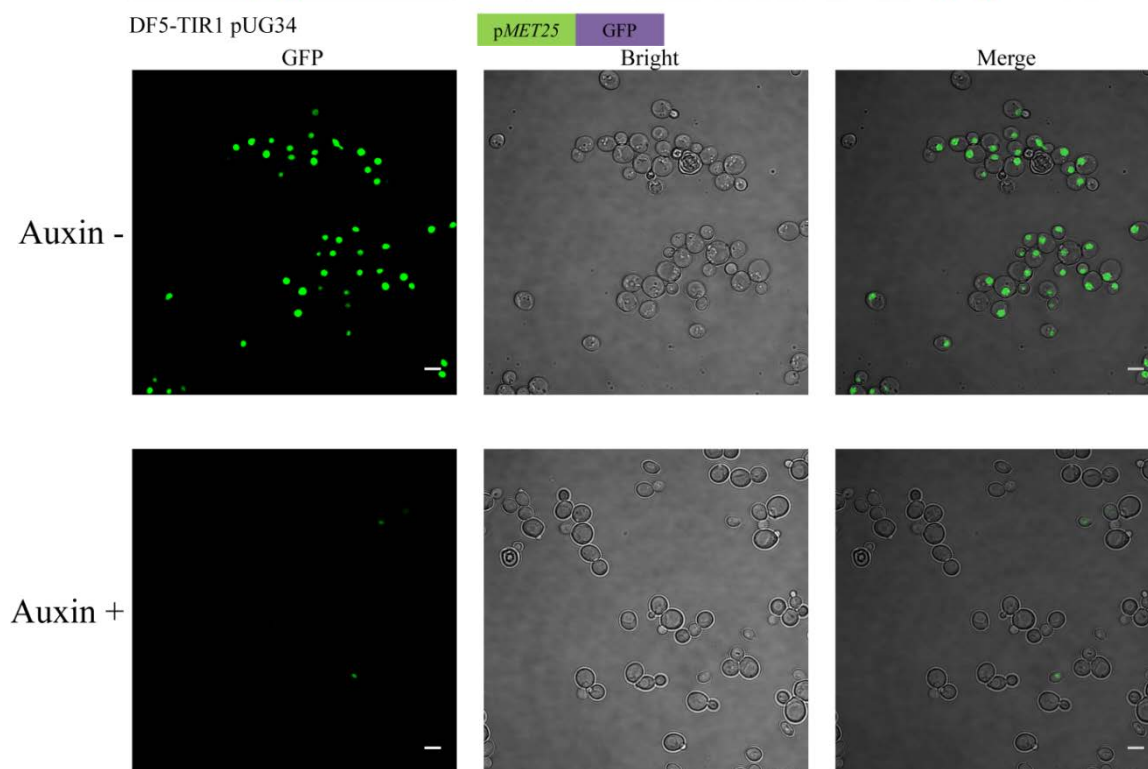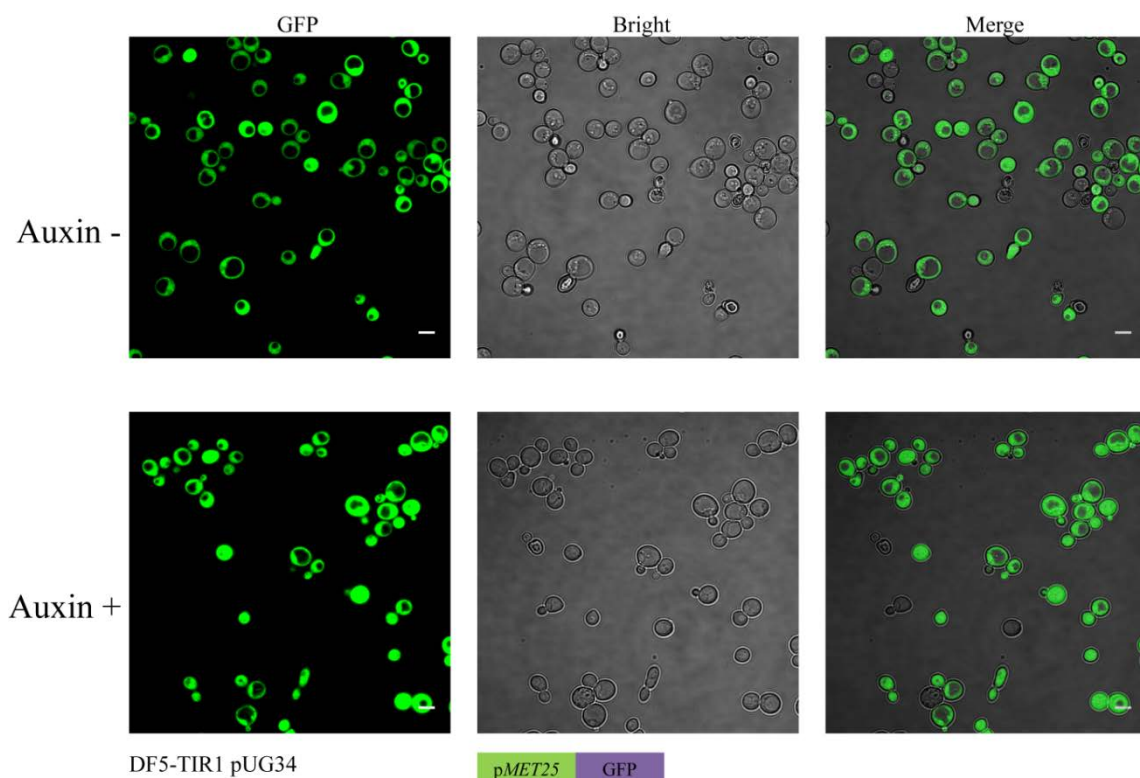
B.

Figure 2 (C and D). Degradation of modified VirD2 protein in yeast can be achieved with the auxin-induced degron system. Yeast cells DF5 stably expressing TIR1-9Myc were transformed with plasmid pUG34 encoding GFP-ViD2^{AID} (C). After cultivation overnight, they were treated with (+) or without (−) 1 mM Auxin (Indole-3-acetic acid) in an 1h incubation and cells were imaged by confocal microscopy and fluorescence was quantified by flow cytometry (D).

C.
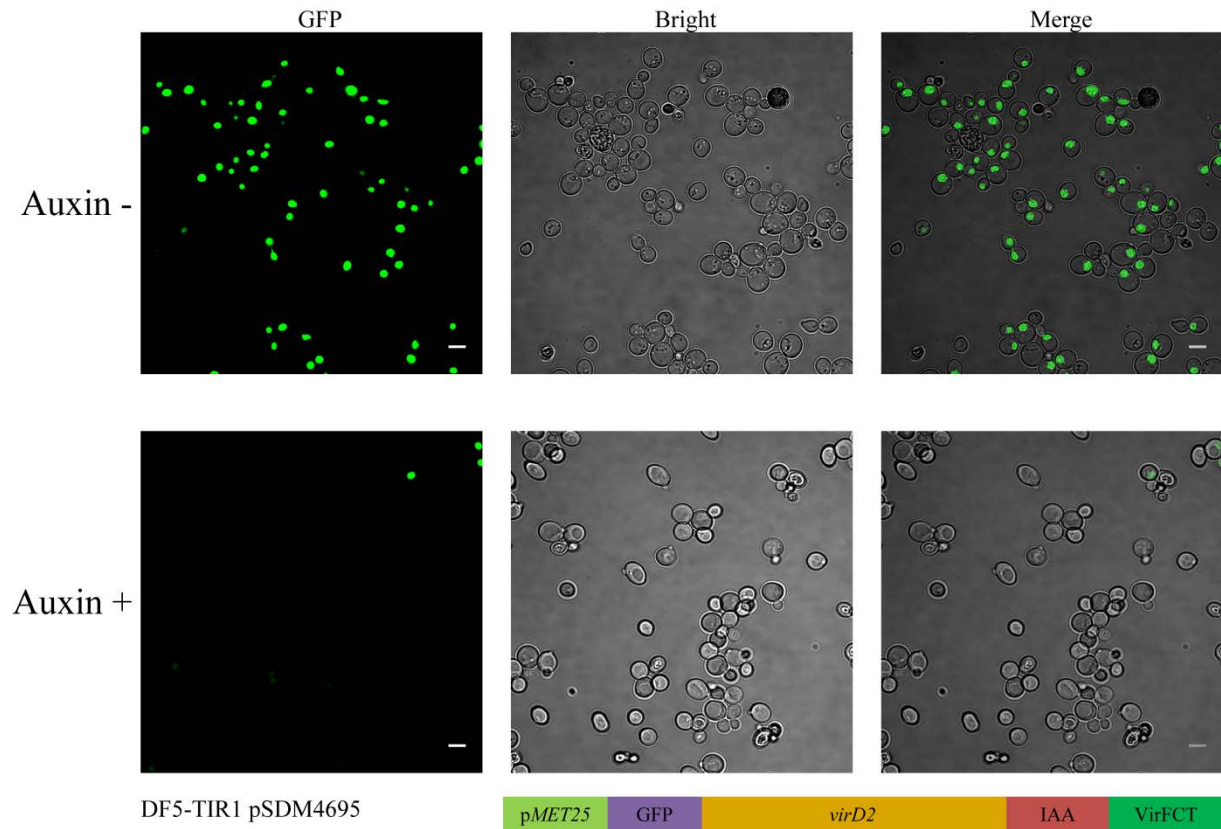
D.
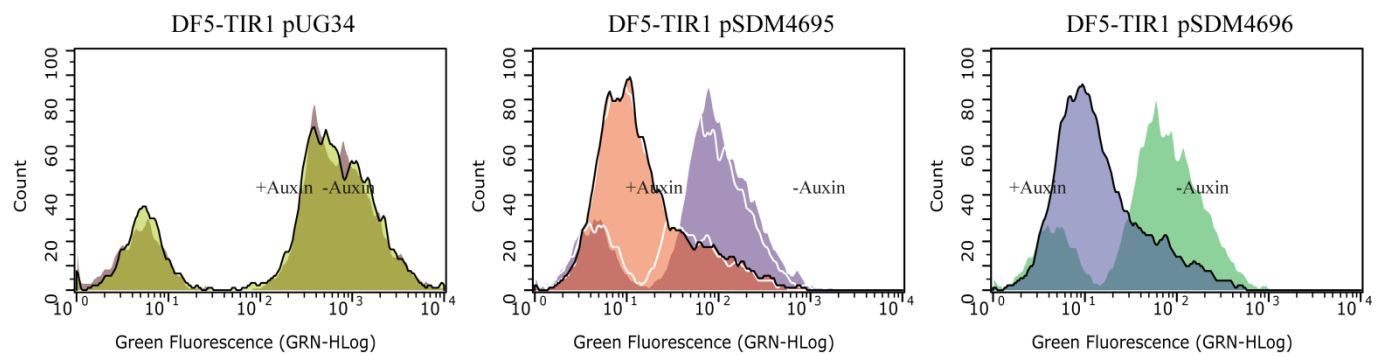
undetectable in most cells by microscopy revealing the presence of an active TIR1-SCF complex in the yeast nucleus (Figure 2B and 2C). The effect of Auxin treatment was further analyzed by flow cytometry. As illustrated in Figure 2D, a significant fluorescent signal could be detected in 95.4% and 95.8% of the cells expressing VirD2^{IAA} or VirD2^{AID}, respectively, but upon auxin treatment it could be detected in only 15.9% and 15.9% percent of the cells, respectively. Average fluorescence intensity was reduced by 72.5% (62.6 to 17.2, arbitrary units) and 67.0% (51.5 to 17.0, arbitrary units) for strains expressing VirD2^{IAA} or VirD2^{AID}, respectively. The levels of free GFP were not affected (Figure 2D). Significant fluorescent signals were detected in around 95% of the cells expressing free GFP when treated or not treated with auxin, while average fluorescence intensities were 268.7 and 272.1 (Arbitrary units), respectively. The flow cytometry experiment was performed twice and in both experiments the same trend was observed. Therefore, it can be concluded that both VirD2^{IAA} and VirD2^{AID} can be degraded by the auxin-inducible degron system within the nucleus. Because of its smaller size VirD2^{AID} will be used in our further research.

Targeted degradation of VirD2 in yeast affects AMT

To investigate the effect of degradation of VirD2 by the AID degron system on AMT, we introduced pBBR6[VirD2] or pBBR6[VirD2^{AID}] expressing native or AID-tagged VirD2, respectively, in the *Agrobacterium virD2* deletion mutant LBA2556. Subsequently, the binary vector pSDM8001 was introduced which allows T-DNA integration into the yeast *PDA1* locus resulting in G418 resistant yeast transformants. The constructed *Agrobacterium* strains were co-cultivated with yeast strain DF5-TIR, in which *TIR1* is expressed continuously and TIR1 binds to the AID-tag to degrade VirD2^{AID} in the presence of auxin. As shown in Figure 3A, in the absence of auxin both *Agrobacterium* strains expressing native VirD2 or AID-tagged VirD2 were able to transform DF5-TIR with a similar efficiency. The efficiency of VirD2^{AID} mutant was slightly lower than that of the VirD2 strain (1.2×10^{-4} vs 0.8×10^{-4}). On the other hand, in the presence of auxin, an approximately 10-fold decrease in the transformation efficiency of the *Agrobacterium* strain expressing VirD2^{AID} was found (0.1×10^{-4}), which was not observed for the *Agrobacterium* strain expressing native VirD2 (1.0×10^{-4}).

Considering the increased AMT efficiency of the *ada2Δ* mutant described in Chapter 2, we tested the possibility whether the enhanced AMT efficiency can be suppressed by the degradation of VirD2. To this end an *ada2Δ* deletion mutant was constructed in the DF5-TIR1 genetic background. This strain and the parent strain DF5-TIR1 were co-cultivated with *A. tumefaciens* strains expressing either VirD2 or VirD2^{AID}. As shown in Figure 3B (left part), in the presence of auxin the *ada2Δ* mutant is transformed with an higher efficiency (6.6×10^{-4}) compared to the parental strain DF5-TIR1 (3.7×10^{-4}) by the *Agrobacterium* strain expressing native VirD2 as observed previously in other strain backgrounds (Chapter 2). On the other hand, in the presence of auxin the transformation efficiencies by *Agrobacterium* expressing VirD2^{AID} were strongly reduced for both DF5-TIR1 and DF5-TIR1 *ada2Δ* strains (1.9×10^{-5} and 2.2×10^{-5}, respectively) (Figure 3B, right part). These results indicate that degradation of VirD2 also strongly reduces AMT efficiency of the *ada2Δ* strain.
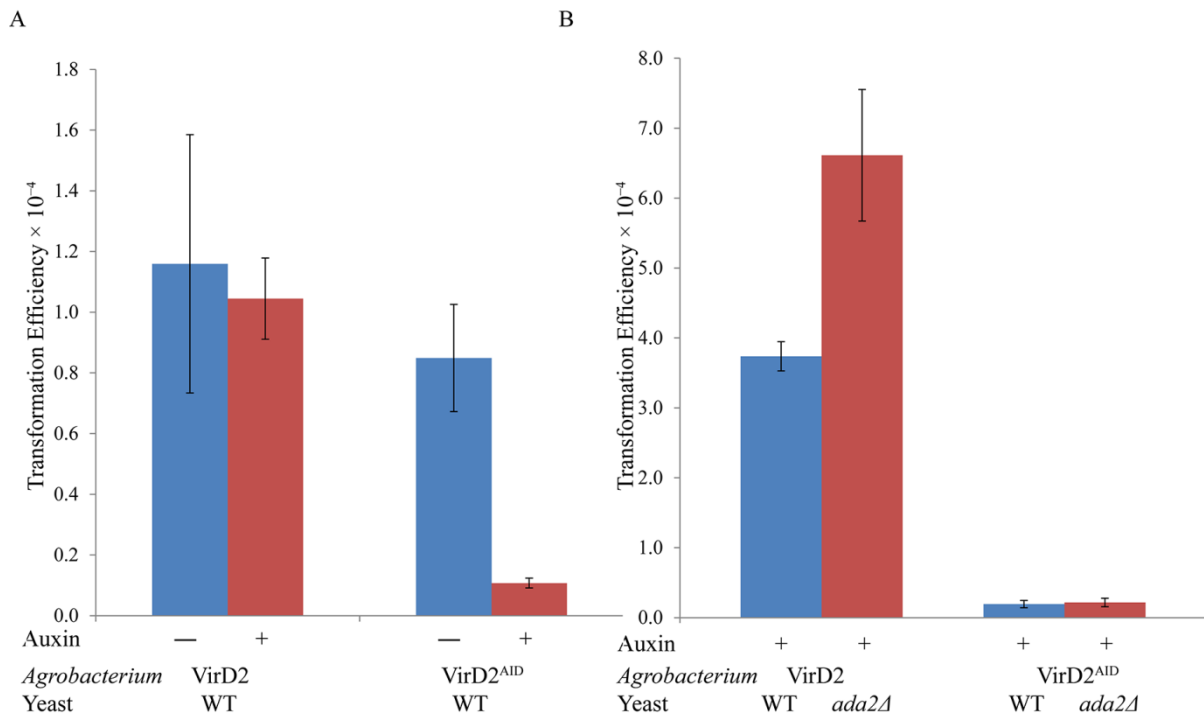
Figure 3. AID-mediated degradation of VirD2 affects AMT efficiency of yeast. (A) DF5-TIR1 yeast cells were co-cultivated for 7 days with *Agrobacterium* strain LBA2556-pSDM3149-pSDM8001 expressing native VirD2 or with LBA2556-pSDM4699-pSDM8001 expressing VirD2^{AID}. The induction medium was either supplemented or not supplemented with Auxin (Indole-3-acetic acid). Transformation efficiency was determined after selection in the presence or absence of G418. Error bars indicate the standard deviations of triplicate independent assays. (B) DF5-TIR1 and its *ada2Δ* mutant were co-cultivated in the presence of auxin with *Agrobacterium* strain LBA1100-pSDM8001 expressing native VirD2 or with LBA2556-pSDM4699-pSDM8001 expressing VirD2^{AID} for 7 days. Error bars indicate the standard deviations of triplicate independent assays.

Targeted degradation of VirD2 affects AMT of plants

In order to test the effect of the AID tag on tumor formation, *virD2^{AID}* , *virD2^{IAA}* and native *virD2* were cloned into the expression vector pBBR6, introduced in the *A. tumefaciens virD2* deletion mutant LBA2569 and tumor formation by these strains on *N. glauca* and *K. tubiflora* was analyzed. The wild type *A. tumefaciens* strain LBA1010 was included as a positive control. It is well known that auxin is present in all parts of the plant, although in different concentrations; therefore, addition of auxin is probably not necessary for VirD2^{AID} degradation in plant cells. As demonstrated in Figure 4A, the tumors on *N. glauca* caused by *Agrobacterium* strains LBA2569 expressing VirD2^{IAA} or VirD2^{AID} were much smaller than those caused by the wild type strain LBA1010, indicating an important role of VirD2 within the plant cell nucleus in transformation. In a second experiment tumor formation on both *N. glauca* and *K. tubiflora* was analyzed. As shown in Figure 4B (upper part) on both plant species tumors formed by *Agrobacterium* strain LBA2569 expressing VirD2^{AID} were smaller than tumors formed by LBA2569 expressing native VirD2. Injection of 100-fold less *Agrobacterium* cells resulted in smaller tumors for all strains (Figure 4B, lower part). The tumors formed after injection with *Agrobacterium* strain LBA2569 expressing VirD2^{AID} at OD$_{600}$ = 1 were similar in size to those formed by the *Agrobacterium*

strain expressing native VirD2, added at OD_{600} = 0.01. These results indicate that degradation of VirD2 within the plant cell nuclei strongly reduces the transformation efficiency.
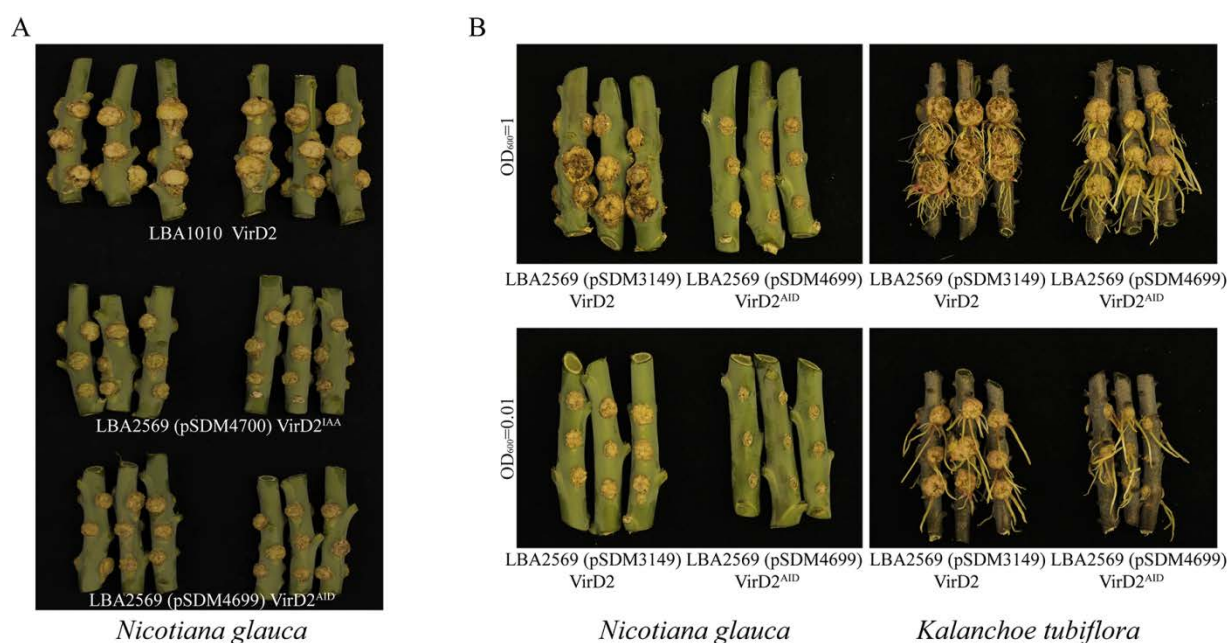

Figure 4. Tumor formation assay on plants. (A) *N. glauca* was infected with the *Agrobacterium* wild type strain LBA1010 and its *virD2Δ* derivative LBA2569 harboring plasmid pSDM4699 or pSDM4700 expressing VirD2^{AID} or VirD2^{IAA}, respectively. Tumors were photographed after 3 weeks. (B) *N. glauca* and *K. tubiflora* were injected with different initial concentrations (OD_{600}=1 or OD_{600}=0.01) of *Agrobacterium* strains LBA2569 harboring plasmid pSDM3149 or pSDM4699 expressing VirD2^{AID} or native VirD2, respectively. Tumors were photographed after 3 weeks.

In order to further investigate the effect of degradation of VirD2 by the AID system, we performed a quantitative transformation assay on *A. thaliana* roots. To this end, the binary vector pCAMBIA2301 containing the kanamycin resistance gene and the *gus* reporter gene was introduced into *A. tumefaciens* strain LBA1100 expressing native VirD2 and into the LBA2556 expressing VirD2^{AID}. The formation of green calli was analyzed after four weeks (Figure 5). The number of calli formed by the strain expressing VirD2^{AID} was only 27% of the number of calli formed by the strain expressing native VirD2. This observation is in line with the results of the tumor formation assay and indicates that degradation of VirD2 within the nucleus of plant cells strongly and negatively affects the efficiency of AMT in plants.

T-DNA transfer is not affected by the degradation of VirD2
As shown in the previous section, degradation of VirD2 by the AID system, resulted in a lower transformation frequency in both yeast and plants. In order to investigate the effect of tagging VirD2 with AID on the transfer of the T-DNA from *Agrobacterium* into the host nuclei, we performed a transient transformation assay on *A. thaliana* roots using the same constructs as were used in the root transformation assay. Upon transfer of T-DNA from pCAMBIA2301 into the plant cells, GUS activity can be detected independently of and prior to integration into one of the chromosomes.
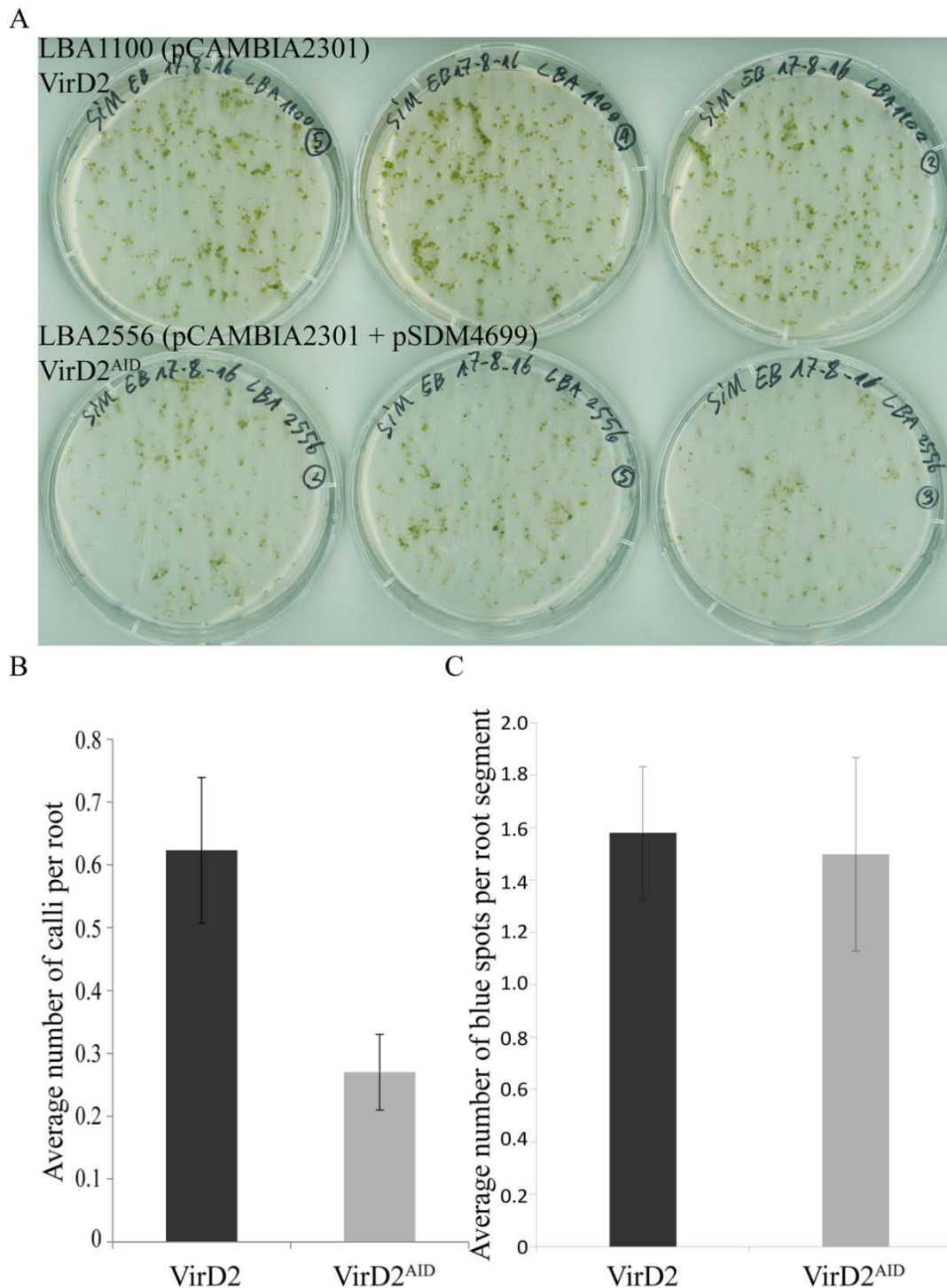
Figure 5. *Arabidopsis* **root transformation assay to study the effect of AID-mediated degradation of VirD2 .** *Arabidopsis* root segments were co-cultivated with *Agrobacterium* strain LBA1100-pCAMBIA2301 expressing native VirD2 or LBA2556-pSDM4699-pCAMBIA2301 expressing VirD2^{AID}. (A and B) Root segments were transferred to shoot induction medium containing kanamycin for selection of T-DNA integration. Root transformation was scored after 4 weeks. Efficiency is presented by the ratio of the average number of green calli per root segment relative to the total number of root segments. (C) Root segments were stained with X-Gluc. Blue spots of co-cultivated *A. thaliana* roots were counted and statistical analysis demonstrated no significant differences. Error bars indicate the standard deviations of triplicate independent assays.

After co-cultivation with *A. tumefaciens* strains for 3 days, the root segments were stained with X-Gluc to reveal transient transformation and the blue spots were quantified. As illustrated in Figure 5C no significant differences were observed in the numbers of blue spots seen after transfer from strains expressing VirD2 or VirD2^{AID}. This indicates that VirD2^{AID} is as active as native VirD2 in transferring T-DNA into the nuclei of host cells, and also that conversion of the T-strand in a double stranded form is not affected by degradation of VirD2.

Discussion

A. tumefaciens induces tumor formation at wound sites in plants by transferring T-DNA into plant cells. VirD2 plays a key role in *Agrobacterium*-mediated T-DNA transfer; it participates in the entire transfer process, from the formation of single-stranded T-strand to the translocation of the T-complex through the T4SS into host cells. Inside the host cell it interacts with importin, which then mediates translocation of the T-complex into the nucleus and it may even be involved in the integration of the T-DNA into the chromosomal DNA. However, a putative role of VirD2 in the integration of the T-DNA into the host chromosomal DNA has remained obscure. In order to study the role of VirD2 inside the host cell nucleus, we exploited the auxin-induced degron (AID) system to destroy VirD2 once it enters the host cell nucleus. To this end, we constructed *Agrobacterium* strains expressing AID-tagged and native VirD2 and studied the effect of VirD2 degradation in the nucleus during the transformation of yeast and plants.

To test whether the AID-degron system could lead to degradation of VirD2 within the nucleus, we expressed an AID-tagged VirD2-GFP fusion protein in yeast strain DF5-TIR, in which the TIR1 F-box protein was expressed allowing auxin-induced activation of the SCF complex. As shown by confocal microscopy (Figure 2) the fluorescence of GFP-tagged VirD2 proteins largely disappeared upon auxin treatment in most cells. Also the fluorescence detected by flow cytometry was greatly reduced, indicating that VirD2 was degraded within the nucleus. On the other hand free GFP was not affected by the auxin treatment. This indicates that in yeast the AID system can be used to strongly reduce the levels of intracellular VirD2. Subsequently, we showed that the efficiency of AMT of yeast by an *Agrobacterium* strain expressing VirD2^{AID} was 10-fold reduced upon auxin-induced degradation of VirD2^{AID} (Figure 3), thus confirming that VirD2 plays a role in the transfer process inside the yeast cell.

With respect to plant cells, the size of the tumors induced by *Agrobacterium* strains expressing VirD2^{AID} was reduced compared to that of the tumors caused by strains expressing native VirD2. Besides, with the root transformation assay the number of calli formed by the strain expressing VirD2^{AID} was only 27% of the number of calli formed by the strain expressing native VirD2 (Figure 5B). These results confirm an important role of VirD2 in the transformation process inside plant cells. To investigate whether this lower efficiency is related to a reduced delivery of tagged VirD2 from *Agrobacterium* into plant cell nuclei, transient transformation assays using *Agrobacterium* strains carrying the binary vector pCAMBIA2301 were performed. The translocation of the T-DNA from *Agrobacterium* to the host cell nucleus was not affected by the AID tag on VirD2 (Figure 5C). Therefore, we can speculate that the observed lower stable transformation efficiency of strains with VirD2^{AID} is caused by degradation of VirD2^{AID} inside the nuclei of host cells rather than by a deficient translocation of the T-DNA from its entry point into the host cell to the cellular nucleus.

It has been shown that VirD2 can interact with histones in yeast (Wolterink-van Loo et al., 2015). It this way it may help direct the T-DNA to the chromatin as a prelude to integration into the host chromosomal DNA. In plants, it is possible that VirD2 interacts with histone proteins or other chromatin factors to drive T-DNA to plant chromatin regions (Gelvin and Kim, 2007), facilitating the integration process. In agreement with this, the local chromatin structure may affect AMT. Overexpression of specific histone proteins increased AMT efficiency in plants (Mysore et al., 2000; Tenea et al., 2009) and deletion of histone-related genes altered AMT efficiency in yeast (Soltani et al., 2009). Taken together, we can speculate that VirD2 can interact with host factors to participate in the integration process as well.

In conclusion, using the degron approach developed in this chapter we found that the efficiency of AMT goes strikingly down after degradation of VirD2 even though the T-DNA was delivered successfully into the nucleus of the host cell. Although direct evidence for a role of VirD2 in the integration process is still lacking, it can be concluded that VirD2 has a role inside the nucleus of the host cell to promote stable transformation. We have to be careful with this conclusion though, because we have not added auxin in the plant experiments and our results may therefore have been influenced by local differences in auxin concentration. Another pitfall, which we have not addressed so far, may be that the SCF-TIR1 complex may be functional to some extent also in the cytoplasm. Further experimentation will have to be done to address these issues.

Acknowledgements

Table 1. Yeast and *Agrobacterium* strains used in this study

Yeast strain	Genotype	Source/Reference
BY4741	MATa *his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	Brachmann et al., 1998
BY4741 *ada2Δ*	BY4741 *ada2::hphMX4*	Chapter 2
DF5	MATa *his3Δ200 leu2Δ3,2Δ112 lys2Δ801 trp1Δ1 ura3Δ52*	Finley et al., 1987
DF5-TIR1	DF5 URA3:: ADH1-AtTIR19myc	Morawska and Ulrich, 2013
DF5-TIR1-pUG34	TIR1 with plasmid pUG34	This study
DF5-TIR1-pSDM4695	TIR1 with plasmid pSDM4695	This study
DF5-TIR1-pSDM4696	TIR1 with plasmid pSDM4696	This study
DF5-TIR1 *ada2Δ*	DF5-TIR1 *ada2::hphMX4*	This study

Agrobacterium strain	Description	Source/Reference
LBA1100	*Agrobacterium* strain C58 cured, pTiB6, *ΔT-DNA, Δocc, Δtra*, Rif, Spc	Beijersbergen et al., 1992
LBA1010	*Agrobacterium* strain C58 cured, pTiB6, Rif	Koekman et al., 1982
LBA2556	LBA1100 *virD2Δ*, Rif, Spc	Jurado-Jacome, 2011
LBA2569	*Agrobacterium* strain LBA1010 *virD2Δ*, Rif, Spc	van Kregten, 2011
LBA1100-pSDM8001	LBA1100 with binary vector pSDM8001, Km	van Attikum and Hooykaas, 2003
LBA1100-pCAMBIA2301	LBA1100 with binary vector pCAMBIA2301, Km	This study
LBA2556-pSDM3149-pSDM8001	LBA2556 with binary vector pSDM3149 and pSDM8001	This study

LBA2556-pSDM3149-pCAMBIA2301	LBA2556 with binary vector pSDM3149 and pCAMBIA2301	This study
LBA2556-pSDM4699-pSDM8001	LBA2556 with binary vector pSDM4699 and pSDM8001	This study
LBA2556-pSDM4699-pCAMBIA2301	LBA2556 with binary vector pSDM4699 and pCAMBIA2301	This study

Table 2. Plasmids used in this study

Plasmid	Specifications	Source/Reference
pSDM3149	pRL662, pvirD-virD2	Jurado-Jacome, 2011
pHyg-IAA$^{1-114}$-8myc	N-terminal tagging signal of IAA17 in parent vector pSM409	Morawska and Ulrich, 2013
pSDM4694	pHyg-AID$^{71-114}$-8myc	This study
pSDM4697(VirFCT)	pHyg-IAA$^{1-114}$-VirFCT-8myc with C-terminal of VirF	This study
pSDM4694(VirFCT)	pHyg-AID$^{71-114}$-VirFCT-8myc with C-terminal of VirF	This study
pSDM4698	pRL662, pvirD-virD2 without stop codon	This study
pSDM4699	pRL662, pvirD-VirD2^{AID}	This study
pSDM4700	pRL662, pvirD-VirD2^{IAA}	This study
pUG34	Centromeric plasmid with a *HIS3* marker to express N-terminal GFP fusions under the control of the *MET25* promoter	Güldener et al., 1996
pRUL1146	pUG34-virD2	Soltani, 2009
pSDM4695	pUG34-VirD2^{IAA}	This study
pSDM4696	pUG34-VirD2^{AID}	This study
pSDM8001	*Agrobacterium* binary vector with KanMX selectable marker and *PDA1* flanking sequences	van Attikum and Hooykaas, 2003
pCAMBIA2301	*Agrobacterium* binary vector with NeoR/KanR selectable marker and GUS genes	Hajdukiewicz et al., 1994
pGPINTAM	Binary vector with a tamoxifen-inducible promoter	Friml et al., 2004

Table 3. Primers used in this study

Primer name	Sequence (5'-3')	Restriction enzyme
VirFCT-Fw	CCCC<u>CCCGGG</u><u>CTCGAG</u>GTTATGGCAGAAGTTC	*XmaI-XhoI*
VirFCT-Rev	CCCC<u>CCCGGG</u>GACGAACAGCACGGATAGTC	*XmaI*
AID-Fw	TT<u>GGTACC</u>CCTAAAGATCCAGCCAAACC	*KpnI*
AID-Rev	CCCC<u>CCCGGG</u>TGATACCTTCACGAACGCCGC	*XmaI*
VirD2-Fw	GCGGCCGC<u>TCTAGA</u>GGATCC	*XbaI*
VirD2-Rev	CG<u>GAATTC</u>GCGCCCATCGTCGCGACGAT	*EcoRI*
IAA-Fw	CG<u>GAATTC</u>CGTACGCTGCAGGTCGAC	*EcoRI*
IAA-Rev	CG<u>GAATTC</u>GACGAACAGCACGGATAG	*EcoRI*
ADA2-Fw	TAAAATATCAGCGTAGTCTGAAAATATATACATTAAG CAAAAAGACAGCTGAAGCTTCGTACGC	
ADA2-Rev	ATAATAACTAGTGACAATTGTAGTTACTTTTCAATTTT TTTTTTGCCGCGGCCGCATAGGCCAC	

Restriction sites are underlined.

References

Beijersbergen, A., Den Dulk-Ras, A., Schilperoort, R. A., & Hooykaas, P. J. J. (1992). Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science*, 256, 1324-1327.

Bian, H., Xie, Y., Guo, F., Han, N., Ma, S., Zeng, Z., Wang, J., Yang, Y., & Zhu, M. (2012). Distinctive expression patterns and roles of the miRNA393/TIR1 homolog module in regulating flag leaf inclination and primary and crown root growth in rice (*Oryza sativa*). *New Phytologist*, 196, 149-161.

Brachmann, C. B., Davies, A., Cost, G. J., Caputo, E., Li, J., Hieter, P., & Boeke, J. D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. *Yeast*, 14, 115-132.

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

Christie, P. J., & Gordon, J. E. (2014). The *Agrobacterium* Ti plasmids. *Microbiology spectrum*, 2.

den Dulk-Ras, A., & Hooykaas, P. J. J. (1995). Electroporation of *Agrobacterium tumefaciens*. In *Plant Cell Electroporation and Electrofusion Protocols*, 63-72. Springer, Totowa, NJ.

De Pater, B. S., Neuteboom, L. W., Pinas, J. E., Hooykaas, P. J. J., & Van Der Zaal, B. J. (2009). ZFN-induced mutagenesis and gene-targeting in *Arabidopsis* through *Agrobacterium*-mediated floral dip transformation. *Plant biotechnology journal*, 7, 821-835.

Dharmasiri, N., Dharmasiri, S., Weijers, D., Lechner, E., Yamada, M., Hobbie, L., Ehrismann, J. S., Jurgens, G., & Estelle, M. (2005). Plant development is regulated by a family of auxin receptor F box proteins. *Developmental cell*, 9, 109-119.

Finley, D., Özkaynak, E., & Varshavsky, A. (1987). The yeast polyubiquitin gene is essential for resistance to high temperatures, starvation, and other stresses. *Cell*, 48, 1035-1046.

Friml, J., Yang, X., Michniewicz, M., Weijers, D., Quint, A., Tietz, O., Benjamins, R., Ouwerkerk, P. B. F., Ljung, K., Sandberg, G., Hooykaas, P. J. J., Palme, K., & Offringa, R. (2004). A PINOID-dependent binary switch in apical-basal PIN polar targeting directs auxin efflux. *Science*, 306, 862-865.

Gietz, R. D., Schiestl, R. H., Willems, A. R., & Woods, R. A. (1995). Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. *Yeast*, 11, 355-360.

Gelvin, S. B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and molecular biology reviews*, 67, 16-37.

Gelvin, S. B., & Kim, S. I. (2007). Effect of chromatin upon Agrobacterium T-DNA integration and transgene expression. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1769, 410-421.

Gelvin, S. B. (2017). Integration of *Agrobacterium* T-DNA into the plant genome. *Annual review of genetics*, 51, 195-217.

Güldener, U., Heck, S., Fiedler, T., Beinhauer, J., & Hegemann, J. H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic acids research*, 24, 2519-2524.

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

Holland, A. J., Fachinetti, D., Han, J. S., & Cleveland, D. W. (2012). Inducible, reversible system for the rapid and complete degradation of proteins in mammalian cells. *Proceedings of the National Academy of Sciences*, 109, E3350-E3357.

Howard, E. A., Zupan, J. R., Citovsky, V., & Zambryski, P. C. (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.

Jefferson, R. A., Kavanagh, T. A., & Bevan, M. W. (1987). GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The EMBO journal*, 6, 3901-3907.

Jurado-Jacome, E. (2011) *Agrobacterium* infection: translocation of virulence proteins and role of VirF in host cells. PhD thesis, Institute of Biology, Leiden University, The Netherlands.

Kaiser, C., Michaelis, S., & Mitchell, A. (1994). Methods in yeast genetics. *Cold Spring Harbor:* Cold Spring Harbor Laboratory Press.

Koekman, B. P., Hooykaas, P. J.J., & Schilperoort, R. A. (1982) A functional map of the replicator region of the octopine Ti plasmid. *Plasmid*, 7, 119–132

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

Lee, L. Y., Wu, F. H., Hsu, C. T., Shen, S. C., Yeh, H. Y., Liao, D. C., Fang, M. J., Liu, N. T., Yen, Y. C., Dokladal, L., Sýkorová, E., Gelvin, S. B., & Lin, C. S. (2012). Screening a cDNA library for protein–protein interactions directly in planta. *The Plant Cell*, 24, 1746-1759..

Morawska, M., & Ulrich, H. D. (2013). An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast*, 30, 341-351.

Mysore, K. S., Nam, J., & Gelvin, S. B. (2000). An *Arabidopsis* histone H2A mutant is deficient in *Agrobacterium* T-DNA integration. *Proceedings of the National Academy of Sciences*, 97, 948-953.

Nester, E. W., Gordon, M. P., Amasino, R. M., & Yanofsky, M. F. (1984). Crown gall: a molecular and physiological analysis. *Annual Review of Plant Physiology*, 35, 387-413.

Nishimura, K., Fukagawa, T., Takisawa, H., Kakimoto, T., & Kanemaki, M. (2009). An auxin-based degron system for the rapid depletion of proteins in nonplant cells. *Nature methods*, 6, 917.

Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, C., & Bellini, C. (2011). *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. *Physiological and molecular plant pathology*, 76, 76-81.

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. *Proceedings of the National Academy of Sciences*, 90,11538–42.

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., & Tinevez, J. Y. (2012). Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9, 676.

Singer, K. (2018). The Mechanism of T-DNA Integration: Some Major Unresolved Questions. *Agrobacterium* Biology: From Basic Science to Biotechnology, 287-317.

Soltani, J. (2009) Host genes involved in *Agrobacterium*-mediated transformation. PhD thesis, Institute of Biology, Leiden University, The Netherlands.

Soltani, J., Van Heusden, G. P. H., & Hooykaas, P. J. J. (2009). Deletion of host histone acetyltransferases and deacetylases strongly affects *Agrobacterium*-mediated transformation of *Saccharomyces cerevisiae*. *FEMS microbiology letters*, 298, 228-233.

Tanaka, S., Miyazawa-Onami, M., Iida, T., & Araki, H. (2015). iAID: an improved auxin-inducible degron system for the construction of a 'tight'conditional mutant in the budding yeast *Saccharomyces cerevisiae*. *Yeast*, 32, 567-581.

Tenea, G. N., Spantzel, J., Lee, L. Y., Zhu, Y., Lin, K., Johnson, S. J., & Gelvin, S. B. (2009). Overexpression of several *Arabidopsis* histone genes increases *Agrobacterium*-mediated transformation and transgene expression in plants. *The Plant Cell*, 21, 3350-3367.

Tinland, B., Schoumacher, F., Gloeckler, V., Bravo‐Angel, A. M., & Hohn, B. (1995). The *Agrobacterium tumefaciens* virulence D2 protein is responsible for precise integration of T-DNA into the plant genome. *The EMBO journal*, 14, 3585-3595.

Tzfira, T., & Citovsky, V. (2006). *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. *Current opinion in biotechnology*, 17, 147-154.

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

van Attikum, H., & Hooykaas, P. J. J. (2003). Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in *Saccharomyces cerevisiae*. *Nucleic acids research*, 31, 826-832.

van Kregten, M., Lindhout, B. I., Hooykaas, P. J. J., & van der Zaal, B. J. (2009). *Agrobacterium*-mediated T-DNA transfer and integration by minimal VirD2 consisting of the relaxase domain and a type IV secretion system translocation signal. *Molecular plant-microbe interactions*, 22, 1356-1365.

van Kregten. (2011). VirD2 of *Agrobacterium tumefaciens*: functional domains and biotechnological applications. Ph.D. thesis, Leiden University, Leiden, The Netherlands.

van Kregten, M., de Pater, S., Romeijn, R., van Schendel, R., Hooykaas, P. J. J., & Tijsterman, M. (2016). T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nature plants*, 2, 16164.

Vergunst, A. C., Schrammeijer, B., den Dulk-Ras, A., de Vlaam, C. M., Regensburg-Tuïnk, T. J., & Hooykaas, P. J. J. (2000). VirB/D4-dependent protein translocation from *Agrobacterium* into plant cells. *Science*, 290, 979-982.

Vergunst, A. C., van Lier, M. C., den Dulk-Ras, A., Stüve, T. A. G., Ouwehand, A., & Hooykaas, P. J. J. (2005). Positive charge is an important feature of the C-terminal transport signal of the VirB/D4-translocated proteins of *Agrobacterium*. *Proceedings of the National Academy of Sciences*, 102, 832-837.

Wolterink-van Loo, S., Ayala, A. A. E., Hooykaas, P. J. J., & van Heusden, G. P. H. (2015). Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. *Microbiology*, 161, 401-410.

Yanofsky, M. F., Porter, S. G., Young, C., Albright, L. M., Gordon, M. P., & Nester, E. W. (1986). The virD operon of *Agrobacterium tumefaciens* encodes a site-specific endonuclease. *Cell*, 47, 471-477.

Young, C., & Nester, E. W. (1988). Association of the virD2 protein with the 5'end of T strands in Agrobacterium tumefaciens. Journal of bacteriology, 170, 3367-3374.

Ziemienowicz, A., Tinland, B., Bryant, J., Gloeckler, V., & Hohn, B. (2000). Plant enzymes but not *Agrobacterium*VirD2 mediate T-DNA ligation *in vitro*. *Molecular and cellular biology*, 20, 6317-6322.

Zonneveld, B. J. (1986). Cheap and simple yeast media. *Journal of microbiological methods*, 4, 287-291.

Chapter 5
Complete sequence of the tumor-inducing plasmid pTiChry5 from the hypervirulent *Agrobacterium tumefaciens* strain Chry5

Shuai Shao, Xiaorong Zhang, G. Paul H. van Heusden and Paul J. J. Hooykaas

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Abstract

Agrobacterium tumefaciens strain Chry5 is hypervirulent on many plants including soybean that are poorly transformed by other *A. tumefaciens* strains. Therefore, it is considered as a preferred vector for genetic transformation of plants. Here we report the complete nucleotide sequence of its chrysopine-type Ti-plasmid pTiChry5. It is comprised of 197,268 bp with an overall GC content of 54.5%. Two T-DNA regions are present and 219 putative protein-coding sequences could be identified in pTiChry5. Roughly one half of the plasmid is highly similar to the agropine-type Ti plasmid pTiBo542, including the virulence genes with an identical *virG* gene, which is responsible for the supervirulence caused by pTiBo542. The remaining part of pTiChry5 is less related to that of pTiBo542 and embraces the *trb* operon of conjugation genes, genes involved in the catabolism of Amadori opines and the gene for chrysopine synthase, which replaces the gene for agropine synthase in pTiBo542. With the exception of an insertion of IS869, these Ti plasmids differ completely in the set of transposable elements present, reflecting a different evolutionary history from a common ancestor.

Introduction

Agrobacterium tumefaciens, a gram-negative and rod-shaped plant pathogen belonging to the family *Rhizobiaceae*, is the causative agent of crown gall disease. It induces tumor formation in plants by transferring a segment of its tumor-inducing plasmid (Ti-plasmid) to plant cells. This transferred DNA (T-DNA) contains genes involved in the synthesis of plant growth regulators and thus causes uncontrolled cell proliferation at infection sites (For review see: Nester et al., 1984; Gelvin, 2003; Tzfira and Citovsky, 2006; Păcurar et al., 2011; Gordon and Christie, 2014). Under laboratory conditions *Agrobacterium* is also able to transform other eukaryotes such as yeast and fungi (Bundock et al., 1995; De Groot et al., 1998). Hence, *A. tumefaciens* and its Ti plasmid have been extensively used as a vector to create transgenic plants and fungi and *Agrobacterium*-mediated transformation (AMT) has become the preferred method of transformation of these organisms over the past decades.

Many different strains of *Agrobacterium* have been isolated from nature and they may differ in their virulence on specific plants. They are often classified on the basis of the opines, tumor-specific metabolites, that are found in the tumors and serve as an energy source for the bacterium (Dessaux et al., 1993). The *A. tumefaciens* strain Chry5, which was originally isolated from chrysanthemum, turned out to have hypervirulence on soybean, which is poorly transformed by other *A. tumefaciens* strains (Bush and Pueppke, 1991). Tumors induced by Chry5 contain a novel opine called chrysopine (Chilton et al., 1995). So far pTiChry5 has only been characterized by restriction endonuclease analysis and partial sequencing (Kovács and Pueppke, 1994; Palanichelvam et al., 2000). Here, we present the complete sequence of pTiChry5.

Materials and methods

After cultivation of strain Chry5, total DNA was extracted (DNeasy Blood & Tissue Kit, QIAGEN) and two libraries were constructed using the HiSeq SBS kit v4 (Illumina). Paired-end 125 cycles sequence reads were generated using the Illumina HiSeq2500 system by BaseClear (Leiden, the Netherlands). After trimming, 4,193,789 paired-end reads (527,368,966 bases) were retained with an average length of 125 bp. Assembly was performed using the CLC Genomics Workbench software (version 7.0.4) and 31 contigs were obtained with a length ranging from 742 bp to 624 kb. Subsequently the Blast 2.2.31+ algorithm (NCBI) was used to map all contigs to the reference sequences, nopaline-type pTiC58 (AE007871) (Wood et al., 2001) and octopine-type pTiAch5 (CP007228) (Henkel et al., 2014) and contigs with high hits were identified. The relative order and relationship of these contigs was subsequently determined by a series of PCRs and small gaps were filled in by sequencing the PCR-generated fragments. The assembled sequence was annotated using the IGS Prokaryotic Annotation Pipeline (Galens et al., 2011) and the Rapid Annotations using Subsystems Technology (RAST) server (Overbeek et al., 2014) with manual modification. We found the following sequencing data deposited in GenBank under accession numbers AF065246.1 GI: 4883694 (4,168bp *traR* and *virH* genes), U88627.1 GI: 1881616 (1,000 bp 6b gene), AF229156.1 GI: 13377005 (3,797 bp *acs* gene), AF065242.2 GI: 12831440 (38,235 bp T_R DNA and Amadori catabolism genes), AF229155.1 GI: 13377004 (1,375 bp T_L border). We have compared our sequence with the previously established sequences and found single nucleotide differences in specific regions, but not evenly distributed over the entire length of the sequence. Similarly, larger differences

including a larger deletion are present in the regions showing multiple nucleotide differences. Therefore, we have carefully rechecked our reads. Although we cannot exclude that some of the differences may be snp's due to maintenance of the strains over time in different labs, it is more likely that the many differences that are located in specific areas are due to sequencing errors. Oger and Farrand proposed several gene names for the genes located in the area which they sequenced (GenBank AF065242.2 GI: 12831440), which we have mostly adopted and used below.

Results and discussion

General features of plasmid pTiChry5

The complete sequence of pTiChry5 was found to comprise 197,268 bp and the overall GC content was 54.5%. In total, 219 putative open reading frames larger than 100 bp were found with an average size of 724 bp (Table 1 and Figure 1). No genes encoding transfer-RNA (tRNA), or ribosomal RNA (rRNA), or genomic islands (GI) were found on pTiChry5 as is the case for other Ti plasmids.

An average nucleotide identity (ANI) analysis was carried out to describe the genetic relatedness with other Ti-plasmids using the Perl script (Konstantinidis and Tiedje, 2005). We extracted all known complete Ti-plasmid sequences from public databases, and calculated ANI values between pTiChry5 and these plasmids (Table 2). The results showed that pTiChry5 shared the highest similarity (97%) with agropine-type pTiBo542 (NC_010929; Oger et al., 2001) within the conserved regions. Similarity with octopine-type plasmids was less (92%) and with nopaline-type Ti-plasmids was even lower (84%). Global plasmid-wide sequence alignment between pTiChry5 and pTiBo542 using GCview showed that the plasmids share large homologous regions in the same order (Figure 1). Therefore, in the following sections we shall focus on a comparison between these two Ti plasmids. A large segment of pTiChry5 stretching from the replication units, over the virulence region, the *tra* operon, the area involved in agrocinopine catabolism up to the left part of the T_L-DNA is most highly conserved with pTiBo542. The remaining parts covering the genes dealing with degradation of the Amadori opines and with conjugation (the *trb* operon) are less or not conserved and probably have a different origin. The presence of two *repABC* replication units in both pTiChry5 and pTiBo542 also suggests that both these plasmids have originated from recombination reactions between at least two different *repABC* plasmids. In their evolution, plasmids pTiChry5 and pTiBo542 have been interrupted by different transposition events. One of these, IS869, is present at exactly the same position in both plasmids, suggesting that its insertion already occurred in the ancestor of both plasmids. Transposable elements IS292, IS1131, IS1312, and IS1313, which are present at different positions in pTiBo542 (Deng et al., 1995a; Ponsonnet et al., 1995; Wabiko, 1992), are all lacking in pTiChry5. Instead, next to the right border repeat of the T_L-DNA of pTiChry5 (see below) a copy of IS66 (96% identity) was inserted and next to it a copy of IS868 (91% identity). The IS66 element was previously discovered in other Ti plasmids (Machida et al., 1984) and IS868 is known from pTiAB3 (Paulus et al., 1991). Adjacent to the right border repeat of the T_R-DNA (see below) copies of ISRel26 (González et al., 2006) (87% identity) and ISAtu3 (81% identity) were found. ISRel26 was identified previously in *Rhizobium etli* plasmid p42a and ISAtu3 is known from accessory plasmid pAtC58 (Wood et al., 2001). Besides, only two other transposable elements were found, a copy of IS21 in the area

between Rep and the virulence region and IS869 just upstream of the *virB* operon (Paulus et al., 1991).
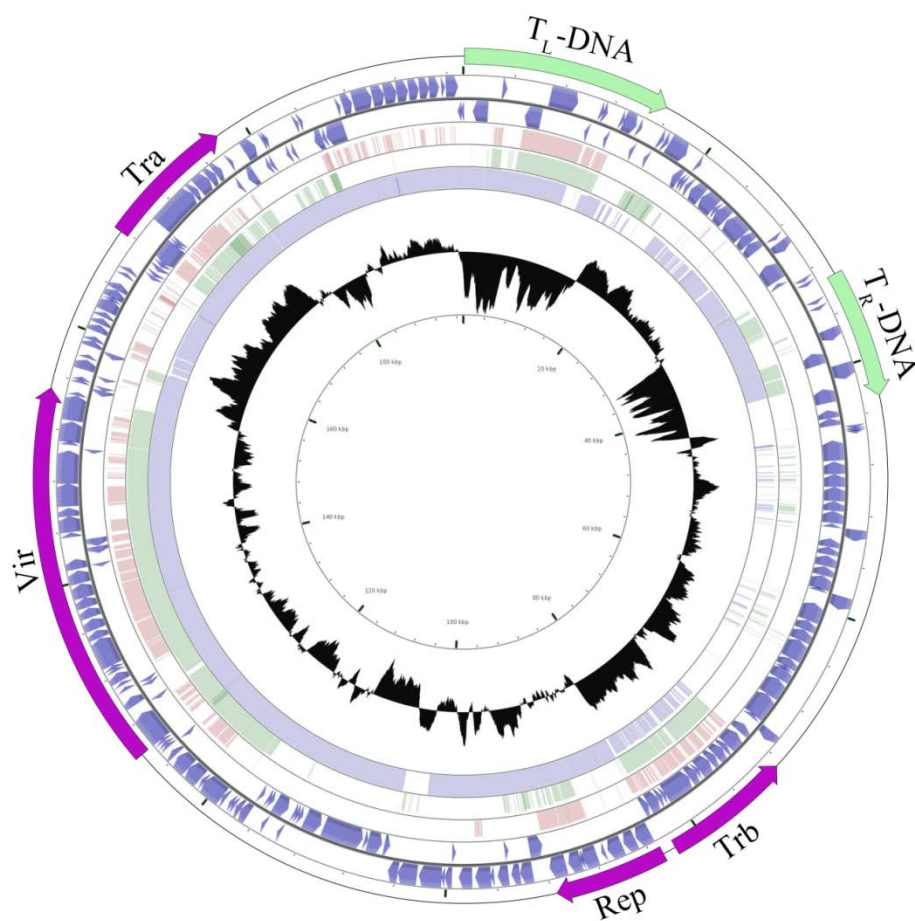
Figure 1. Schematic circular map of Ti-plasmid pTiChry5. Circle ranges from 1 (outer circle) to 6 (inner circle). Circle 1, location of T-DNA, *tra*, *trb*, *vir* and *rep* gene clusters. Circles 2 and 3, predicted open reading frames on the plus and minus strand, respectively. Circles 4, 5 and 6, coordinated mapping of pTiChry5 against Ti-plasmids pTiC58 (red), pTiAch5 (green) and pTiBo542 (light blue) respectively. The single nucleotide based similarities were denoted by color on the circles from blank to fully filled representing 0-100% similarity. Circle 7, GC content percentages. Sequence comparisons and designing the figure was performed using the CGView program (Stothard and Wishart. 2005).

Table 1. General features of Ti-plasmid pTiChry5.

Feature	pTiChry5
Size	197,268 bp
GC Content	54.5%
Protein coding regions	219
Hypothetical	50
Average ORF size	724bp
T-region	2 (15,632bp and 9,626bp)
GC % of T-regions	45.0% and 44.5%
rRNA	-
tRNA	-
IS elements	6

Table 2. Average nucleotide identity (ANI) values of sequenced Ti-plasmids.

ANI (%)	pTiSAKURA	pTiC58	pTiAch5	pTi15955	pTiBO542	pTiChry5
pTiSAKURA	-					
pTiC58	97.99	-				
pTiAch5	82.36	82.14	-			
pTi15955	82.24	82.24	99.82	-		
pTiBo542	84.92	84.32	91.35	91.22	-	
pTiChry5	84.01	84.06	92.04	92.23	97.48	-

T-DNA regions

Like octopine-type and agropine-type Ti-plasmids, pTiChry5 contains two T-DNA regions indicated as T_L-DNA (15,632 bp) and T_R-DNA (9,626 bp). Both T-DNAs are surrounded by border repeats that match the previously determined border sequences described by Palanichelvam et al. (2000). The GC content of the T-DNA regions is 45.0% and 44.5%, respectively, which is significantly lower than that of the rest of the Ti-plasmid. The T_L-DNA containing the onc-genes and the genes for l,l-succinamopine/leucinopine (*les*) and agrocinopine (*acs*) synthesis, is very similar to that of pTiBo542, but the pTiChry5 T_L-DNA harbors gene *6a*, which is not present in the pTiBo542 T_L-DNA (Figure S1). Also within the pTiBo542 T_L-DNA an IS1312-like element is inserted near the left border repeat, but this is not the case in pTiChry5. It has been speculated that this IS1312 insertion containing a "pseudo-border" sequence may cause the transfer of a truncated T-DNA lacking gene5 into plants (Deng et al., 1995b). As gene5 encodes an enzyme which can form the inactive auxin analog indole 3-lactate which may compete with auxin production (Korber et al., 1991), the absence of gene5 in tumors may be the cause of the necrosis seen on tumors induced by Bo542 (Deng et al., 1995b). With regard to T_R-DNA, most genes are well conserved between pTiChry5 and pTiBo542 except the *ags* gene (Figure S2). This gene is located adjacent to the T_R-DNA right border repeat of pTiBo542 and is involved in agropine synthesis (Figure S2). Previously, it was found that in Chry5 tumors chrysopine is produced instead of agropine (Chilton et al., 1995) and that the locus responsible for chrysopine biosynthesis is located adjacent to the right border repeat on the T_R-DNA (Palanichelvam et al., 2000). Instead of the *ags* gene, the T_R-DNA of pTiChry5 harbours the *chsA* gene responsible for chrysopine synthase (Palanichelvam et al., 2000). This gene (AgrTiChry5_50) is distantly related to the *ags* gene (36% identity), which is expected as both encode enzymes carrying out a similar lactonization reaction. It is remarkable that the pTiChry5 T_R-DNA contains not only a gene *chsC* homologous to *mas2*, which mediates formation of deoxy-fructosyl-glutamate (dfg), the direct precursor for lactonizing to chrysopine by chrysopine synthase, but also a gene *chsB* homologous to *mas1*, which encodes the enzyme responsible for the reduction of dfg into mannopine (Ellis et al., 1984). Evidently, tumors induced by Chry5 contain dfg and chrysopine, but no mannopine (Chilton et al., 1995). It is likely therefore that the gene product encoded by pTiChry5 *chsB* can no longer carry out this reduction reaction. Indeed, although the N-terminal part of *chsB* is highly conserved, the C-terminal part, which contains the reductase motifs in *mas1*, is very different from *mas1* (Figure S13).

Plasmid backbone

For further analysis of the evolutionary relationship between pTiChry5 and other Ti-plasmids, we compared their *trb*, *tra*, and *rep* gene clusters at the nucleic acid level by using ProgressiveMauve (Darling et al., 2010) (Figure 2A). With regard to the conjugation genes (Figure S5, S10), we found that the similarity of the *tra* operon with that of pTiBo542 (99% similarity) is much stronger than that of the *trb* operon (82% similarity), reflecting a likely different origin of these two operons in pTiChry5. Indeed, reversely, the similarity of the *tra* operon with that of the octopine Ti plasmid pTi15955 (similarity 81%) is weaker than that of the *trb* operon (88% similarity).
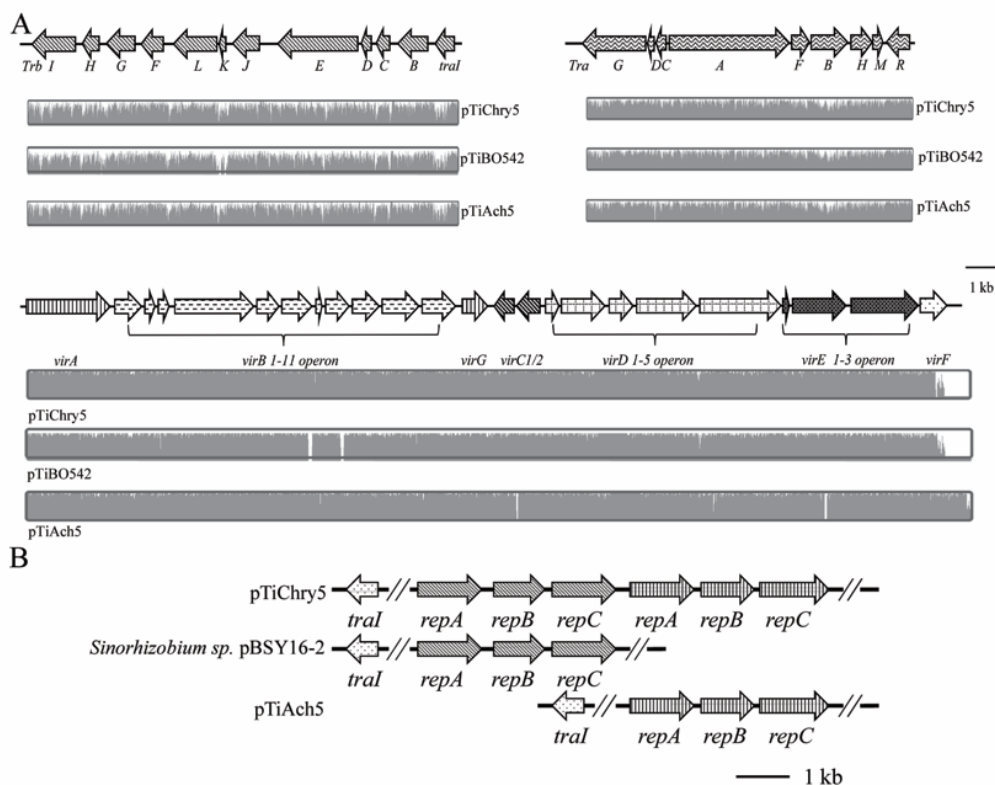

Figure 2. Comparison of the *trb*, *tra*, *vir* and *rep* gene clusters of pTiChry5 with those of other Ti-plasmids. A, Comparison of the *trb*, *tra* and *vir* gene clusters of pTiChry5, pTiBo542 and pTiAch5. The arrows indicate the genes included in these operons. ProgressiveMauve was used to generate pairwise alignments with default parameters and the single nucleotide based similarities are expressed as height of the panels from blank to filled to represent 0-100% similarity. **B,** Fine structure of the *repABC* operons of pTiChry5, pBSY16-2 (CP016451) and pTiAch5. Locations of open reading frames are shown by arrows filled with different patterns and the same pattern represents homologous genes.

Similar to pTiBo542, we found two adjacent copies of the *repABC* operon in pTiChry5 (Figure 2B, S6). The *repABC* operon (*repA* AgrTiChry5_98, *repB* AgrTiChry5_100, *repC* AgrTiChry5_101) is not only almost identical to *repABC*2 of pTiBo542, but also shares high similarity (88% with pTiC58 and 80% with pTiAch5 at the amino acid levels) with *repABC* of incRh1 Ti-plasmids. It remains to be determined whether this operon like that of pTiBo542 (Yamamoto et al., 2017) has lost its role in replication, but has only a function in

incompatibility and maybe plasmid maintenance. The other *repABC* operon, consisting of *repA* (AgrTiChry5_94), *repB* (AgrTiChry5_95) and *repC* (AgrTiChry5_96), is almost identical to *repABC1* of pTiBo542, but is only distantly related to the *repABC* operons of the *incRh1* Ti plasmids (only 40% similarity at the amino acid level). Blasting this rare *repABC* operon with the NCBI non-redundant database revealed that a very similar *repABC* operon (*repA* 94%, *repB* 89% and *repC* 98%) can be found on the accessory plasmid pBSY16-2 from the nitrogen-fixing bacterium *Sinorhizobium* sp. RAC02, (CP016451). A large set of genes of unknown function shared with pTiBo542 is located between the replication region and the virulence region (Figure S6, S7), and between the virulence region and the *tra*-operon (figure S9).

Virulence genes

As to the virulence genes (Figure S8, S10), most of the *vir* genes of pTiChry5 are highly similar to those of pTiBo542 (99% similarity) and to those of pTiAch5 (97% similarity). The *virF*, *virP*, *virQ* and *virR* genes are separated from the other *vir* genes by a large set of genes of unknown function and the *tra* genes. The *virF* gene, which enhances the virulence of octopine type strains (Melchers, et al., 1990), is replaced in pTiChry5 by a gene which shares only weak homology with *virF* of pTiAch5. This gene encodes a protein that still has the features of an F-box domain (Schrammeijer et al., 2001) and is also present in pTiBo542 (Figure S12). A previous study has shown that a small fragment containing *virG* is responsible for the enhanced virulence caused by pTiBo542 (Jin et al., 1987). The encoded VirG protein of pTiChry5 is completely identical to that of pTiBo542 and differs from octopine-type pTiAch5 VirG protein in two amino acid substitutions (at residue 7 and 106 in VirG of pTiAch5). Thus it is likely that the hypervirulence of Chry5 is at least partially due to the presence of this *virG* allele.

Opine catabolism genes

A large set of genes of Ti plasmids is involved in the catabolism of opines. We found major differences between these genes present on pTiBo542 and on pTiChry5 in line with the known catabolic properties of the host bacteria (Bush and Pueppke, 1991; Vaudequin-Dransart et al., 1995). The two Ti plasmids share a large area located between the T_L-DNA and the T_R-DNA containing genes involved in the uptake and catabolism of L,L succinamopine and leucinopine (Figure S3). Genes involved in the transport and catabolism of agrocinopines are located adjacent to the left border of the T_L-DNA, separated from this by an IS1131 element in pTiBo542, but not in pTiChry5 (Figure S11). We identified the complete *accR* regulator, which was shown to control not only the induction of the *acc*-genes by agrocinopines C and D, but also the truncated *arc*-operon comprising the *traR* regulator of conjugation (Oger and Farrand, 2001). A major difference between pTiBo542 and pTiChry5 can be seen in the segment adjacent to the T_R-DNA right border containing genes involved in the transport and catabolism of the Amadori opines (Figure S4). The pTiBo542 genes located here are involved in transport and catabolism of mannopine, mannopinic acid, agropine and agropinic acid. Genes with homology to mannopine and mannopinic acid transport and catabolism are not present in pTiChry5, which indeed does not confer catabolism of these compounds on it hosts (Chilton et al., 1995; Vaudequin-Dransart et al., 1995). However, although pTiChry5 neither confers catabolism of agropine and agropinic acid on its hosts, genes 68% similar to the genes involved in agropine transport, agropine catabolism and agropinic acid transport are nevertheless present in pTiChry5. As Chry5 does not catabolize agropine, we suspect that the four "agropine

transport" genes and the three "agropine catabolism" genes in pTiChry5 are in fact involved in the transport and catabolism of chrysopine, which is structurally related. An insertion mutation in *Bam*H1 fragment 18 of pTiChry5 was previously shown to inactivate the capacity to degrade chrysopine (Vaudequin-Dransart et al., 1998). From the sequence we can see that this fragment contains the genes AgrChry5_56 -59 with homology to the agropine transport genes *agtA-agtD*. This indicates that genes AgrChry5_56 -59 determine a chrysopine transport system. However, we find a loss of the stop codon of the *agtC* homolog and also a loss of the start codon of the *agtD* homolog. Apparently, these genes have fused in pTiChry5 and we propose to call these three genes *chtA-chtC*. The gene *chcA* with strong homology to the *agcA* gene, which encodes the agropine delactonase, probably encodes the delactonase which converts chrysopine into deoxyfructosylglutamine (dfg). Dfg in turn can be broken down into a sugar and an amino acid by enzymes encoded by *mocE* and *mocD* (Kim and Farrand, 1996). In pTiChry5 homologous genes are present, which were called *chcE* and *chcD* by Oger and Farrand (Genbank AF065242). A *mocC* gene encoding an oxido-reductase is absent as predicted before (Baek et al., 2005) and explains why Chry5, in contrast to Bo542, cannot catabolize mannopine. A regulator ChcR with high similarity to MocR, is encoded by an adjacent open reading frame. Next are five genes with some similarity to *socA*, *socB* and *socD*, which are known from the pAtC58 plasmid and involved in the uptake and catabolism of dfg, also called santhopine (Baek et al., 2003). Tumors formed by Chry5 contain not only chrysopine, but also the Amadori compounds dfg and deoxyfructosylproline (dfop), which can be catabolized by Chry5 (Chilton et al., 1995; Vaudequin-Dransart et al., 1995). However, transfer of a closely related chrysopine Ti plasmid into C58 cured of both its plasmids did not confer the ability to degrade dfg on the recipients, but this required the co-transfer of a megaplasmid indicating that the "*soc*" -genes present on pTiChry5 may not (all) be functional. In pAtC58 an uptake system for dfg is encoded by only two genes, *socA* and *socB*. While *socA* encodes a periplasmic substrate-binding protein, the *socB* gene encodes the different parts of the membrane transporter itself (Baek et al., 2003). Intriguingly in pTiChry5 the *socA*-like gene is surrounded by three genes that share homology with parts of *socB*. It needs to be tested whether these genes together still encode a functional membrane transporter. These putative "*soc*"- transport genes are followed by a catabolic gene with similarity to *socD* encoding an amadoriase. A *socC* homolog is not present and this may also be a reason that the pTiChry5 plasmid genes are not sufficient for dfg catabolism. The *socD* gene is followed by a regulatory gene, which may control the "*soc*"- genes. On the other side of this regulator a large operon is located consisting of 8 genes. At the 3'end of the operon four genes (AgrChry5_71 –74) are present with high similarity to the agropinic acid transport genes *agaA-agaD*. As Chry5 does not catabolize agropinic acid, these genes, named *dfpA-dfpD* by Oger and Farrand (Genbank AF065242), may encode a transport system for the uptake of deoxyfructosylproline (dfop), which is a lactam opine, which has structural sinilarity to agropinic acid. The four catabolic genes (AgrChry5_75 -78), which are present at the 5'end of the operon have similarity to *mocE, mocD* and to *hyuA* and *hyuB* genes encoding hydantoinases and oxoprolinases. The proteins encoded by the latter two genes may be involved in the opening of the oxoproline ring of dfop, converting it back into dfg, which can be subsequently degraded by enzymes encoded by the *mocD-* and *mocE* -like genes, which we propose to name *dfpH* and *dfpI* following the proposed nomenclature of *dfpF* and *dfpG* for the *hyu* homologs. In octopine Ti plasmids two genes with sinilarity to *hyuA* and *hyuB* were

previously identified and called *agaF* and *agfG* as they were involved in the catabolism of agropinic acid, which also requires lactam ring opening (Lyi et al., 1999). Some evidence for an involvement of this "dfop-operon" in the catabolism of dfop could be found in Vaudequin-Dransart et al.(1998) who reported that an insertion into the *Bam*HI fragment 8 of pTiChry5 inactivated the ability to degrade dfop. From the DNA sequence we can see that this fragment contains the genes AgrChry5_74 -79. Although this would suggest that pTiChry5 should confer not only growth on dfop, but also on dfg, this is evidently not the case. It may be that the *dfp* operon only becomes active in the presence of dfop and not by the presence of dfg, thus precluding the catabolism of dfg in the absence of dfop. Similarly, the *mocD* and *mocE* genes, which are involved in the catabolism of chrysopine and which should also allow breakdown of dfg, may need induction by chrysopine, precluding its degradation in the absence of chrysopine. It may be beneficial for the bacteria to activate these Ti plasmid-encoded opine degradation systems only in tumors when there is an abundance of opines and to prevent activation in soil, where dfg may be present. Separate dfg degradation genes, the *soc* genes, are present on non-Ti plasmids in strains such as C58 and Chry5 (Baek et al., 2005; Vaudequin-Dransart et al., 1998). These are used for degradation of dfg from soil and possibly also dfg from tumors. In pTiChry5 such genes became integrated into the Ti plasmid, but as explained above it seems that they became (partially) inactive so that also pTiChry5 requires the *soc* genes on an accessory plasmid to scavenge dfg from soil.

Conclusions

We report here the complete nucleotide sequence of pTiChry5 and the detailed alignments between pTiChry5 and pTiBo542 in order to more comprehensively understand their evolutionary relationship. As illustrated in the Figures 1 and S1-13, in general these two plasmids have a similar structure and order of functional areas with different interruptions by transposable elements and losses or gains of small sets of genes. Nevertheless, based on similarity, unexpectedly roughly one half of these plasmids (*rep*, *vir*, *tra*, *acc*, T_L-DNA) is much more similar than the other half (*trb*, Amadori opine catabolism up to the chrysopine synthase gene), suggesting that these plasmids are chimaeric due to recombination with other related plasmids. Presence of two *repABC* units in both these plasmids is in line with this hypothesis. Especially, the development of new opine profiles may have conferred evolutionary advantage on their host bacteria in some specific environments. The complete sequence presented here will be helpful in further detailed analysis of pTiChry5 and its development into a more efficient transfer vector.

Nucleotide sequence accession numbers

The complete sequence of pTiChry5 has been deposited at GenBank under the accession number KX388536 and the strain Chry5 has been deposited in the collection of the Institute of Biology, Leiden University, the Netherlands.
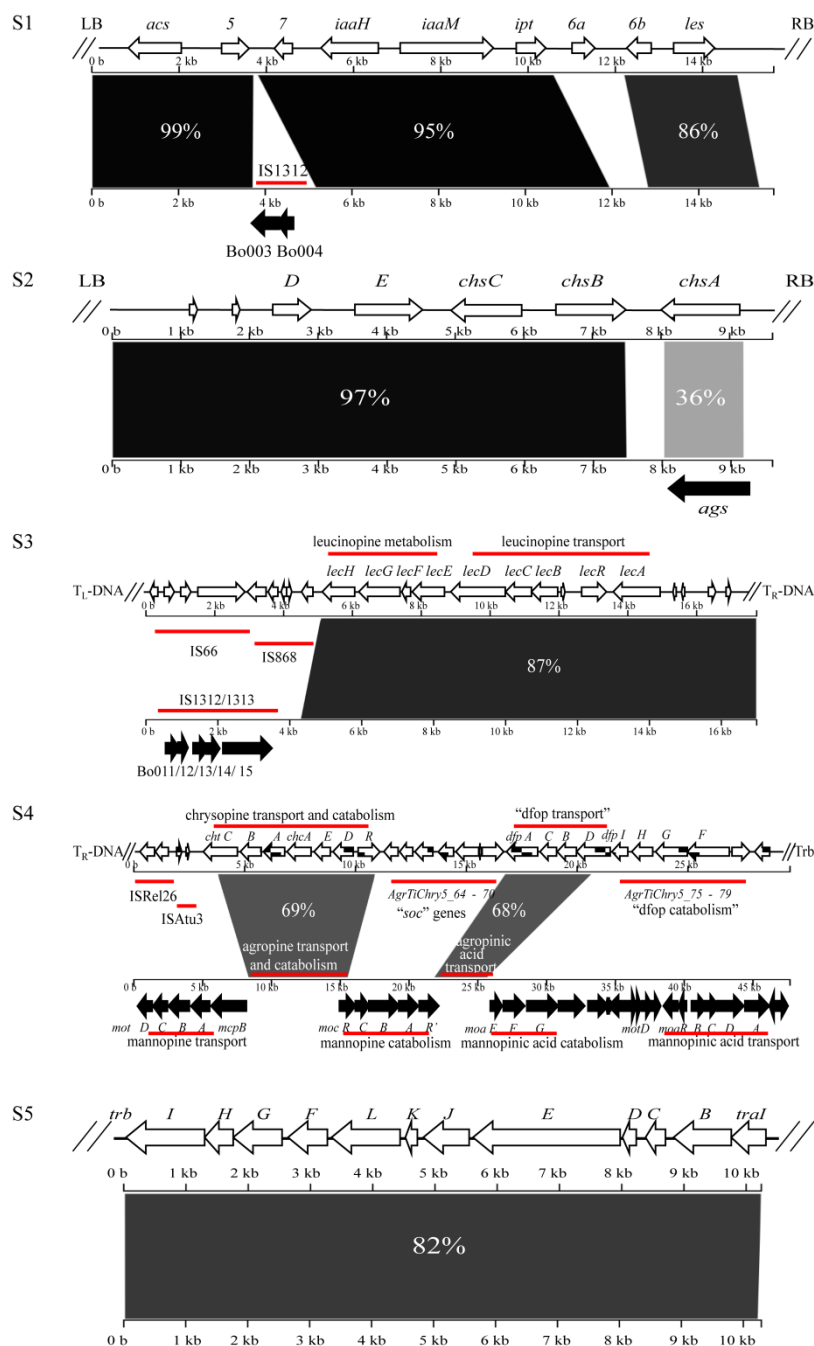
Acknowledgements

References

Baek, C., Farrand, S.K., Lee, K., Park, D., Lee, J.K., Kim, K., 2003. Convergent evolution of Amadori opine catabolic systems in plasmids of *Agrobacterium tumefaciens*. J.Bacteriol. 185, 513-524.

Baek,C., Farrand,S.K., Park,D., Lee,K., Hwang,W., Kim,K., 2005. Genes for utilization of deoxyfructosyl glutamine (DFG), an amadori compound, are widely dispersed in the family *Rhizobiaceae*. FEMS Microbiol.Ecol. 53, 221-233.

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

Bush, A. L., Pueppke, S. G., 1991. Characterization of an unusual new *Agrobacterium tumefaciens* strain from *Chrysanthemum moriflorum* Ram. Appl. Env. Microbiol. 57, 2468-2472.

Chilton, W. S., Stomp, A. M., Beringue, V., Bouzar, H., Vaudequin-Dransart, V., Petit, A., Dessaux, Y., 1995. The chrysopine family of Amadori-type crown gall opines. Phytochemistry 40, 619-628.

Darling, A. E., Mau, B., Perna, N. T., 2010. ProgressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. Plos One 5, e11147.

De Groot, M. J., Bundock, P., Hooykaas, P. J.J., Beijersbergen, A. G., 1998. *Agrobacterium tumefaciens*-mediated transformation of filamentous fungi. Nature Biotechnol. 16, 839-842.

Deng, W., Gordon, M.P., Nester, E.W., 1995. Sequence and distribution of IS1312: evidence for horizontal DNA transfer from *Rhizobium meliloti* to *Agrobacterium tumefaciens*. J.Bacteriol. 177, 2554-2559.

Deng, W., Pu, X., Goodman, R.N., Gordon, M.P., Nester, E.W., 1995. T-DNA genes responsible for inducing a necrotic response on grape vines. Mol. Plant Microbe Int. 8, 538-548.

Dessaux, Y., Petit, A., Tempe, J., 1993. Chemistry and biochemistry of opines, chemical mediators of parasitism. Phytochemistry 34, 31-38.

Ellis, J.G., Ryder, M.H., Tate, M.E., 1984. *Agrobacterium tumefaciens* T_R-DNA encodes a pathway for agropine biosynthesis. Mol.Gen. Genet. 195, 466-473.

Galens, K., Orvis, J., Daugherty, S., Creasy, H. H., Angiuoli, S., White, O., Wortman, J., Mahurkar, A., Giglio, M. G., 2011. The IGS standard operating procedure for automated prokaryotic annotation. Stand. Genomic Sci. 4, 244.

Gelvin, S.B., 2003. Agrobacterium-mediated plant transformation: the biology behind the "gene-jockeying" tool. Microbiol Mol. Biol. Rev. 67, 16–37.

González, V., Santamaría, R. I., Bustos, P., Hernández-González, I., Medrano-Soto, A., Moreno-Hagelsieb, G., Janga, S.C., Ramírez, M.A., Jiménez-Jacinto, V., Collado-Vides,J., Dávila, G., 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc.Natl.Acad. Sci. USA. 103, 3834-3839.

Gordon, J.E., Christie, P.J. 2014. The Agrobacterium Ti plasmids. Microbiol Spectr. 2, doi:10.1128 PLAS-0010-2013.

Henkel, C. V., den Dulk-Ras, A., Zhang, X., Hooykaas, P.J.J., 2014. Genome sequence of the octopine-type *Agrobacterium tumefaciens* strain Ach5. Genome Announc. 2, e00225-14.Jin, S. G., Komari, T., Gordon, M. P., Nester, E. W., 1987. Genes responsible for the supervirulence phenotype of *Agrobacterium tumefaciens* A281. J Bacteriol. 169, 4417-4425.

Kim, K., Farrand, S.K., 1996. Ti plasmid-encoded genes responsible for catabolism of the crown gall opine mannopine by *Agrobacterium tumefaciens* are homologs of the T-region genes responsible for synthesis of this opine by the plant tumor. J.Bacteriol. 178, 3275-3284.

Kim, K., Chilton, W.S., Farrand, S.K., 1996. A Ti plasmid-encoded enzyme required for degradation of mannopine is functionally homologous to the T-region-encoded enzyme required for synthesis of this opine in crown gall tumors. J.Bacteriol. 178, 3285-3292.

Konstantinidis, K. T., Tiedje, J. M., 2005. Genomic insights that advance the species definition for prokaryotes. Proc.Natl.Acad. Sci. USA. 102, 2567-2572

Korber, H., Strizhov, N., Feldwisch,J., Olsson, O., Sandberg, G., Palme, K., Schell, J., Koncz, C. 1991. T-DNA gene 5 of Agrobacterium modulates auxin response by autoregulated synthesis of a growth hormone antagonist in plants. EMBO J. 10, 3983-3991.

Kovács, L. G., Pueppke, S. G, 1994. Mapping and genetic organization of pTiChry5, a novel Ti plasmid from a highly virulent *Agrobacterium tumefaciens* strain. Mol. Gen. Genet. 242, 327-336.

Lyi, S.M., Jafri, S., Winans, S.C., 1999. Mannopinic acid and agropinic acid catabolism region of the octopine-type Ti plasmid pTi15955. Mol.Microbiol. 31, 339-347.

Machida, Y., Sakurai, M., Kiyokawa, S., Ubasawa, A., Suzuki,Y., Ikeda, J.-E., 1984. Nucleotide sequence of the insertion sequence found in the T-DNA region of mutant Ti plasmid pTiA66 and distribution of its homologues in octopine Ti plasmid. Proc.Natl.Acad.Sci. USA 81,7495-7499.

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

Nester, E.W., Gordon, M.P., Amasino, R.M,, Yanofsky, M.F. 1984. Crown gall: a molecular and physiological analysis. Ann. Rev. Plant Physiol. 35, 387-413.

Oger,P., Farrand, S.K., 2001. Co-evolution of the agrocinopine opines and the agrocinopine-mediated control of TraR, the quorum-sensing activator of the Ti plasmid conjugation system. Mol.Microbiol. 41, 1173-1185.

Oger,P., Reich, C., Olsen, G.J., Farrand, S.K., 2001. Complete nucleotide sequence and analysis of pTiBo542. Plasmid 45, 169-170.

Overbeek, R., Olson, R., Pusch, G. D., Olsen, G. J., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Parrello, B., Shukla, M., Vonstein, V., Wattam, A. R., Xia, F., Stevens, R., 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). Nucleic Acids Res. 42, D206-D214.

Păcurar, D. I., Thordal-Christensen, H., Păcurar, M. L., Pamfil, D., Botez, C., Bellini, C., 2011. *Agrobacterium tumefaciens*: From crown gall tumors to genetic transformation. Physiol. Mol Plant Pathol. 76, 76-81.

Palanichelvam, K., Oger, P., Clough, S. J., Cha, C., Bent, A. F., Farrand, S. K., 2000. A second T-region of the soybean-supervirulent chrysopine-type Ti plasmid pTiChry5, and construction of a fully disarmed vir helper plasmid. Mol. Plant Microbe Int. 13, 1081-1091.

Paulus, F., Canaday, J., Otten, L., 1991. Limited host range Ti plasmids: recent origin from wide host range Ti plasmids and involvement. Mol. Plant Microbe Int. 4, 190-197.

Ponsonnet, C., Normand, P., Pilate, G., Nesme, X., 1995. IS292: a novel insertion element from *Agrobacterium*. Microbiology 141, 853-861.

Schrammeijer, B., Risseeuw, E., Pansegrau, W., Regensburg-Tuïnk, T. J., Crosby, W. L., Hooykaas, P.J.J, 2001. Interaction of the virulence protein VirF of *Agrobacterium tumefaciens* with plant homologs of the yeast Skp1 protein. Current Biology 11, 258-262.

Stothard, P., Wishart, D. S., 2005. Circular genome visualization and exploration using CGView. Bioinformatics 21, 537-539.

Tzfira, T., Citovsky, V., 2006. *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. Curr. Opin. Biotech. 17, 147-154.

Vaudequin-Dransart,V., Petit,A., Poncet,C., Ponsonnet,C., Nesme,X., Jones, J.B., Bouzar, H., Chilton, W.S., Dessaux, Y., 1995. Novel Ti plasmids in Agrobacterium strains isolated from fig tree and chrysanthemum tumors and their opinelike molecules. Mol. Plant Microbe Int. 8, 311-321

Vaudequin-Dransart,V., Petit,A., Chilton, W.S., Dessaux, Y., 1998. The cryptic plasmid of *Agrobacterium tumefaciens* cointegrates with the Ti plasmid and cooperates for opine degradation. Mol. Plant Microbe Int. 11, 583-591.

Wabiko, H., 1992. Sequence analysis of an insertion element, IS1131, isolated from the nopaline-type Ti plasmid of *Agrobacterium tumefaciens*. Gene 114, 229-233.

Wintersinger, J. A., Wasmuth, J. D., 2015. Kablammo: an interactive, web-based BLAST results visualizer. Bioinformatics. 31, 1305-1306.

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

Yamamoto,S., Agustina, V., Sakai, A., Moriguchi, K., Suzuki, K., 2017. An extra *repABC* locus in the incRh2 Ti plasmid pTiBo542 exerts incompatibility toward an incRh1 plasmid. Plasmid 90, 20-29.

Figure S1-11. Schematic comparison between Ti plasmid pTiChry5 and reference pTiBo542. The alignment was calculated using the Blast 2.2.31+ algorithm (NCBI) and visualized by the web-tool Kablammo (Wintersinger and Wasmuth, 2015). The arrows represent genes on pTiChry5 at the top and the black arrows indicate the different genes on pTiBo542 at the bottom. S1, T_L-DNA; S2, T_R-DNA; S3, the region between T_L-DNA and T_R-DNA; S4, the region between T_R-DNA and *trb* operon; S5, *trb* operon; S6, *repABC* operon and some neighboring genes; S7, the region between the *repABC* operon and the main *vir* genes; S8, *vir* genes; S9, the region between the main *vir* genes and the *tra* operon; S10, *tra* operon; S11, the region between the *tra* operon and T_L-DNA.
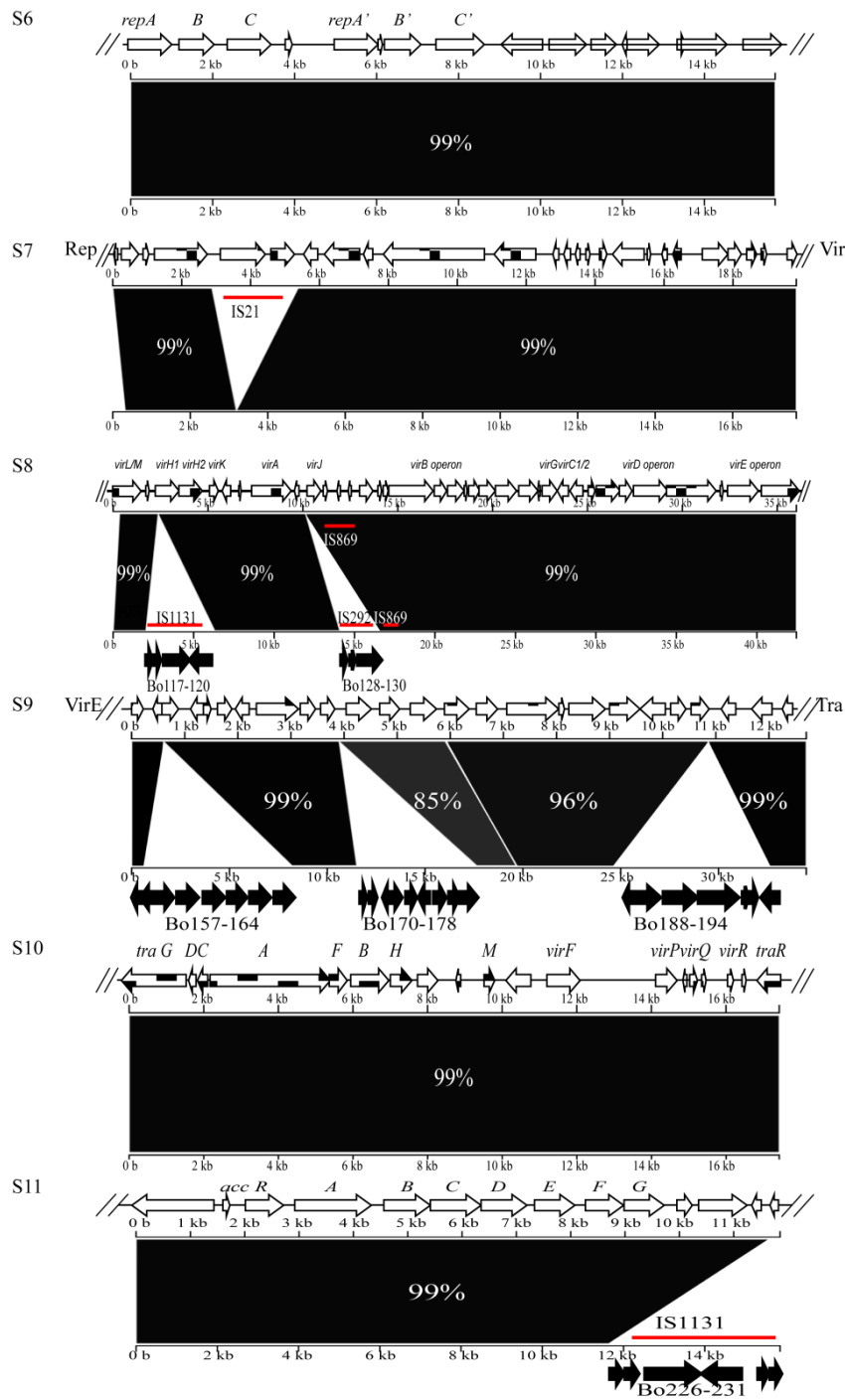
Figure S12. Alignment of the VirF protein from the octopine Ti plasmid pTi15955 in comparison with F-box proteins from nopaline pTiC58, agropine pTiBo542 and pTiChry5. The F-box is indicated in the figure by a red box.

Figure S13. Alignment of the Mas1 proteins from pTiChry5 and pTiBo542; the differences are indicated in the figure by a red box.

S12

```
pTiChry5   MVGSQDFRGVRYMQNIVSNSQTQRQPIEKEYSPFHRVMSDDLISRIVDGMVTNDPVETAR
pTiBo542   ------------MQNIVSNSQTQRQPIEKEYSPFHRVMSDDLISRIVDGMVTNDPVETAR
pTiC58     ---------MEPSQRSLSD-ESEASRTEVHSTSTNRDMPPELLAKVATYIPTQDPVETAR
pTi15955   ------------MRNSSLRD-ASGSNDAQVPNKTELLNLPDHVLTEVAKRLATNNPVESAF
                       :  : :   :   .  :       :  .::::.:.   :  *::***:*.

pTiChry5   NISNLKLSNKSVKESVERSAGGRFHGQINRLGAASKALYDHAVPTEGFSEHPEFRPALPR
pTiBo542   NISNLKLSNKSVKESVERSAGGRFHGQINRLGAASKALYDHAVPTEGFSEHPEFRPALPR
pTiC58     NLGSLERTGRAGREAVTSDPVGKYHARMKRIGASAKTVFDTVIPGNQLPEWETNRPSPTA
pTi15955   NIANFSKSHRFTRDAVRTEPLEKFSSRLKILSRNAKLLSHAVRHAATLPDGEQLSEAQLS
           *:..:.  : :   :::*   .      :: .:::  :.   :*  : . .       ::        :

pTiChry5   DYALAGERIDAIGGTLKLQTPQRKSALVDH-----ILNIAEPYDQGHALEYVAPHVGEFS
pTiBo542   DYALAGERIDAIGGTLKLQTPQRKSALVDH-----ILNIAEPYDQGHALEYVAPHVGEFS
pTiC58     R---TRAVGPILK----FQSEAGKSRFVTN-----ILNLPESA-QCDAILSVIKHLNDLG
pTi15955   QMRSEVATRPVLGVAYTHQDGQPEERLSGNHLDHKINNIPNL------------------
                  :       *      :. :  :       * *:  :

pTiChry5   SEDRRRLVEKMLEHFSTPFTDELPSDENFSGTYALFKALPHMEDSLK----ARMLDTVVN
pTiBo542   SEDRRRLVEKMLEHFSTPFTDELPSDENFSGTYALFKALPHMEDSLK----ARMLDTVVN
pTiC58     EANKTRLIERSIEILRLEP----PLTWNMGQKCPAVDVLVQGEKYLNADQLARIQQERSS
pTi15955   -------------VFNVAE----PIMFNEI---SALEVMAEVR------PIAR-------
                 :          *    *      ...: . .         **

pTiChry5   NPDMAALIQSQKERYDWLRAQDEYEKATWRSTQKPTVDERLADIESSVRSNLSL--GQGS
pTiBo542   NPDMAALIQSQKERYDWLRAQDEYEKATWRSTQKPTVDERLADIESSVRSNLSL--GQGS
pTiC58     RPRLSPLFGRAIADFEVQKTMD-ANPRRYRDAAGPEVRAVDTVIETIERYSASLARGGPN
pTi15955   ------------------------------------------------------------


pTiChry5   DLEKMRSARPLQQTISDTLNLAREELKSRAPDVRGR----
pTiBo542   DLEKMRSARPLQQTISDTLNLAREELKSRAPDVRGR----
pTiC58     AVGLLTTNESFEKNINEVYNRTRAELDASSRDRSRSGLSR
pTi15955   -------------SIKTAHDDARAELMSADRPRSTRGL--
                .*.  . : :* ** :
```

S13

```
                         20          40          60          80
                         |           |           |           |
chsB pTiChry5  MPISQAYLGS AEPPSGIKSF YFIRHGTTDL NEKEMVVNGE KHWGVQGSGT NIGLNEKGVR QALLAGNVLR TLPISKVVCS 80
mas1 pTiBO542  MPISQAYLGS AELPSGIKSF YFIRHGTTDL NEKEMVVNGE KHWGVQGSGT NIGLNEKGVR QALLAGNVLR TLPISKVVCS 80
                         100         120         140         160
                         |           |           |           |
chsB pTiChry5  PLLRALQTAF IANPGCPTFQ IEDDLQERDF GEHEGGFGPL QMFEDDYPDC EPTEIFSIHV AKALKHARKE NVLLVAHGGV 160
mas1 pTiBO542  PLLRALQTAF IANPGCPTFQ IEDDLQERDF GEHEGGFGPL QMFEDDYPDC EPTEIFSIHV AKALEHARKE NVLLVAHGGV 160
                         180         200         220         240
                         |           |           |           |
chsB pTiChry5  LRVVAALLGV AITDEHTDNG RVLHFSLDAG NWSVRVIQSR VVMVTGATRG IGKAIAQELI RHGYRVSLGA RNVQDLEAVF 240
mas1 pTiBO542  LRVVAALLGV AITDEHTANG RVLHFSLDAG NWSVRVIQSP VVMVSGATRG IGKAIAQDLV RHGYRVSLGA RNVQALEAVF 240
                         260         280         300         320
                         |           |           |           |
chsB pTiChry5  GDQNEALHYA CFDALDHSSM KDWVSTTVAK FNKIDGLVNN AGCVDHVDLE KEINVEMLQK QWDINCVAPL IMTKLCMPYL 320
mas1 pTiBO542  GDQNEALHYA RFDALDHSSM KDWVSTTFAK FNRIDGLVNN AGCGDHVDLE KEINVELLQK QWDINCVAPL IMTKLCMPYL 320
                         340         360         380         400
                         |           |           |           |
chsB pTiChry5  IESGSGRIVN LNSMSRQRGF NSFVGYNVTN HACGSLTQII NHIGLDHRVG AIDICLAVSA TNMSSLTDTV GPEH------ 394
mas1 pTiBO542  IESGSGRIVN LNSMSGQRVA NSLVGYNMTK HGLGGLTKTT QHVGWDHGLR AVDICPGFVA TDMSSWTNLI GPDEMIQPQD 400
                         420
                         |
chsB pTiChry5  IAKLIREEIE RPNIAF---- --------N* 412
mas1 pTiBO542  IAKLVREAME RPNRAFVPKN EVLCMKEATR 430
```

Chapter 6
Complete sequence of succinamopine Ti-plasmid pTiEU6 reveals its evolutionary relatedness with nopaline-type Ti-plasmids

Shuai Shao, G. Paul H. van Heusden and Paul J. J. Hooykaas

Department of Molecular and Developmental Genetics, Plant Cluster, Institute of Biology, Leiden University, Leiden, 2333 BE, The Netherlands

Abstract

Agrobacterium tumefaciens is the etiological agent of plant crown gall disease, which is induced by the delivery of a set of oncogenic genes into plant cells from its tumor-inducing plasmid (Ti-plasmid). Here we present the first complete sequence of a succinamopine-type Ti-plasmid. Plasmid pTiEU6 is comprised of 176,375 bp with an overall GC content of 56.1% and 195 putative protein-coding sequences could be identified. This Ti-plasmid is most closely related to nopaline-type Ti-plasmids. It contains a single T-region which is somewhat smaller than that of the nopaline-type Ti-plasmids and in which the gene for nopaline synthesis is replaced by a gene (*sus*) for succinamopine synthesis. Also in pTiEU6 the nopaline catabolic genes are replaced by genes for succinamopine catabolism. In order to trace the evolutionary origin of pTiEU6, we sequenced six nopaline Ti-plasmids to enlarge the scope for comparison to this class of plasmids. Average nucleotide identity analysis revealed that pTiEU6 was most closely related to nopaline Ti-plasmids pTiT37 and pTiSAKURA. In line with this traces of several transposable elements were present in all the nopaline Ti plasmids and in pTiEU6, but one specific transposable element insertion, that of a copy of IS1182, was present at the same site only in pTiEU6, pTiT37 and pTiSAKURA, but not in the other Ti plasmids. This suggests that pTiEU6 evolved after diversification of nopaline Ti-plasmids by DNA recombination between a pTiT37-like nopaline Ti-plasmid and another plasmid, thus introducing amongst others new catabolic genes matching a new opine synthase gene for succinamopine synthesis.

Introduction

Agrobacterium tumefaciens, a gram-negative plant pathogen belonging to the family *Rhizobiaceae*, is the causative agent of crown gall disease, which causes severe losses in agriculture. It induces tumor formation in plants by transferring a segment of its tumor-inducing (Ti) plasmid, the T-DNA, to plant cells, and has been developed as a vector to create transgenic plants and fungi and *Agrobacterium*-mediated transformation (AMT) has become the preferred method of transformation of these organisms over the past decades.

In the tumors induced by *A. tumefaciens*, tumor-specific metabolites called opines are formed under control of genes present on the T-DNA. According to the different opines found in the tumors, Ti-plasmids are classified as nopaline-type, octopine-type, agropine-type, vitopine-type, succinamopine-type, and chrysopine-type. Examples of most of these Ti-plasmid types have now been characterized in detail by DNA sequencing. The octopine Ti-plasmid (Zhu et al. 2000) and the nopaline Ti-plasmid pTiSAKURA (Suzuki et al. 2000) were the first that were fully characterized by DNA sequencing. Later with the advent of next generation sequencing (NGS) sequences of several other Ti-plasmids including nopaline-type pTiC58 (Goodner et al. 2001; Wood et al. 2001), octopine-type pTiAch5 (Henkel et al. 2014; Huang et al. 2015), vitopine-type pTiS4 (Slater et al. 2009), agropine-type pTiBo542 (Oger et al. 2001), and chrysopine-type pTiChry5 (Shao et al. 2018) became available.

Here we present the first sequence of a succinamopine-type Ti-plasmid, that of pTiEU6. Previously, it was shown that succinamopine strains such as EU6, AT181, and T10/73 contain a Ti plasmid which has a restriction profile that resembles that of nopaline Ti plasmid pTiT37 to some extent (Sciaky et al. 1978). These strains induce tumors in which succinamopine/asparaginopine (SAP), succinamopine lactam (SAL), and succinopine lactam (SOL) are present, but not nopaline (Chang and Chen 1983; Chilton et al. 1984). Succinamopine is a conjugate of α-ketoglutaric acid and asparagine, which is structurally related to nopaline, which is a conjugate of α-ketoglutaric acid and arginine. Although it was originally reported that these strains could degrade nopaline (Montoya et al. 1977), later it was reported that this was not the case, but that they could degrade leucinopine, a conjugate of α-ketoglutaric acid and leucine, besides SAP and SOL (Chilton et al. 1985). Agropine strains such as Bo542 and chrysopine strains such as Chry5 produce tumors that besides agropine and chrysopine, respectively, also contain leucinopine and succinamopine, which they also can degrade. Leucinopine and succinamopine produced in tumors induced by Bo542 and Chry5 have the unusual L,L-stereochemistry instead of the D,L-isomer, which is common for opines including leucinopine and succinamopine made in tumors induced by strains carrying pTiEU6. By comparing the pTiEU6 sequence to that of a set of nopaline Ti-plasmids, including six newly sequenced nopaline Ti-plasmids, we found that pTiEU6 lacks the nopaline synthase gene in the T-region and the adjacent genes for nopaline import and catabolism, but instead has obtained a gene for succinamopine synthase in the T-region and an adjacent set of genes, which show similarity to the pTiBo542 and pTiChry5 genes involved in succinamopine and leucinopine import and catabolism. We conclude that pTiEU6 is derived from a nopaline Ti-plasmid related to the group of pTiT37 and pTiSAKURA by DNA recombination with another so far unknown plasmid.

Material and methods

Bacterial growth and DNA extraction

A. tumefaciens strains LBA941 and LBA942, are Ti-plasmid-cured derivatives of strain C58, containing pTiEU6 and pTiKerr27, respectively, and were obtained as strains A289 and A293, respectively, from dr. E. W. Nester (Seattle, USA). Ti-plasmid pTiT37 is harbored in a derivative of strain T37, LBA8374 which is from our own strain collection. Wildtype strain LBA9670 was obtained as strain 1 from dr. S. Süle (Budapest, Hungary); strain LBA9280 was obtained from dr. A. Kerr as Kerr108 and strains LBA8670 and LBA8790 were received from dr. X. Nesme (Lyon, France) as strains CFBP2178 and CFBP1935, respectively. All *A. tumefaciens* strains were cultured in LC medium (Hooykaas et al. 1977) with appropriate antibiotics at 29 °C.

After growth on LC medium at 29 °C, genomic DNA was isolated using the DNeasy Blood & Tissue Kit (QIAGEN) with a few modifications. Briefly, after the cells were harvested, the pellets were re-suspended in TES buffer (0.01M Tris-HCl pH8.0, 0.05M NaCl, 0.05M EDTA). Then the cells were treated with 400µg/ml lysozyme for 30 minutes on ice. After these steps, the kit protocol was followed.

DNA sequencing, sequence assembly and analysis

Total genomic DNA was isolated from *A. tumefaciens* strains and two libraries were constructed using the HiSeq SBS kit v4. Paired-end 125 cycles sequence reads were generated using the Illumina HiSeq2500 system by BaseClear (Leiden, the Netherlands). After quality analysis, paired-end reads were retained with an average length of 125 bp. Initially, assembly was carried out using CLC Genomics Workbench software (version 7.0.4). Subsequently, the Blast 2.2.31+ algorithm (NCBI) was used to map all contigs to the reference sequences, nopaline-type pTiC58 (AE007871) (Wood et al. 2001) and octopine-type pTiAch5 (CP007228) (Henkel et al. 2014). Based on sequence conservation among Ti-plasmids and topology, possible contigs with high hits were identified. The relative order and relationship of these contigs was determined by a series of PCRs and gaps were filled in by sequencing relevant positive PCR products. PCR products were purified by agarose gel electrophoresis and sequenced by Macrogen (Amsterdam, The Netherlands) and all primers were obtained from Sigma-Aldrich.

Protein-coding genes were predicted using the program Glimmer (version 3.02) (Delcher et al. 2007) and submitted to the RAST annotation system (http://rast.nmpdr.org/) for functional annotation (Overbeek et al. 2014). The assembled sequence was annotated using the IGS Prokaryotic Annotation Pipeline as well (Galens et al. 2011). Functional characterization was generated using eggNOG-Mapper which employed the functional categories in COG, KOG and arCOG databases (Huerta-Cepas et al. 2017). Insertion elements (IS elements), tRNAs and rRNAs were identified using IS finder (https://www-is.biotoul.fr/), tRNAscan-SE (http://lowelab.ucsc.edu/tRNAscan-SE/) and RNAmmer (http://www.cbs.dtu.dk/services/RNAmmer/), respectively. IslandViewer was used to predict genome islands (Dhillon et al. 2015) and CGView was used to generate circular maps of Ti plasmids (Stothard and Wishart 2005). Average nucleic acids identity (ANI) values between the query plasmid and the reference plasmid were calculated by the Perl script (Konstantinidis and Tiedje 2005). Comparative maps of complete Ti-plasmid sequences were constructed by CGView using BlastN (Stothard and Wishart 2005) and the Mauve program (Darling et al.

2010). The alignments were visualized by the web-tool Kablammo (Wintersinger and Wasmuth 2015) and sequence similarity of T-DNA borders was illustrated using Weblogo (http://weblogo.berkeley.edu/).

Phylogenetic analysis

Phylogenetic analysis was performed using MEGA version 7.0 (Kumar et al. 2016). Distances (distance options according to the Kimura two-parameter model) and clustering using the neighbor-joining (NJ) and maximum likelihood (ML) methods were determined using bootstrap values based on 1000 replicates. The consensus trees represent combined data from 3 independent runs, with 100% consensus required for inference of relatedness.

The kSNP3 program (Gardner et al. 2015) was performed to identify core Single Nucleotide Polymorphisms (SNPs) in the set of pTiEU6 and nopaline Ti-plasmids and to estimate phylogenetic tree based upon those SNPs. The optimal value of K-mer size was obtained by Kchooser of the kSNP3 package. The kSNP3 was conducted using the default parameters with the core and the maximum likelihood tree options. After that a phylogenetic maximum-likelihood tree was generated using FigTree v1.4.3 software (Rambaut 2016). The branch lengths are expressed in terms of changes per number of SNPs.

Nucleotide sequence accession numbers

The complete nucleotide sequences of the Ti-plasmids sequenced in this study are deposited at GenBank and are available with the following accession numbers: pTiEU6, KX388535; pTiSule1, MK439381; pTiCFBP2178, MK439382; pTiCFBP1935, MK439383; pTiKerr108, MK439384; pTiKerr27, MK439385, and pTiT37, MK439386.

Results

Genomic overview of succinamopine Ti-plasmid pTiEU6

The complete sequence of pTiEU6 comprises 176,375 bp and has an average GC content of 56.1%. In total, 195 open reading frames (ORF) were identified with an average size of 761bp (table 1 and fig. 1). To analyze the functional content of ORFs of pTiEU6, COG functional analysis was performed by eggNOG-Mapper. In total, 120 ORFs were assigned to COG categories, whereby the intracellular trafficking and secretion (group U) and unknown function (group S) categories contained the highest number of proteins. Ti-plasmid pTiEU6 carries one T-region of 18,789 bp with a GC content of 45.7%, which is significantly lower than other regions of the Ti-plasmid. Genes encoding transfer-RNA (tRNA) or ribosomal RNA (rRNA) and genome islands (GI) could not be identified.

Table 1. General features of Ti-plasmid pTiEU6 and newly sequenced nopaline Ti-plasmids

Feature	pTiEU6	pTiT37	pTiKerr27	pTiSule1	pTiKerr108	pTiCFBP1935	pTiCFBP2178
Size (bp)	176,375	203,781	243,905	217,820	220,307	213,092	217,821
G+C (%)	56.1	55.9	57.2	56.8	56.8	56.8	56.8
Open reading frames (ORFs)	195	198	233	202	207	196	202
ORFs assigned into COG	120 (61.5%)	138 (69.7%)	167 (71.7%)	145 (71.8)	148 (71.5%)	142 (72.4%)	146 (72.3%)

114

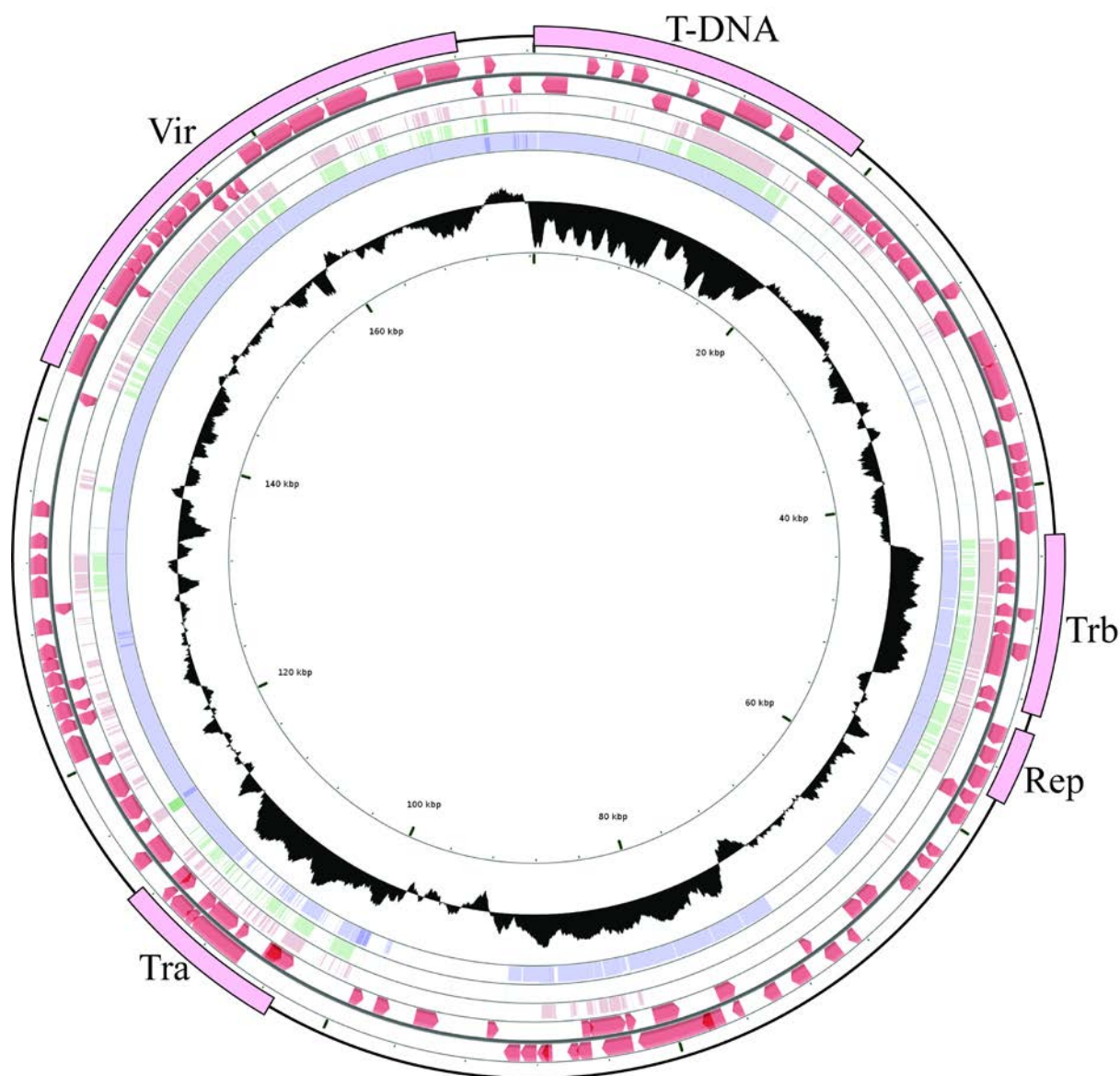| T-region (bp) | 18,789 | 23,415 | 24,835 | 24,477 | 24,477 | 24,477 | 24,477 |
| G+C (%) of T-regions | 45.7 | 45.8 | 46.1 | 45.8 | 45.8 | 45.8 | 45.8 |
| IS elements | 4 | 4 | 8 | 4 | 4 | 4 | 4 |

Figure 1. Schematic circular map of Ti-plasmid pTiEU6. Circle ranges from 1 (outer circle) to 6 (inner circle). Circle 1, location of T-DNA, *tra*, *trb*, *rep* and *vir* operons; Circles 2 and 3, predicted open reading frames on the plus strand and minus strand, respectively; Circles 4, 5, and 6, coordinated mapping of Ti-plasmids pTiBo542, pTiAch5, and pTiC58 against reference pTiEU6, respectively; Circle 7, G+C content percentages, median line represents the average GC content of the entire sequence. Single nucleotide based identities were expressed as height of the circles from blank to filled to represent 0-100% identity.

Genetic relatedness of pTiEU6 with other Ti-plasmids

An average nucleotide identity (ANI) analysis was performed to determine the genetic relatedness of pTiEU6 with other Ti- and Ri-plasmids. To this end, we extracted the complete sequences of nopaline pTiC58 (AE007871) and pTiSAKURA (AB016260), octopine pTiAch5 (CP007228), agropine pTiBo542 (DQ058764), chrysopine pTiChry5 (KX388536), mikimopine pRi1724 (AP002086, Moriguchi et al. 2001), cucumopine pRi2659 (EU186381, Mankin et al. 2007) and vitopine pTiS4 (CP000637) from public databases and calculated ANI values between pTiEU6 and them. As demonstrated in figure 2A, pTiEU6 shared the highest ANI values with nopaline Ti-plasmids pTiSAKURA (98.78%) and pTiC58 (97.67%), while ANI values of pTiEU6 were less than 85% with other opine type Ti- and Ri-plasmids. In addition, it became apparent that Ti-plasmids can be divided into two groups basically, which are nopaline and succinamopine Ti-plasmids on the one hand versus agropine, chrysopine and octopine Ti-plasmids on the other hand. Pair-wise alignments of pTiC58, pTiAch5, and pTiBo542 with reference pTiEU6 made with the BlastN algorithm (fig. 1) confirmed that pTiEU6 shared most of the conserved regions with pTiC58 except for the opine catabolism region between T-DNA and *trb* operon. These results demonstrate that pTiEU6 is most related to the nopaline Ti-plasmids.
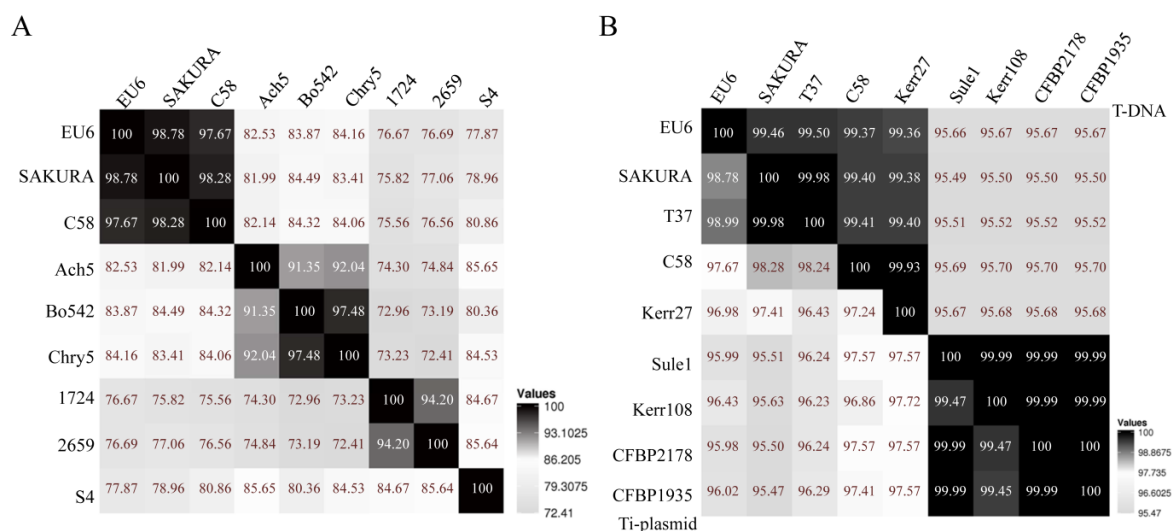

Figure 2. Heatmap of average nucleotide identity (ANI) values. The ANI values between the query sequence and the reference sequence were calculated by the Perl script (Konstantinidis and Tiedje 2005) and visualized by Heatmapper (Babicki et al., 2016). A, the ANI values of representative Ti-plasmids. B, the ANI values of pTiEU6 and sequenced nopaline Ti-plasmids. The ANI values of T-DNA were shown in the upper right part and those of whole Ti-plasmids were present in the lower left part.

To confirm this evolutionary relationship at the gene level, a phylogenetic tree was constructed using the concatenated nucleotide sequences of the genes *virD2, repC, traR/M* and *trbE*, which are involved in virulence, replication, regulation and conjugation, respectively (fig. 3). In the tree, besides the corresponding genes from nopaline pTiC58, octopine pTiAch5, agropine pTiBo542, chrysopine pTiChry5 and vitopine pTiS4, the corresponding genes from two hairy root inducing plasmids, *i.e.* mikimopine pRi1724 and cucumopine pRi2659 were

included as well. In line with the results from ANI analyses, pTiEU6 was clustered with nopaline-type Ti-plasmids, which form an individual branch. The trees confirm the close relatedness of pTiEU6 with the nopaline-type Ti-plasmids.
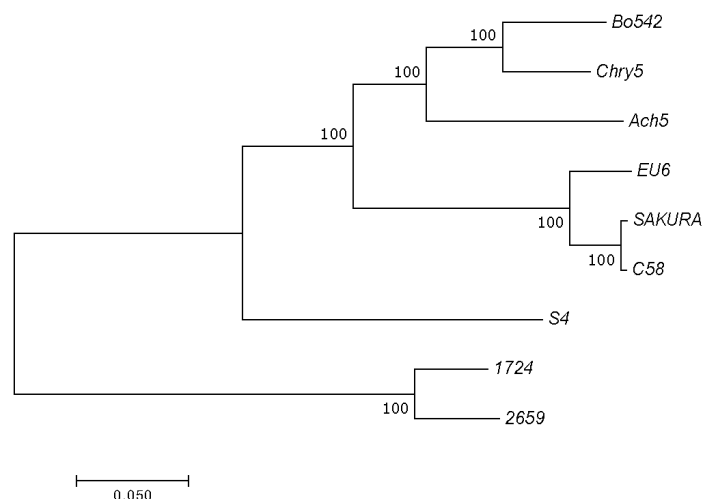

Figure 3. Phylogenetic tree constructed from alignments for concatenated nucleotide sequences of the genes *virD2*, *repC*, *traR/M* and *trbE* using the MEGA7.0 software. The topological structures of each tree are built and confirmed by ML and NJ methods. Bootstrap values from 1,000 replicates are located next to the branches and the evolutionary distances were computed with the Poisson correction. The scale bar represents the nucleotide substitutions rate units.

Evolutionary relationship between pTiEU6 and the class of nopaline Ti-plasmids
In order to get a more detailed insight into the evolutionary relationship between pTiEU6 and the nopaline Ti-plasmids, we went on to sequence six other nopaline Ti-plasmids from various geographical origins. The general features of these plasmids (pTiCFBP1935, pTiCFBP2178, pTiKerr27, pTiKerr108, pTiSule1 and pTiT37) are described in Table 1, which shows that they have similar features and their sizes range from 203,781 bp up to 243,905 bp with a similar GC content of about 56-57%. In addition, all of them present similar functional COG profiles.

The ANI values between pTiEU6 and all available nopaline Ti-plasmids were then calculated. As can be seen from figure 2B (lower left part), pTiEU6 shares the strongest similarities with pTiSAKURA (98.78%) and pTiT37 (98.99%). Similarities with the other nopaline Ti-plasmids are lower (95.98-97.67%). Similar results were obtained by ANI analyses of the T-regions (upper right part). Subsequently, global plasmid-wide sequence alignments were carried out, which are presented as linear maps by ProgressMauve (fig. 4A and supplementary fig. S1). The single nucleotide based similarities were displayed as height of the panels from blank to full to represent 0-100% similarity and color frames indicate the sequences in these regions are conserved. A pair-wise nucleotide sequence content distance was used to perform cluster analysis. At the same time, another phylogenetic relationship tree was deduced using SNP-based core sequence variations by the kSNP3 program (fig. 4B) (Gardner et al. 2015). The optimal value of K-mer size was 17 and obtained by Kchooser of the kSNP3 package. The number of core SNPs was 1091 and the total number detected was 4079. Using MUMmer, the numbers of SNPs were determined for each pair of Ti-plasmids (Kurtz et al. 2004). Compared with pTiEU6, the number of SNPs in nopaline Ti-plasmids were 1556 (pTiSAKURA), 1541 (pTiT37), 3745 (pTiC58), 4422 (pTiKerr27), 4764 (pTiKerr108), 4798 (pTiSule1), 4831 (pTiCFBP1935) and 4797 (pTiCFBP2178), respectively. All these data show that pTiEU6 is most similar to pTiSAKURA and pTiT37. We shall refer to these latter two
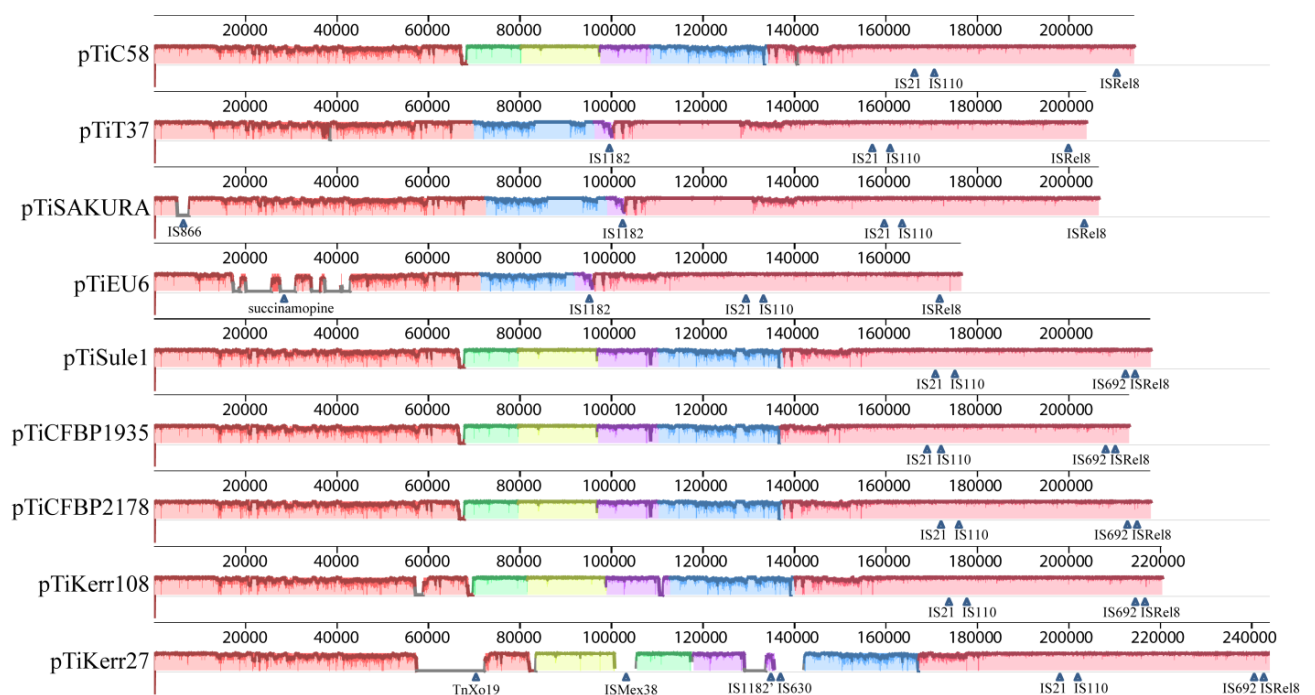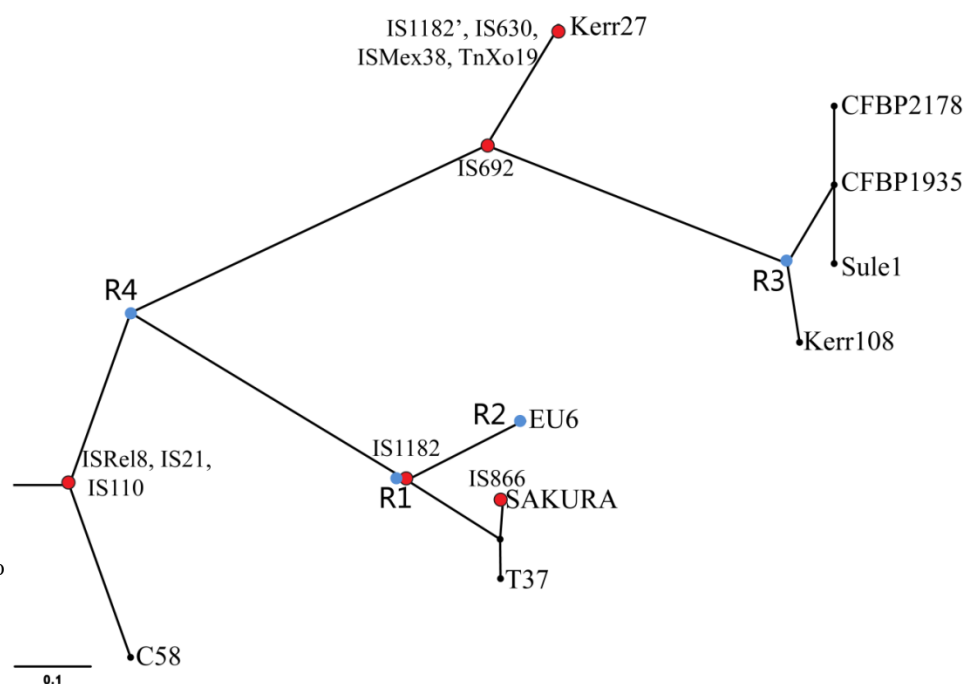
Figure 4. A, Schematic comparison of pTiEU6 and all sequenced nopaline Ti-plasmids. ProgressiveMauve (Darling et al., 2010) was used to generate pairwise alignments with default parameters and the single nucleotide based similarities were expressed as height of the panels from blank to filled to represent 0-100% similarity. Different colored blocks represent conserved regions and arrows indicated the site of transposons. B, Core SNPs analysis. SNP-based phylogenetic tree of pTiEU6 and nopaline Ti-plasmids was visualized using FigTree v1.4.3 software (Rambaut, 2016). Core SNPs were identified by the kSNP3.0 program (Gardner et al., 2015). Red nodes represent the insertion of transposable elements. Blue nodes R1, R2, R3 and R4 represent DNA recombination events leading to exchanges in the area between *rep* and *tra* operon (R1), opine synthesis and catabolism area (R2), the area containing genes *6a* and *6b* (R3) and the area of *tra* operon containing genes *traG/D/C/A* (R4). The R4 event followed the insertion of three transposable elements ISRel8, IS21 and IS110, which are present in all nopaline Ti-plasmids. Later the R2 event took place after the acquirement of IS1182. Whether the recombination events R1 and R3 coincided or not in line with the acquirement of IS1182 and IS692, respectively, cannot be deduced from the data. The scale bar represents the nucleotide substitutions rate units.

plasmids as forming subgroup A of the nopaline Ti plasmids, the other studied nopaline Ti-plasmids forming subgroup B (pTiKerr27, pTiKerr108, pTiSule1, pTiCFBP1935 and pTiCFBP2178).

Transposable elements

Over time transposable elements may insert in chromosomes and plasmids and their presence at an identical position in different Ti-plasmids may thus point to a common descent. Several (remnants of) transposable elements were found in pTiEU6 and the nopaline type Ti-plasmids, which hereinafter will be referred to by the name of the most closely related element found in the IS finder database. However, this may reflect only limited similarity to these elements in short segments only. Nevertheless these turned out to be very useful in tracing the evolutionary relationships between these Ti plasmids. Transposable element ISRel8 (AgrTiEU6_207, IS66 family), IS21 (AgrTiEU6_149) and IS110 (AgrTiEU6_154) are present at an identical position in all nopaline Ti-plasmids and in pTiEU6. They apparently integrated before the diversification of these plasmids and their spread over the world. The insertion element IS1182 (AgrTiEU6_104) is present at an identical position only in the subgroup A nopaline Ti-plasmids and in pTiEU6. Apparently, this IS1182 copy was integrated in a group A nopaline plasmid before it recombined with another plasmid to generate pTiEU6. Plasmid pTiKerr27 also contains a copy of IS1182, in the figure 4 called IS1182', but at a completely different position and thus representing an independent insertion event. Another insertion element IS692 (IS66 family), occurs at a site close to ISRel8 in the subgroup B nopaline Ti-plasmids, except pTiC58. An insertion of IS866 is present specifically in the T-DNA of pTiSAKURA and must represent a relatively recent insertion event. A similar IS866 element was previously found to inactivate the *iaaH* gene in the T-DNA of several octopine Ti-plasmids (Paulus et al. 1989).

Plasmid pTiKerr27 is the largest nopaline Ti-plasmid discovered so far. This plasmid is a group B nopaline Ti plasmid, but it contains four additional transposable elements in areas with segments of novel genes. These areas may therefore have been introduced into pTiKerr27 by transposition. A gene cluster was found with partial similarity to the transposon module TnXo19 with genes *tniA/B/Q* (AgrTiKerr27_65/66/67) (Mindlin et al. 2001). Another transposable element related to ISMex38 (AgrTiKerr27_00095), is present in another non-homologous region of pTiKerr27. Finally, a large set of non-conserved hypothetical genes was introduced along with insertion element IS630 (AgrTiKerr27_00128), which is located closely to IS1182' (AgrTiKerr27_00126). Overall the evolutionary path revealed by the (remnants of) transposable elements thus concurs with the analyses by similarity scores of the whole Ti-plasmid and their individual genes.

Regions shared between pTiEU6 and nopaline Ti plasmids

As stated in the previous paragraph, we compared the main gene clusters at the nucleic acid level to determine the evolutionary relationship between pTiEU6 and nopaline Ti-plasmids by using ProgressMauve (fig. 4A). Taken altogether, the homologies are very high with the subgroup A nopaline Ti-plasmids in all areas except for the region responsible for nopaline and succinamopine catabolism, respectively. A large part of the region between *repABC* and the *tra* operon of almost 30 kb is strongly different between the subgroup A and subgroup B nopaline Ti-plasmids, as can be seen by the colored blocks in figure 4A. In pTiKerr27 the insertion of transposable element ISMex38 (AgrTiKerr27_00095) in the middle of this region has led to a

local inversion, as shown in figure 4A, the green and yellow blocks. These results are again in line with results from SNPs and ANI analyses showing that pTiEU6 is more closely related to pTiSAKURA and pTiT37 than to other nopaline Ti-plasmids.
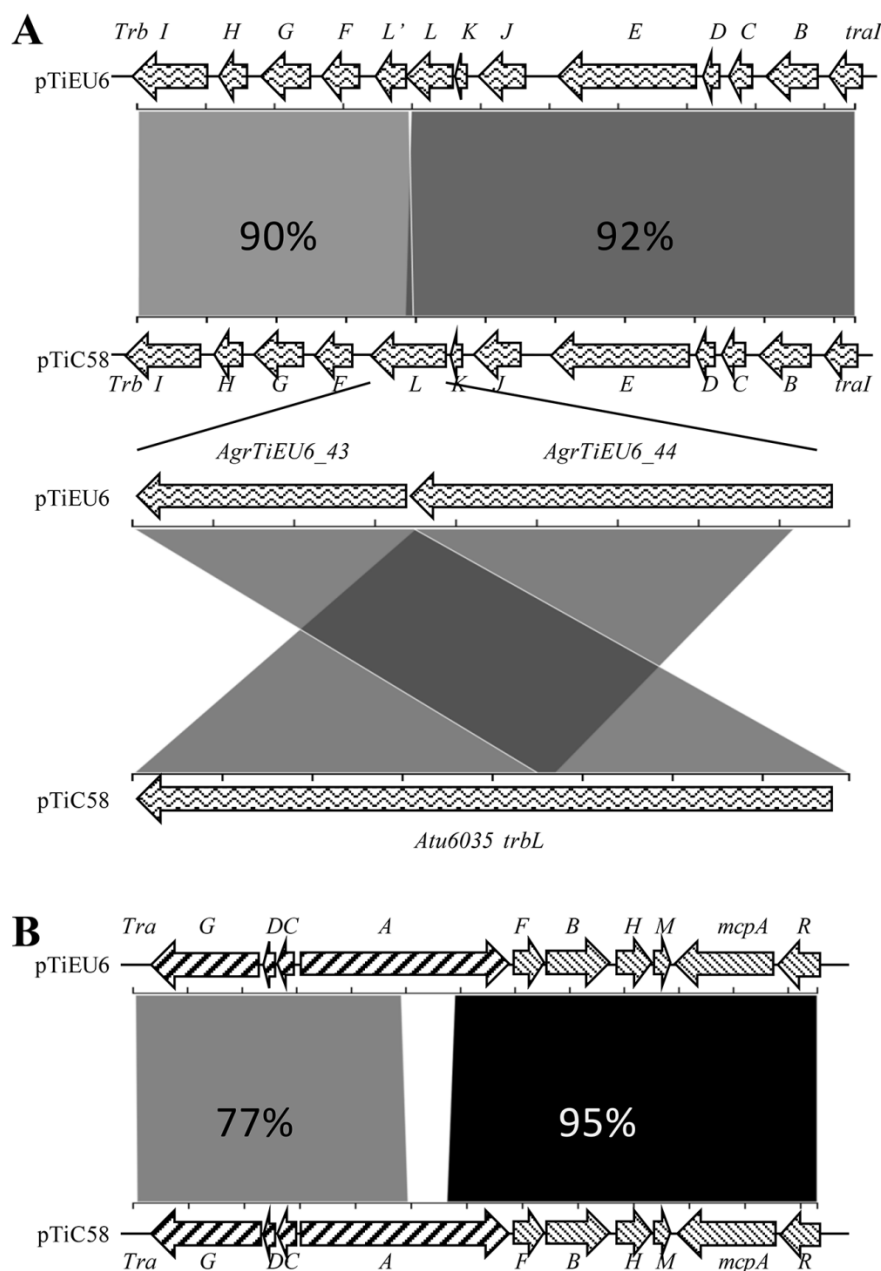

Figure 5. Schematic comparison of the *tra* and *trb* operons of pTiEU6 with those of pTiC58. The alignment was calculated using the Blast 2.2.31+ algorithm (NCBI) at nucleotide level and visualized by the web-tool Kablammo (Wintersinger and Wasmuth, 2015). The arrows represent genes in pTiEU6 at the top and in pTiC58 at the bottom, respectively. The identity values of each single gene at amino acid level are present in the supplementary table S1.

Gene clusters associated with plasmid replication and maintenance (*rep*) and virulence (*vir*), share very strong similarities with those in all nopaline Ti-plasmids (approximately 99% identical at nucleotide level) (supplementary fig. S2). A similar strong similarity of 90-92% was seen for the *trb* operon encoding the type4 secretion system (T4SS) involved in conjugative plasmid transfer (fig. 5 and supplementary table S1). It was remarkable that in contrast to the nopaline Ti-plasmids the *trbL* gene was split into two individual ORFs (AgrTiEU6_43 and AgrTiEU6_44) in reverse order in pTiEU6. Also the pTiEU6 plasmid has an almost identical set of *acc* genes involved in agrocinopine import and catabolism (99% identity with those of

pTiC58 and pTiSAKURA at the nucleotide level). It suggests that the *rep, vir, acc* and *trb* operons evolved together. Although the *tra* operons of pTiEU6, responsible for conjugative DNA processing of the Ti-plasmid, were almost identical to those of pTiSAKURA and pTiT37, the *traCDG* and *traA* operons shared only 77% similarities with those of pTiC58 (fig. 5 and supplementary table S1). Pairwise alignments of the *tra* operons of nopaline Ti-plasmids revealed that the *tra* operons of nopaline Ti-plasmids were closely related to each other, with the exception of those of pTiC58. This suggests that a recombination event took place in this area in pTiC58 or in an early descendant of pTiC58 before the diversification of the nopaline Ti-plasmids. The *tra* operons of pTiC58 are more related to those of pAtS4c (94% nucleotide identity) and the nopaline catabolic plasmid pAtK84b (88% nucleotide identity).

T-region

The T-region of pTiEU6 defined by conserved border repeats, has a length of 18,789 bp (fig. 6A). Comparing the T-DNA of pTiEU6 to those of nopaline Ti-plasmids as first described by Otten et al. 1999 for pTiC58, it was found that the left border is identical to that of nopaline-type Ti-plasmids, but the right border is somewhat different from that of nopaline Ti-plasmids (fig. 6B). The T-DNA carries the essential oncogenes involved in the synthesis of the plant hormones auxin (*iaaH* and *iaaM*) and cytokinin (*ipt*). Also for the remaining part the T-region of pTiEU6 is collinear with those of the nopaline Ti plasmids, including a gene for agrocinopine synthase (*acs*). However, in comparison to the nopaline T-region it is truncated at both ends. At the left a gene encoding a putative second agrocinopine synthase torf6/orf172 (*acs*) and gene 5 are absent and at the right end the area occupied by oncogene *6b* and the nopaline synthase gene is replaced by a novel gene (AgrTiEU6_18). Blundy et al. (1986) showed that the sequences at the right end of the T-region are involved in the biosynthesis of succinamopine. Gene AgrTiEU6_18 encodes a protein which contains a NAD/NADP dehydrogenase domain as is also present in nopaline synthase, but otherwise is very different This gene must be responsible for succinamopine synthesis and we shall call AgrTiEU6_18 from now on *sus* as the gene for succinamopine synthase. It is important to remark here that *sus* is very different from the gene *les*, sometimes also called *sus* and responsible for succinamopine and leucinopine biosynthesis in tumors induced by agropine and chrysopine strains. This is not unexpected as succinamopine strains such as EU6 induce tumors with D^{glu}, L^{asp}-succinamopine (Chilton et al. 1984), whereas agropine strains such as Bo542 and chrysopine strains such as Chry5 induce tumors with L^{glu}, L^{asp}-succinamopine (Chilton et al. 1985). In order to avoid confusion we propose to call the previously described Bo542 and Chry5 gene encoding the enzyme catalyzing the production of the L,L-isomer *susL* and the EU6 gene encoding the enzyme catalyzing the formation of the D,L-isomer *susD*. The susD and susL genes are very different, but it is intriguing that they contain a small area of similarity, which is present in the N-terminal part of the protein encoded by *susL* and in the C-terminal part of the protein encoded by *susD*.

When comparing the T-regions of the nopaline Ti-plasmids, it was found that pTiKerr108, pTiCFBP1935, pTiCFBP2178, and pTiSule1 lack gene *6a* and instead have another gene called *3'* inserted between gene *6b* and the *nos* gene (fig. 6A). Although a *6b* gene is still present, this exhibited only a low nucleotide similarity of 65% with the gene *6b* encoded by the other nopaline Ti-plasmids in this study. The gene *3'* showed 71% nucleotide similarity with the corresponding gene of octopine-type Ti-plasmids. This suggests that genes 6a and 6b were

replaced by gene 3' and a different version of 6b by a local DNA recombination event with another Ti-plasmid. These differences with pTiC58 were noted before by the Otten lab, after sequencing this part of the T-regions of the nopaline Ti plasmids pTi82.139, which is identical to pTiCFBP2178, and pTiAB4 (Drevet et al. 1994; Otten and De Ruffray 1994).
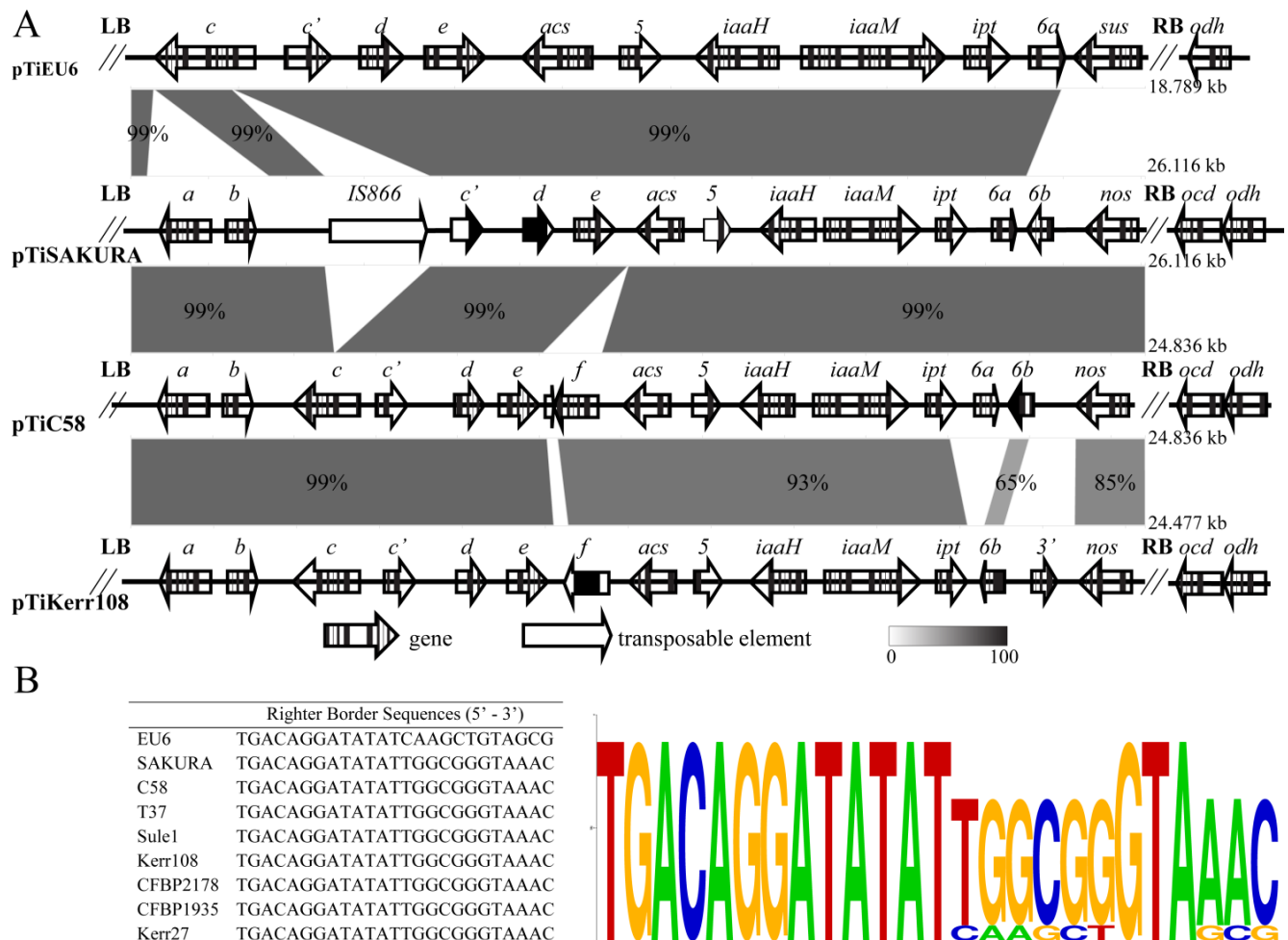

Figure 6. A, Structure of T-DNAs of pTiEU6 and nopaline pTiC58, pTiKerr108 and pTiSAKURA. Locations of open reading frames are shown by arrows. The pairwise alignments were calculated using the BlastN algorithm and visualized by the web-tool Kablammo (Wintersinger and Wasmuth, 2015). B, Multiple sequence alignments of T-DNA right border repeats from pTiEU6 and nopaline Ti-plasmids were generated by WebLogo and a consensus is shown.

Succinamopine catabolism

Nopaline Ti-plasmids have strongly conserved genes involved in nopaline import and catabolism (*noc*), which are located adjacent to the right border of the T-region. In pTiEU6 these genes are absent, but replaced by another set of genes. Tumors induced by strains containing pTiEU6 do not contain nopaline, but instead succinamopine/asparaginopine is present, which is related to nopaline, but in which arginine is replaced by asparagine (Chang and Chen 1983; Chilton et al. 1984). Besides succinamopine (SAP), succinamopine lactam (SAL) and succinopine lactam (SOL) were found in these tumors (Chilton et al. 1984). The set

of genes adjacent to the RB of the T-region (AgrTiEU6_20 - AgrTiEU6_28) shares significant homology with the genes involved in succinamopine/leucinopine import of agropine plasmid pTiBo542 and chrysopine plasmid pTiChry5 (table 2). Directly adjacent to the RB are three genes (AgrTiEU6_20 - AgrTiEU6_22, SacBCD) for an ABC transport system, which components have 79%, 87% and 82% similarities at amino acid level with the components LecBCD of the succinamopine/leucinopine import system from agropine pTiBo542 and chrysopine pTiChry5, but with undetectable or only low similarity with those of the nopaline import system. A gene encoding a periplasmic substrate binding protein can be found close by (AgrTiEU6_28, SacA), which shows significant similarity to the leucinopine/succinamopine binding proteins from pTiBo542 and pTiChry5 (LecA, 70% similarity at amino acid level), but which has much less similarity to the corresponding nopaline binding protein NocT. Although it was thought that flavin-containing opine NAD(P)H-dependent dehydrogenase activities were encoded by two genes in the nopaline Ti plasmid called *noxA* and *noxB* (Zanker et al. 1994), recently evidence was provided that in fact such activity requires a small third gene called *noxC*, which provides an [2Fe-2S] iron-sulfur cluster to the enzyme (Watanabe et al. 2015). The enzymes encoded by genes AgrTiEU6_24/25/26 (*sacGFE*) show about 40% similarity with NoxACB at amino acid level. Remarkably, a fourth gene is present AgrTiEU6_23 (*sacH*) which encodes a protein which shows much less similarity with that encoded by *noxB*, but still has a similar architecture as NoxB and which shows about 50% similarity with the proteins encoded by the *lecH* genes from pTiBo542 and pTiChry5 at amino acid level. The *lecH* gene together with genes *lecG* and *lecF* encodes the dehydrogenase necessary for leucinopine/succinamopine catabolism. Both in AgrTiEU6_23 (*sacH*) and in AgrTiEU6_26 (*sacE*) the signature for FAD-binding can be found. Such a cluster of four genes with two "*noxB*" components was previously found in *Bradyrhizobium japonicum*, where it was shown that two different ABC complexes can be formed, both with catalytic activity (Watanabe et al. 2015). It is likely, therefore, that also in strains carrying the pTiEU6 plasmid two different complexes are formed for the catabolism of different opines including succinamopine and leucinopine.

Table 2. Relatedness among the predicted gene products of pTiEU6 opine catabolism region and those of other Ti-plasmids.

locus_tag	gene	pTiC58	pTiBo542	pTiChry5	pTiAch5
AgrTiEU6_20	*sacD*	*nocP* 33%	*lecD* 79%	*lecD* 80%	-
AgrTiEU6_21	*sacC*	-	*lecC* 87%	*lecC* 87%	-
AgrTiEU6_22	*sacB*	-	*lecB* 82%	*lecB* 82%	-
AgrTiEU6_23	*sacH*	-	*lecH* 50%	*lecH* 49%	-
AgrTiEU6_24	*sacG*	*noxA* 42%	*lecG* 47%	*lecG* 48%	*ooxA* 39%
AgrTiEU6_25	*sacF*	-	*lecF* 54%	*lecF* 56%	-
AgrTiEU6_26	*sacE*	*noxB* 43%	-	-	*ooxB* 41%
AgrTiEU6_27	*sacR*	*nocR* 33%	*lecR* 59%	*lecR* 61%	-
AgrTiEU6_28	*sacA*	-	*lecA* 70%	*lecA* 70%	-

Values of relatedness are percentage of identities for the pairs of amino acid sequences. The dash means no similarities.

Directly adjacent to the RB of pTiEU6 gene AgrTiEU6_19 is present, predicted to encode an NAD-dependent opine dehydrogenase (*odh*). The ODH encoded by AgrTiEU6_19 shows strong similarity to the ODH encoded by nopaline Ti plasmids (64% similarity with that of pTiC58 at amino acid level). Its role in opine catabolism has not been clarified as yet. Nopaline catabolism liberates α-ketoglutaric acid and arginine and the nopaline Ti plasmid has genes encoded on the Ti plasmid for arginine and ornithine catabolism. These genes are absent from pTiEU6, which has a gene encoding an asparaginase instead (AgrTiEU6_32), which can cope with the asparagine liberated when succinamopine is split into asparagine and α-ketoglutaric acid. Genes (AgrTiEU6_30 and 31) encoding a couple of proteins with the signature of hydantoinase/oxoprolinase A and B are present adjacent to the gene encoding asparaginase. Such genes are also present in the nopaline catabolism area of the nopaline Ti-plasmids. These proteins could open lactam rings and may therefore be necessary for the catabolism of the lactam opines SAL and SOL in the case of pTiEU6. A LysR type transcriptional regulator (AgrTiEU6_27) is present within the area and this may control expression of the succinamopine import and catabolism genes. Finally this region of the Ti plasmid contains another set of genes for an ABC transporter (AgrTiEU6_34 - AgrTiEU6_37), most related to putrescine (*potABCD*) transport systems, which is accompanied by a gene for a second LysR type transcriptional regulator (AgrTiEU6_33), which shares 30% similarity with AgrTiEU6_27. Finally this segment of the Ti-plasmid ends with a gene *mmgE* (AgrTiEU6_38), encoding an enzyme with similarity to 2-methylcitrate dehydratase, which is necessary for catabolism of (potentially toxic) propionyl-CoA propionate, which may be formed when catabolizing leucinopine.

Discussion

Traditionally, Ti-plasmids are classified on the basis of the opines that are generated in the infected tumors. In this study we reported the first complete sequence of a succinamopine-type Ti-plasmid pTiEU6 and its close relatedness to nopaline Ti-plasmids. Six nopaline Ti-plasmids from strains from different geographical origins were newly sequenced in order to trace the evolutionary history of nopaline and succinamopine Ti-plasmids.

The succinamopine Ti-plasmid pTiEU6 shares with nopaline Ti-plasmids a number of highly conserved regions including most of T-region, replication, conjugation, and virulence region. Based on whole plasmid sequence comparison, pTiEU6 is most similar to pTiT37 and pTiSAKURA, which form a subgroup A of nopaline Ti-plasmids. They only differ considerably in the region related to the biosynthesis and catabolism of succinamopine or nopaline, even though the structures of these opines are similar. The area between *rep* and *tra* operon is conserved in pTiEU6 and subgroup A, but is different from that of the subgroup B nopaline Ti-plasmids. Similarly, the *tra* operon of pTiEU6 was found to be highly homologous to that in subgroup A nopaline Ti-plasmids, but this operon only shares 77% similarity with that of pTiC58. Moreover, pTiEU6 and subgroup A nopaline Ti-plasmids carry the same transposable elements. Taken together, it can be assumed that pTiEU6 is derived from a common ancestor with subgroup A nopaline Ti-plasmids (pTiT37 and pTiSAKURA) in which the opine catabolic region and the right part of the T-region was replaced by that of an unknown succinamopine plasmid by DNA recombination.

By comparison of the complete Ti-plasmid sequences, the nopaline Ti-pasmids can be classified into two subgroups, which we have called A and B and which differ in an area located between the *rep* and *tra* genes. The insertion of transposable elements can be used to follow the evolutionary path. Although the nopaline Ti-plasmids studied all contain (the remnants of) three transposable elements (IS21, IS110 and ISRel8) at an identical site, only the subgroup A plasmids specifically contain an copy of IS1182, whereas only the subgroup B plasmids carry an copy of IS692 near ISRel8. Interesting, plasmid pTIC58 lacks both of them. In general, major evolutionary changes only take place once. We therefore suggest that pTiC58 represents the most ancestral of the nopaline Ti-plasmids in our study. It has all conserved sectors and lacks both transposable elements IS1182 and IS692. The insertion event of IS1182 defines subgroup A, which has kept the *6a* and *6b* genes as in C58, but has undergone a major exchange (recombination event R1, fig. 4B) with another plasmid in the segment between replication *rep* and conjugation *tra* operons. One or more further recombination events R2 (fig. 4B) led to exchanges of the right part of the T-region and the adjacent catabolic region and resulted in the formation of the succinamopine Ti-plasmid. The acquirement of IS692 marks the evolution of subgroup B. Acquisition of IS692 was followed by a putative DNA recombination event R3 (fig. 4B) with another Ti-plasmid leading to a replacement of the area occupying genes 6a and 6b with a new segment containing a very different gene 6b and gene *3'*. The result of this event can be seen in all the subgroup B plasmids except pTiKerr27, which has otherwise undergone severe changes at three positions probably coinciding with transposition events. Although pTiC58 seems to be the most closely related to the ancestral nopaline Ti-plasmid, it has undergone changes that are not seen in any of the other nopaline Ti-plasmids. By a putative recombination event R4 part of the *tra* region of pTiC58 is deviant from that of the other nopaline Ti-plasmids and also it has a unique deletion in the T-region removing gene f as previously reported by Broer et al. 1995. During the preparation of our manuscript, the characterization of the closely related nopaline Ti-plasmids pTiC5.7/pTiC6.5, which are almost identical to plasmid pTiKerr108 from our study, was reported and the relatedness of nopaline Ti-plasmids to nopaline catabolic plasmids was determined leading to the hypothesis that the nopaline Ti-plasmids are derived from such catabolic plasmids (Kuzmanović and Puławska 2019). Interestingly, the *tra* genes of the pTiC58 are more similar to those of nopaline catabolic plasmid pAtK84b (88% nucleotide identity) than to those of the other nopaline Ti-plasmids in our study, which would be in line with pTiC58 being the most ancestral of the studied plasmids and recombination event R4 occurring after formation of the ancestral nopaline Ti-plasmid, but before the transposition events leading to subgroups A and B.

Conclusion

The sequence of the succinamopine Ti plasmid confirms its close relatedness to the nopaline Ti plasmids. By a detailed comparison to the full sequences of the two published and 6 newly sequenced nopaline Ti-plasmids we can infer that the succinamopine Ti plasmid is evolutionary derived from the group A nopaline Ti plasmids (pTiSAKURA, pTiT37), which is characterized by the presence of a copy of IS1182 at a specific location and a recombination event leading to a new set of genes between the *tra* genes and the *rep* genes in comparison to other nopaline Ti-plasmids. In pTiEU6 the area occupying the right part of the T-region (including gene *6b* and *nos*) and the cluster of genes involved in nopaline import and

catabolism up to the conserved *trb* cluster involved in conjugation was replaced by a new set of genes that is overall more similar to that present in chrysopine and agropine Ti plasmids and involved in succinamopine/leucinopine synthesis, import and catabolism. It is likely that this was brought about by a recombination event between a group A nopaline Ti plasmid and another type of Ti plasmid. Precedent for recombination events between Ti plasmids has been seen in *in vitro* experiments, when introducing octopine Ti plasmids into nopaline strains (Hooykaas et al. 1980). Initially this leads to co-integrate formation, but eventually such co-integrates may disintegrate leading to new types of (Ti) plasmids.

Acknowledgements

References

Babicki S, et al. 2016. Heatmapper: web-enabled heat mapping for all. *Nucleic Acids Res.* 44: W147-W153.

Blundy KS, White J, Firmin JL, Hepburn AG. 1986. Characterisation of the T-region of the SAP-type Ti-plasmid pTiAT181: identification of a gene involved in SAP synthesis. *Mol Gen Genet.* 202: 62-67.

Broer I, Droge-Laser W, Barker RF, Neumann K, Klipp W, Puhler A. 1995. Identification of the *Agrobacterium tumefaciens* C58 T-DNA genes *e* and *f* and their impact on crown gall tumour formation. Plant Mol Biol. 27: 41-57.

Chang CC, Chen CM. 1983. Evidence for the presence of N2-(1, 3-dicarboxypropyl)-L-amino acids in crown-gall tumors induced by *Agrobacterium tumefaciens* strains 181 and EU6. *FEBS Lett.* 162: 432-435.

Chilton WS, Tempé J, Matzke M, Chilton MD. 1984. Succinamopine: a new crown gall opine. *J Bacteriol.* 157: 357-362.

Chilton WS, Hood E, Chilton MD. 1985. Absolute stereochemistry of leucinopine, a crown gall opine. *Phytochemistry.* 24: 221-224.

Darling AE, Mau B, Perna NT. 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PloS One.* 5: e11147.

Dhillon BK, et al.. 2015. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. *Nucleic Acids Res.* 43:W104-W108.

Drevet C, Brasileiro ACM, Jouanin L. 1994. Oncogene arrangement in a shooty strain of *Agrobacterium tumefaciens. Plant Mol Biol.* 25: 83-90.

Delcher AL, Bratke KA, Powers EC, Salzberg SL. 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics.* 23: 673-679.

Galens K, et al.. 2011. The IGS standard operating procedure for automated prokaryotic annotation. *Stand Genomic Sci.* 4: 244.

Gardner SN, Slezak T, Hall BG. 2015. kSNP3. 0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. *Bioinformatics.* 31: 2877-2878.

Goodner B, et al.. 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. *Science.* 294: 2323-2328.

Henkel CV, den Dulk-Ras A, Zhang X, Hooykaas PJJ. 2014. Genome sequence of the octopine-type *Agrobacterium tumefaciens* strain Ach5. *Genome Announc.* 2: e00225-14.

Hooykaas PJJ, Klapwijk PM, Nuti MP, Schilperoort RA, Rörsch A. 1977. Transfer of the *Agrobacterium tumefaciens* Ti plasmid to avirulent *agrobacteria* and to *Rhizobium* ex planta. *Microbiology.* 98: 477-484.

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

Huerta-Cepas J, et al.. 2017. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol Biol Evol* 34: 2115-2122.

Huang YY, et al.. 2015. Complete genome sequence of *Agrobacterium tumefaciens* Ach5. *Genome Announc.* 3: e00570-15.

Konstantinidis KT, Tiedje JM. 2005. Genomic insights that advance the species definition for prokaryotes. *Proc. Natl. Acad. S*ci. USA 102: 2567-2572.

Kumar S, Stecher G, Tamura K. 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 33: 1870-1874.

Kurtz S, et al.. 2004. Versatile and open software for comparing large genomes. *Genome Biol.* 5: R12.

Kuzmanović N, Puławska J. 2019. Evolutionary Relatedness and Classification of Tumour-Inducing and Opine-Catabolic Plasmids in Three *Rhizobium rhizogenes* Strains Isolated from the Same Crown Gall Tumour. *Genome Bio. Evol.*

Liu W, et al.. 2015. IBS: an illustrator for the presentation and visualization of biological sequences. *Bioinformatics.* 31: 3359-3361.

Mankin SL, et al.. 2007. Disarming and sequencing of *Agrobacterium rhizogenes* strain K599 (NCPPB2659) plasmid pRi2659. *In Vitro Cell Dev Biol Plant.* 43: 521-535.

Mindlin S, et al.. 2001. Mercury resistance transposons of Gram-negative environmental bacteria and their classification. *Res Microbiol.* 152: 811-822.

Montoya AL, Chilton MD, Gordon MP, Sciaky D, Nester EW. 1977. Octopine and nopaline metabolism in *Agrobacterium tumefaciens* and crown gall tumor cells: role of plasmid genes. *J Bacteriol.* 129: 101-107.

Moriguchi K, et al.. 2001. The complete nucleotide sequence of a plant root-inducing (Ri) plasmid indicates its chimeric structure and evolutionary relationship between tumor-inducing (Ti) and symbiotic (Sym) plasmids in rhizobiaceae1. *J Mol Biol* 307: 771-784.

Oger P, Reich C, Olsen GJ, Farrand SF. 2001. Complete nucleotide sequence and analysis of pTiBo542: what genomics tells us about structure and evolution of plasmids in the family *Rhizobiaceae. Plasmid.* 45: 169-170.

Otten L, De Ruffray P. 1994. *Agrobacterium vitis* nopaline Ti plasmid pTiAB4: relationship to other Ti plasmids and T-DNA structure. *Mol Gen Genet.* 245, 493-505.

Otten L, et al.. 1999. Sequence and functional analysis of the left-hand part of the T-region from the nopaline-type Ti plasmid, pTiC58. *Plant Mol Biol.* 41: 765-776.

Overbeek R, et al.. 2013. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res.* 42: D206-D214.

Paulus F, Ride M, Otten L. 1989. Distribution of two *Agrobacterium tumefaciens* insertion elements in natural isolates: evidence for stable association between Ti plasmids and their bacterial hosts. *Mol Gen Genet.* 219: 145-152.

Rambaut A. 2016. FigTree version 1.4. 0 Available at http://tree.bio.ed.ac.uk/software/figtree..

Sciaky D, Montoya AL, Chilton MD. 1978. Fingerprints of *Agrobacterium* Ti plasmids. *Plasmid.* 1: 238-253.

Shao S, Zhang X, van Heusden GPH., Hooykaas PJJ. 2018. Complete sequence of the tumor-inducing plasmid pTiChry5 from the hypervirulent *Agrobacterium tumefaciens* strain Chry5. *Plasmid.* 96: 1-6.

Slater SC, et al.. 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. *J Bacteriol.* 191: 2501-2511.

Stothard P, Wishart DS. 2004. Circular genome visualization and exploration using CGView. *Bioinformatics.* 21: 537-539.

Suzuki K, et al.. 2000. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid. *Gene.* 242: 331-336.

Watanabe S, Sueda R, Fukumori F, Watanabe Y. 2015. Characterization of flavin-containing opine dehydrogenase from bacteria. *PloS One.* 10: e0138434.

Wintersinger JA, Wasmuth JD. 2014. Kablammo: an interactive, web-based BLAST results visualizer. *Bioinformatics.* 31: 1305-1306.

Wood DW, et al.. 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science.* 294: 2317-2323.

Zanker H, et al.. 1994. Octopine and nopaline oxidases from Ti plasmids of *Agrobacterium tumefaciens*: molecular analysis, relationship, and functional characterization. *J Bacteriol.* 176: 4511-4517.

Zhu J, et al.. 2000. The bases of crown gall tumorigenesis. *J Bacteriol.* 182: 3885-3895.

Supplementary Table S1. The alignments of *tra* and *trb* operon at amino acid level.

Identity	TraG	TraD	TraC	TraA	TraF	TraB	TraH	TraM	TraR	McpA
pTiC58	82%	77%	76%	60%	67%	95%	99%	100%	99%	99%
pTiSAKURA	99%	99%	99%	98%	94%	99%	98%	99%	100%	98%
RSF1010	–	–	–	30%	–	–	–	–	–	–
RP4	24%	–	–	–	35%	–	–	–	–	–

Identity	TrbI	TrbH	TrbG	TrbF	TrbL	TrbK	TrbJ	TrbE	TrbD	TrbC	TrbB	TraI
pTiEU6 vs pTiC58	88%	87%	95%	95%	94%	81%	94%	95%	100%	98%	100%	100%
pTiEU6 vs pTiSAKURA	88%	87%	98%	95%	95%	78%	94%	95%	100%	99%	99%	99%
pTiC58 vs pTiSAKURA	95%	99%	98%	100%	100%	97%	100%	99%	100%	99%	100%	99%

FIG. S1. Schematic comparison of pTiC58 and subgroup A nopaline Ti-plasmids (A) and pTiC58 and subgroup B nopaline Ti-plasmids (B). ProgressiveMauve (Darling et al., 2010) was used to generate pairwise alignments with default parameters and the single nucleotide based similarities were expressed as height of the panels from blank to filled to represent 0-100% similarity. Different colored blocks represent conserved regions.
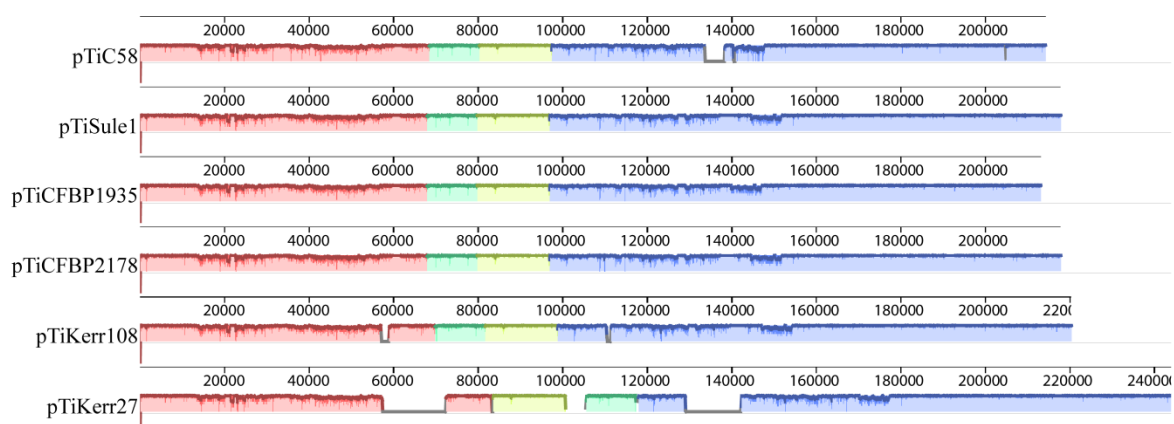
A
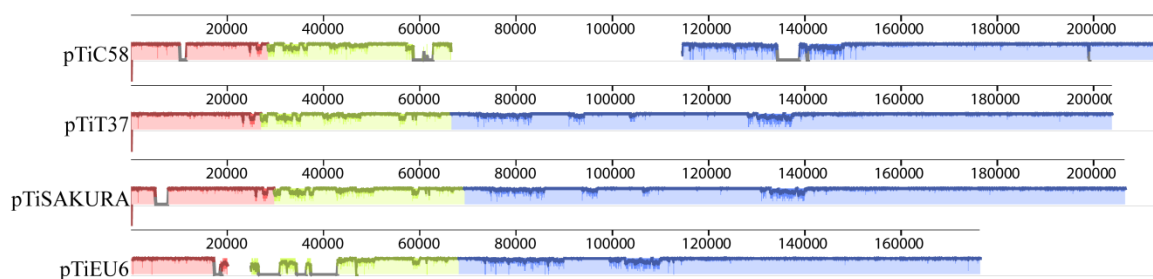
B

FIG. S2. A, Schematic comparison of the *vir* and *rep* operons of pTiEU6 with those of pTiC58. The alignment was calculated using the Blast 2.2.31+ algorithm (NCBI) at nucleotide level and visualized by the web-tool Kablammo (Wintersinger and Wasmuth, 2015). The arrows represent genes in pTiEU6 at the top and in pTiC58 at the bottom, respectively. The virulence genes are indicated by pattern. B, Schematic genetic organization of leucinopine catabolism region in pTiChry5, succinamopine catabolism region in pTiEU6 and nopaline catabolism region in pTiC58. The corresponding homologous genes are indicated by pattern.
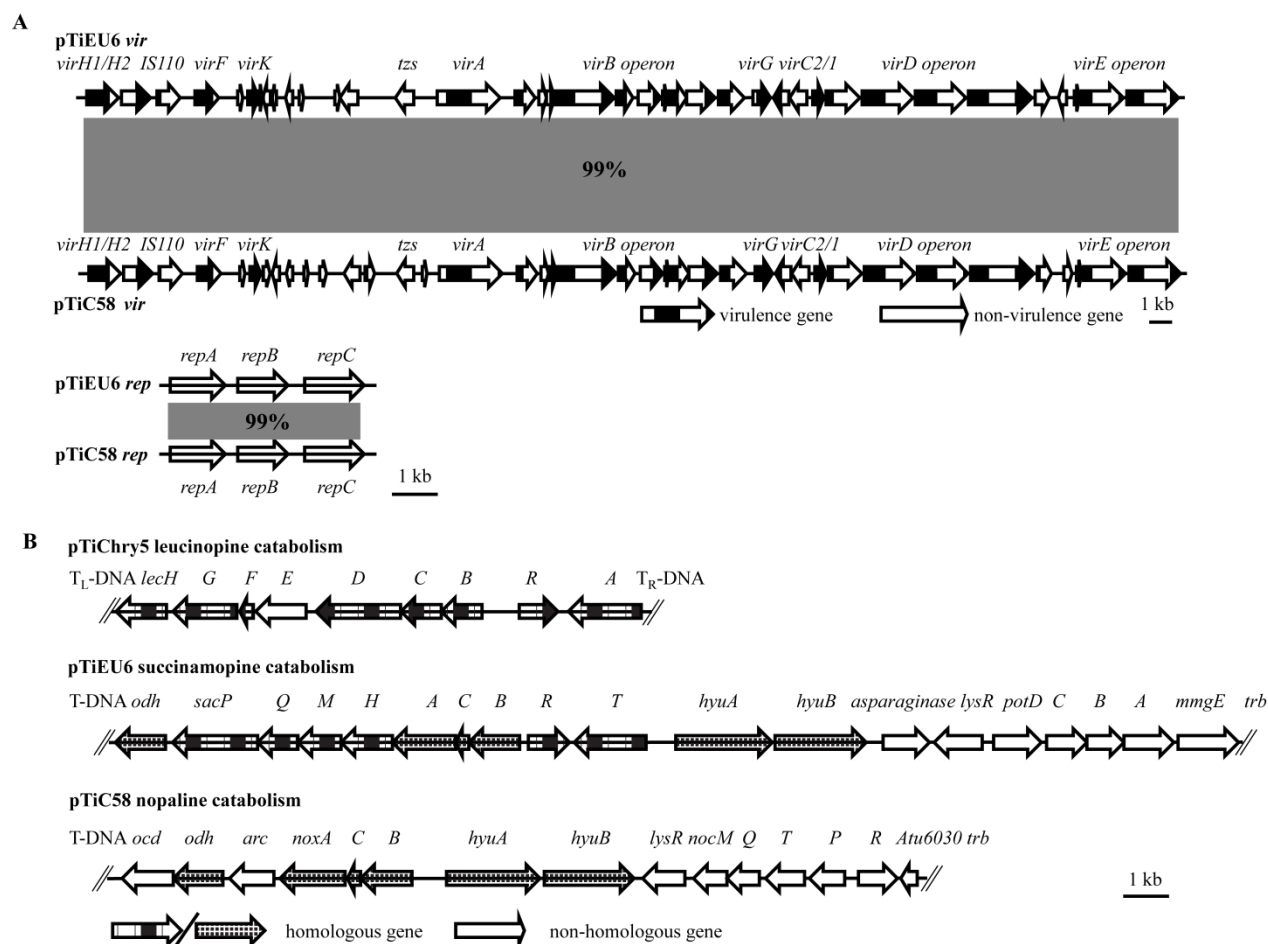
Summary

Summary

Agrobacterium tumefaciens, a gram-negative plant pathogen belonging to the family *Rhizobiaceae*, is the causative agent of crown gall disease and plays critical roles in worldwide agriculture. It induces tumor formation in plants by transferring a segment of its tumor-inducing plasmid (Ti-plasmid) to the plant cell. This transferred DNA (T-DNA) contains genes involved in the synthesis of auxin, cytokinin and opines resulting in uncontrolled cell proliferation and production of opines. Under laboratory conditions it is also able to transform other eukaryotes such as the yeast *Saccharomyces cerevisiae* and other fungi (Bundock et al., 1995; de Groot et al., 1998). Hence, the many experimental tools available for research on *S. cerevisiae* can be applied to obtain new insights into the mechanism of *Agrobacterium*-mediated transformation (AMT). These insights can contribute to the development of more efficient transformation methods for use in the plant field. In this respect more insight in the role of host factors in AMT is important and useful.

The yeast *S. cerevisiae* has emerged as an excellent host model to study the mechanism of AMT (Bundock et al., 1995; van Attikum et al., 2001, 2003; Ohmine et al., 2016; Hooykaas et al., 2018). The bacterial side of AMT has been investigated rather well but the host factors involved in this process remain obscure. Previous research in our group revealed that the efficiency of AMT of the yeast *ada2Δ* deletion mutant is higher than that of the wild type strain when the T-DNA is able to integrate by homologous recombination (Soltani, 2009). In Chapter 2, it was confirmed that the efficiency of AMT of yeast cells lacking *ADA2* was increased when the T-DNA can be integrated by homologous recombination (HR). In addition, we showed that the AMT efficiency of the *ada2Δ* mutant was also highly increased when the T-DNA has to be integrated by non-homologous end joining (NHEJ). Deficiency of *ADA2* confers hypersensitivity to DNA damaging agents such as methyl methanesulfonate (MMS) and hydroxyurea (HU). It has been reported that the homologous protein (ADA2b) from *Drosophila* and *Arabidopsis* is involved in the DNA damage response (Qi et al., 2004 and Lai et al., 2018). Even though the link between the increased AMT efficiency and impaired double strand break (DSB) repair in *ada2Δ* deletion mutant remains obscure, one hypothesis is that the enhanced sensitivity to DNA damage leads to the formation of more DSBs as a landing pad for T-DNA integration. In general, DNA damage events may occur at random chromosomal sites. Consistent with this, AMT efficiency by NHEJ increased 50 fold while that by HR was only 2 fold higher in the absence of *ADA2* compared to wild-type strain. A genome-wide rescue screen indicated that overexpression of *SFP1* in the *ada2Δ* background enhances the resistance to DNA damaging agents and attenuate the increased AMT efficiency to wildtype levels in the *ada2Δ* deletion mutant. That the increased AMT efficiency of the *ada2Δ* deletion mutant is the consequence of incorrect regulation of unknown genes involved in the integration process, however, could not be ruled out. Furthermore, considering the role of *ADA2* in defining the boundary of heterochromatin regions, the increased AMT efficiency of the *ada2Δ* deletion mutant may be the result of expression of the selection gene at integration sites which may otherwise be silenced and could therefore not be detected.

Upon AMT a considerable amount of T-DNA is present as circular extrachromosomal structures rather than integrated into chromosomal DNA (Singer et al., 2012). Such structures are present at a higher frequency in the *ada2Δ* deletion mutant than in the wild type strain. In Chapter 3, by analysis of the junctions formed between the LB and RB, T-circles isolated from *Nicotiana benthamiana* or the yeast *ada2Δ* deletion mutant were characterized. These

extrachromosomal T-DNA structures contained various formats such as perfect fusions of LB and RB, binary plasmid sequences, filler host DNA sequences and T-DNA truncations. The sequence of the LB- RB junctions resembled that of repaired DNA double-strand breaks sites and therefore the formation of extrachromosomal T-circles was proposed to exploit the same DNA repair mechanisms as those used to explain T-DNA integration. Collectively, our results lead to a conceptually simple mechanism for T-DNA junction formation: HR in yeast and TMEJ mediated by *Pol θ* in plants. The distinct ability of *Pol θ* to extend minimally paired 3′ ends may provide a possible explanation for the existence of filler sequences to a large extent.

The *Agrobacterium* virulence protein VirD2 is potentially involved in the T-DNA integration process. To expand our knowledge about the function of VirD2 during AMT, in Chapter 4 we exploited the auxin-inducible degron (AID) system (Morawska and Ulrich, 2013) to specifically degrade VirD2 after translocation into the host nucleus. The efficiency of AMT of yeast was strongly reduced upon auxin-induced degradation of VirD2. With respect to plant cells, the size of the tumors was reduced and the number of calli decreased strikingly, thus indicating T-DNA integration was greatly diminished. Combined with the results that T-DNA could be delivered into the plant nucleus successfully with the guide of AID-tagged VirD2, we can conclude that VirD2 plays a role inside the host nucleus, possibly in the T-DNA integration process. Previously it has been shown that when the VirD2 omega domain was mutated, the efficiency of transient transformation was reduced to 20-25% and stable transformation was almost eliminated (only 1-3% of that of wild type) (Narasimhulu et al., 1996). Besides, VirD2 was reported to interact with yeast histones (Wolterink-van Loo et al., 2015) and probably plant histone proteins or other chromatin structural proteins (Gelvin and Kim, 2007). In this scenario, T-DNA may target DSB sites with the help of VirD2 and VirD2 could interact with host DNA repair proteins to serve as a prelude to guide T-DNA to chromatin and facilitate the integration process.

Many different *Agrobacterium* strains have been isolated from nature and they may differ in their virulence on specific plants. They are often classified on the basis of the opines, tumor-specific metabolites, that are found in the tumors and serve as a nitrogen and carbon source for the bacterium (Dessaux et al., 1993). Several Ti-plasmids have been completely sequenced, such as those from both nopaline and octopine-*Agrobacterium* strains. On the other hand, until now Ti-plasmids from chrysopine and succinomopine *Agrobacterium* strains have not been sequenced.

In Chapter 5, we report the complete nucleotide sequence of pTiChry5, a Ti-plasmid isolated from a chrysopine *Agrobacterium* strain. A detailed alignment between pTiChry5 and agropine plasmid pTiBo542 was made in order to more comprehensively understand their evolutionary relationship. In general, these two plasmids have a similar order of functional areas with different interruptions by transposable elements and losses or gains of small sets of genes. Based on similarity, unexpectedly, only roughly one half of the plasmid containing *rep, vir, tra, acc*, TL-DNA is very similar in the two plasmids, whereas the other half (*trb*, Amadori opine catabolism up to the chrysopine synthase gene), is much more different suggesting that these plasmids are chimeric due to recombination with other opine-related plasmids. The presence of two *repABC* units in both these plasmids is in line with this hypothesis. Especially, the development of new chrysopine profiles may have conferred evolutionary advantage on their host bacteria in some specific environments.

Based on phylogenetic analysis, the evolutionary relatedness of nopaline-type Ti-plasmids was elucidated in Chapter 6. We elucidated the sequence of the succinamopine Ti plasmid pTiEU6. Analysis of the sequence confirms its close relatedness to the nopaline Ti-plasmids in line with the fact that succinamopine has a similar chemical structure as nopaline. By a detailed comparison to the sequences of 8 nopaline Ti plasmids (of which 6 were newly sequenced) we could infer that the succinamopine Ti plasmid is evolutionary derived from the group A nopaline Ti plasmids (pTiSAKURA, pTiT37), which is characterized by the presence of a copy of IS1182 at a specific location and a set of *tra* genes that has only 77% similarity to that of other nopaline Ti plasmids. In pTiEU6 the area occupying the right part of the T-region (including gene *6b* and *nos*) and the cluster of genes involved in nopaline import and catabolism up to the conserved *trb* cluster involved in conjugation was replaced by a new set of genes that is overall more similar to that present in chrysopine and agropine Ti plasmids and involved in succinamopine/leucinopine synthesis, import and catabolism. It is likely that this was brought about by a recombination event between a group A nopaline Ti plasmid and another type of Ti plasmid.

Reference

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

De Groot, M. J., Bundock, P., Hooykaas, P. J. J., & Beijersbergen, A. G. (1998). *Agrobacterium* tumefaciens-mediated transformation of filamentous fungi. *Nature biotechnology*, 16, 839.

Dessaux, Y., Petit, A., Tempe, J., 1993. Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry*, 34, 31-38.

Gelvin, S. B., & Kim, S. I. (2007). Effect of chromatin upon Agrobacterium T-DNA integration and transgene expression. *Biochimica et Biophysica Acta (BBA)-Gene Structure and Expression*, 1769, 410-421.

Hooykaas, P. J. J., van Heusden, G. P. H., Niu, X., Roushan, M. R., Soltani, J., Zhang, X., & van der Zaal, B. J. (2018). *Agrobacterium*-Mediated Transformation of Yeast and Fungi, 349-374.

Lai, J., Jiang, J., Wu, Q., Mao, N., Han, D., Hu, H., & Yang, C. (2018). The transcriptional coactivator ADA2b recruits a structural maintenance protein to double-strand breaks during DNA repair in plants. *Plant physiology*, pp-00123.

Morawska, M., & Ulrich, H. D. (2013). An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast*, 30, 341-351.

Narasimhulu, S. B., Deng, X. B., Sarria, R., & Gelvin, S. B. (1996). Early transcription of *Agrobacterium* T-DNA genes in tobacco and maize. *The Plant Cell*, 8, 873-886.

Ohmine, Y., Kiyokawa, K., Yunoki, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2018). Successful Transfer of a Model T-DNA Plasmid to *E. coli* Revealed Its Dependence on Recipient RecA and the Preference of VirD2 Relaxase for Eukaryotes Rather Than Bacteria as Recipients. *Frontiers in Microbiology*, 9.

Ohmine, Y., Satoh, Y., Kiyokawa, K., Yamamoto, S., Moriguchi, K., & Suzuki, K. (2016). DNA repair genes *RAD52* and *SRS2*, a cell wall synthesis regulator gene SMI1, and the membrane sterol synthesis scaffold gene *ERG28* are important in efficient *Agrobacterium*-mediated yeast transformation with chromosomal T-DNA. *BMC microbiology*, **16:** 58.

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. *Proceedings of the National Academy of Sciences*, 90,11538–42.

Qi, D., Larsson, J., & Mannervik, M. (2004). Drosophila Ada2b is required for viability and normal histone H3 acetylation. *Molecular and cellular biology*, 24, 8080-8089.

Rolloos, M., Dohmen, M. H., Hooykaas, P. J. J., & van der Zaal, B. J. (2014). Involvement of Rad52 in T-DNA circle formation during *Agrobacterium* tumefaciens-mediated transformation of *Saccharomyces cerevisiae. Molecular microbiology*, 91, 1240-1251.

Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. (2012). Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant physiology*, pp-112.

Soltani, J. (2009) Host genes involved in *Agrobacterium*-mediated transformation. PhD thesis, Institute of Biology, Leiden University, The Netherlands.

Van Attikum, H., Bundock, P., Hooykaas, P. J. J. (2001). Non-homologous end-joining proteins are required for *Agrobacterium* T-DNA integration. *The EMBO journal*, 20, 6550-6558.

Van Attikum, H., Hooykaas, P. J. J. (2003). Genetic requirements for the targeted integration of *Agrobacterium* T-DNA in Saccharomyces cerevisiae. *Nucleic acids research*, 31, 826-832.

Van Kregten, M., de Pater, B. S., Romeijn, R., van Schendel, R., Hooykaas, P. J. J., Tijsterman, M. (2016). T-DNA integration in plants results from polymerase-θ-mediated DNA repair. *Nature plants*, 2, 16164.

Wolterink-van Loo, S., Ayala, A. A. E., Hooykaas, P. J. J., & van Heusden, G. P. H. (2015). Interaction of the *Agrobacterium tumefaciens* virulence protein VirD2 with histones. *Microbiology*, 161, 401-410.

Samenvatting

Agrobacterium tumefaciens, een gram-negatieve plantenpathogene bacterie behorende tot de familie *Rhizobiaceae,* is de veroorzaker van de plantenziekte wortelhalsknobbel (engels: crown gall). Deze bacterie induceert tumorvorming in planten door een segment van zijn tumor inducerend plasmide (Ti-plasmide) over te dragen naar de plantencel. Dit overgedragen DNA (T-DNA) bevat genen die betrokken zijn bij de synthese van auxine, cytokinine en opines resulterend in ongecontroleerde deling van cellen waarin opines worden geproduceerd. Deze natuurlijke vorm van genetische manipulatie wordt nu wereldwijd toegepast voor de genetische modificatie van planten voor onderzoek en in de plantenbiotechnologie. Onder laboratoriumomstandigheden kan *Agrobacterium* ook andere eukaryoten transformeren, zoals de gist *Saccharomyces cerevisiae* en andere schimmels (Bundock et al. 1995; de Groot et al. 1998). Het modelorganisme *S. cerevisiae* wordt gebruikt om nieuwe inzichten te verkrijgen in het mechanisme van *Agrobacterium*-gemedieerde transformatie (AMT). Terwijl de bacteriële kant van AMT vrij goed onderzocht is, is van de gastheerfactoren die bij dit proces betrokken zijn, nog veel onbekend.

In eerder onderzoek in onze groep is gevonden dat de efficiëntie waarmee de gist *ada2Δ* deletiemutant getransformeerd wordt door *Agrobacterium* hoger is dan die van de overeenkomstige wildtype stam. In deze experimenten werd een T-DNA gebruikt dat in het gistgenoom kan integreren via homologe recombinatie (Soltani, 2009). In hoofdstuk 2 werd bevestigd dat de efficiëntie van AMT van gistcellen zonder *ADA2* was verhoogd, wanneer het T-DNA kan integreren via homologe recombinatie (HR). Daarna toonden we aan dat de AMT-efficiëntie van de *ada2Δ* mutant relatief nog sterker was verhoogd, wanneer het T-DNA alleen kan integreren via niet-homologe recombinatie (NHEJ). Een tekort aan *ADA2* maakt gist overgevoelig voor DNA-beschadigende middelen zoals methylmethaansulfonaat (MMS) en hydroxyurea (HU). Over het homologe eiwit (ADA2b) van *Drosophila* en *Arabidopsis* is gerapporteerd dat het betrokken is bij de respons op DNA-schade (Qi et al., 2004 en Lai et al., 2018). Wij vermoeden dan ook dat de verhoogde gevoeligheid voor DNA-schade in de *ada2Δ* -mutant een gevolg is van een DNA reparatiedefect en dat het (gemiddeld) grotere aantal DNA beschadigingen in de cel meer kansen biedt voor T-DNA integratie. Een genoombrede zoektocht naar factoren die bij overexpressie in de *ada2Δ* mutant de weerstand van deze mutant tegen DNA-beschadigende middelen verhogen, leverde niet alleen het *ADA2* gen zelf op, maar ook *SFP1*. Overexpressie van *SFP1* in de *ada2Δ* -mutant leidde ook weer tot een verlaging van de frequentie van AMT.

Bij AMT kan vroeg in het transformatieproces een aanzienlijke hoeveelheid T-DNA aanwezig zijn in circulaire extrachromosomale structuren (Singer et al., 2012). Dergelijke structuren zijn aanwezig met een hogere frequentie in de *ada2Δ* -mutant dan in de wildtype stam. In hoofdstuk 3 werden dergelijke T-cirkels geïsoleerd uit de plant *Nicotiana benthamiana* en uit gist en vervolgens gekarakteriseerd. Deze T-cirkels waren veelal gevormd door perfecte fusie van LB en RB, maar bevatten soms ook binaire vector plasmide-sequenties of sequenties afkomstig uit het genoom van de gastheer.

Het *Agrobacterium* virulentie-eiwit VirD2 is nodig voor overdracht van het T-DNA naar de gastheercel via het Type 4 secretie systeem (T4SS) en vervolgens voor translocatie naar de celkern, maar mogelijk ook daarna nog bij het T-DNA-integratieproces. Om onze kennis hierover uit te breiden, hebben we in hoofdstuk 4 het auxine-induceerbare degron (AID) -

systeem (Morawska en Ulrich, 2013) gebruikt om VirD2 specifiek af te breken na translocatie in de gastheerkern. De efficiëntie van AMT van gist ging sterk omlaag wanneer auxine werd toegevoegd tijdens de co-cultivatie. Ook de transformatie van plantencellen bleek sterk verlaagd te zijn. De gevormde tumoren waren kleiner en het aantal calli was opvallend afgenomen, hetgeen erop duidde dat de T-DNA-integratie sterk was verminderd. Dit duidt erop dat VirD2 een rol speelt in de celkern, mogelijk in het T-DNA-integratieproces.

Veel verschillende *Agrobacterium* stammen zijn geïsoleerd uit de natuur en ze kunnen verschillen in hun virulentie voor specifieke planten. Ze worden vaak geclassificeerd op basis van de opines, tumorspecifieke metabolieten, die in de tumoren worden gevormd en als stikstof- en koolstofbron voor de bacterie kunnen dienen (Dessaux et al., 1993). Verschillende typen Ti-plasmiden zijn al volledig gesequenst, zoals de nopaline, de octopine en de agropine Ti-plasmiden. Echter de DNA volgorde van chrysopine en succinamopine Ti-plasmiden was nog onbekend. In hoofdstuk 5 beschrijven we de complete nucleotidesequentie van pTiChry5, een chrysopine-Ti-plasmide. Een gedetailleerde vergelijking tussen pTiChry5 en agropine Ti-plasmide pTiBo542 leerde dat deze twee plasmiden ruwweg een vergelijkbare volgorde van functionele gebieden hebben met verschillende onderbrekingen door onder andere transposons. Onverwacht bleek toch slechts de helft van het plasmide (de gebieden met *rep*, *vir, tra, acc,* en TL-DNA) zeer sterk verwant, terwijl de DNA volgorde in de andere helft (*trb*, Amadori opine katabole genen tot het chrysopine synthase gen), veel minder verwantschap suggereert. We concluderen daarom dat deze plasmiden chimeren zijn en mogelijk ontstaan zijn door recombinatie tussen twee plasmiden die gerelateerd zijn aan verschillende opines. De aanwezigheid van twee *repABC* replicator-gebieden in beide plasmiden ondersteunt deze hypothese. Het ontstaan van nieuwe opine profielen heeft wellicht evolutionair voordeel opgeleverd aan hun gastheerbacteriën in sommige specifieke niches.

In hoofdstuk 6 hebben we de sequentie van het succinamopine Ti-plasmide pTiEU6 opgehelderd. Analyse van de sequentie bevestigde zijn nauwe verwantschap met de nopaline Ti-plasmiden. Door een gedetailleerde vergelijking met de sequenties van 8 nopaline Ti-plasmiden (waarvan er 6 nieuw werden gesequenst) konden we afleiden dat het succinamopine Ti-plasmide evolutionair is afgeleid van de groep A nopaline Ti-plasmiden (pTiSAKURA, pTiT37), die wordt gekenmerkt door de aanwezigheid van een kopie van IS1182 op een specifieke locatie en een reeks *tra*-genen die slechts 77% gelijkenis vertoont met die van andere nopaline Ti-plasmiden. In pTiEU6 werd echter het rechter deel van het T-gebied (inclusief gen 6b en *nos*) en het cluster van genen betrokken bij nopaline-import en katabolisme tot het geconserveerde *trb*-cluster betrokken bij conjugatie vervangen door een nieuwe set genen. Deze genen zijn over het algemeen meer vergelijkbaar met de overeenkomstige genen die aanwezig zijn in de chrysopine en agropine Ti-plasmiden en betrokken zijn bij de synthese, import en katabolisme van succinamopine / leucinopine. Het is waarschijnlijk dat dit werd veroorzaakt door een recombinatiegebeurtenis tussen een groep A nopaline Ti-plasmide en een ander type Ti-plasmide.

Referenties

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

De Groot, M. J., Bundock, P., Hooykaas, P. J. J., & Beijersbergen, A. G. (1998). *Agrobacterium* tumefaciens-mediated transformation of filamentous fungi. *Nature biotechnology*, 16, 839.

Dessaux, Y., Petit, A., Tempe, J., 1993. Chemistry and biochemistry of opines, chemical mediators of parasitism. *Phytochemistry*, 34, 31-38.

Lai, J., Jiang, J., Wu, Q., Mao, N., Han, D., Hu, H., & Yang, C. (2018). The transcriptional coactivator ADA2b recruits a structural maintenance protein to double-strand breaks during DNA repair in plants. *Plant physiology*, pp-00123.

Morawska, M., & Ulrich, H. D. (2013). An expanded tool kit for the auxin-inducible degron system in budding yeast. *Yeast*, 30, 341-351.

Qi, D., Larsson, J., & Mannervik, M. (2004). Drosophila Ada2b is required for viability and normal histone H3 acetylation. *Molecular and cellular biology*, 24, 8080-8089.

Singer, K., Shiboleth, Y. M., Li, J., & Tzfira, T. (2012). Formation of complex extrachromosomal T-DNA structures in *Agrobacterium*-infected plants. *Plant physiology*, pp-112.

Soltani, J. (2009) Host genes involved in *Agrobacterium*-mediated transformation. PhD thesis, Institute of Biology, Leiden University, The Netherlands.

Curriculum vitae

Shuai Shao was born on the 17th of August 1989 in Nanchang, Jiangxi province, China. He grew up there and received his high school diploma at the affiliated middle school of Jiangxi Normal University. In 2006, he started his bachelor study at Jiangnan University. His major was Food Science and Technology and he obtained his Bachelor degree in 2010. One year later, he decided to pursue his academic study at East China University of Science and Technology under the supervision of Prof. Dr. Qiyao Wang. From 2011 to 2014, his work focused on multiple phylogenetic analyses of *Edwardsiella* species and their type III and VI secretion virulence proteins. In 2014, he received his Master degree and then started his PhD research under the supervision of Dr. G. Paul H. van Heusden in the laboratory of Prof. Dr. Paul J. J. Hooykaas at the Institute of Biology, Leiden University, with a financial support by China Scholarship Council. As described in this thesis, his PhD work concerned the involvement of host and bacterial factors in *Agrobacterium*-mediated transformation.