Involvement of host and bacterial factors in Agrobacterium-mediated transformation

Shao, S.

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.

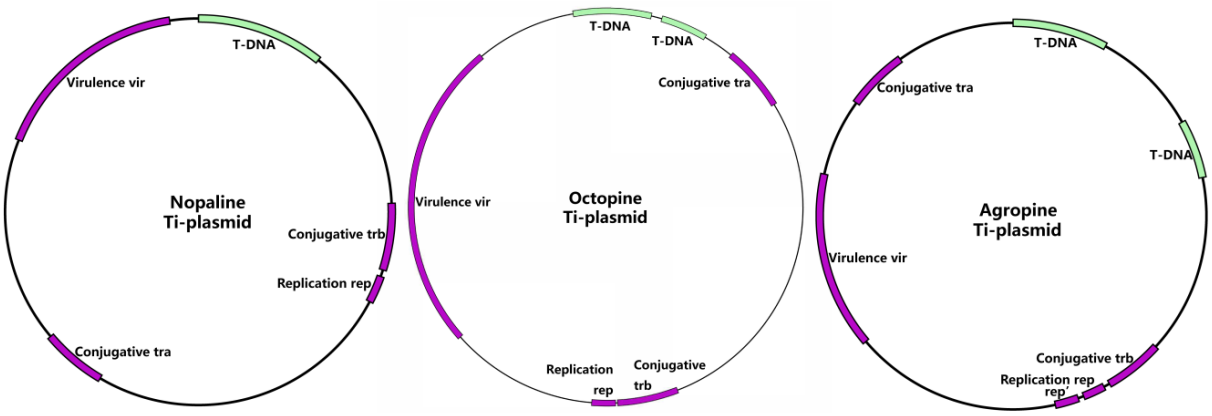
Table 3. All known complete sequences of Ti-plasmids (as of April 2019).

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.

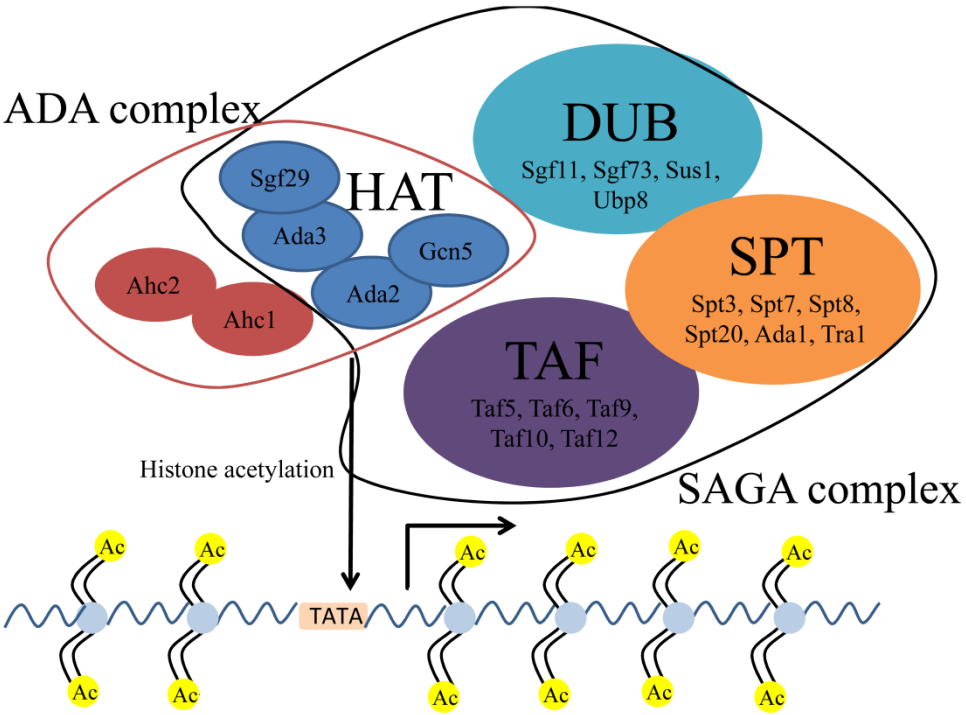


Figure 2. Schematic presentation of the structure of the SAGA and ADA complexes. Different colors are used to represent the different modules, such as the histone acetyltransferase (HAT) module, the histone deubiquitinase (DUB) module, the core structural (SPT) module, and the activator-binding (TAF) 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-Garcidueñas, 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.