ABSTRACT Agrobacterium tumefaciens is a plant pathogen with the capacity to deliver a segment of oncogenic DNA carried on a large plasmid called the tumor-inducing or Ti plasmid to susceptible plant cells. A. tumefaciens belongs to the class Alphaproteobacteria, whose members include other plant pathogens (Agrobacterium rhizogenes), plant and insect symbionts (Rhizobium spp. and Wolbachia spp., respectively), human pathogens (Brucella spp., Bartonella spp., Rickettsia spp.), and nonpathogens (Caulobacter crescentus, Rhodobacter sphaeroides). Many species of Alphaproteobacteria carry large plasmids ranging in size from ~100 kb to nearly 2 Mb. These large replicons typically code for functions essential for cell physiology, pathogenesis, or symbiosis. Most of these elements rely on a conserved gene cassette termed repABC for replication and partitioning, and maintenance at only one or a few copies per cell (1). The subject of this review is the ~200-kb Ti plasmids carried by infectious strains of A. tumefaciens. We will summarize the features of this plasmid as a representative of the repABC family of megaplasmids. We will also describe novel features of this plasmid that enable A. tumefaciens cells to incite tumor formation in plants, sense and respond to an array of plant host and bacterial signal molecules, and maintain and disseminate the plasmid among populations of agrobacteria. At the end of this review, we will describe how this natural genetic engineer has been adapted to spawn an entire industry of plant biotechnology and review its potential for use in future therapeutic applications of plant and nonplant species.

NOMENCLATURE AND TYPES OF Ti PLASMIDS
Agrobacterium species that are pathogenic on plants, including Agrobacterium tumefaciens, A. vitis, A. rubi, and A. rhizogenes, all carry megaplasmids. By contrast, nonpathogenic strains either lack these plasmids entirely or carry mutant forms of plasmids. A strict requirement of the Ti plasmid for virulence was established through mutational analyses and by a demonstration that the introduction of Ti plasmids into Rhizobium or Phyllobacterium spp. converts these nonpathogenic species into tumor-inducing pathogens (2, 3). Ti plasmids induce a disease called crown gall, which is typified by the formation of undifferentiated plant tumors at the plant crown (the subterranean-to-aerial transition zone). The related root-inducing or Ri megaplasmids carried by A. rhizogenes instead induce hairy root disease, which is typified by the formation of entangled masses of roots at the infection site (4).

As discussed in more detail below, Ti plasmid-carrying strains of agrobacteria induce not only plant tumor formation but also the production of various amino acid and sugar phosphate derivatives termed opines. The transformed plant cell secretes opines, which can then be taken up and catabolized for use as a food source by the infecting bacterium. The Ti plasmid carries the genes for opine synthesis by plant cells as well as the corresponding catabolism genes. Ti plasmids traditionally have been classified by opine type, and here we will retain this classification scheme with a focus mainly on the two best-characterized Ti plasmids designated as the octopine and nopaline types. Several octopine-type (pTiA6, B6, Ach5, 15955, R10) and nopaline-type (pTiC58, pTi37) Ti plasmids have been extensively characterized, and many have been sequenced at this time. All Ti plasmids code for functions associated with (i) plasmid replication and maintenance, (ii) conjugative transfer, (iii) virulence,
(iv) opine utilization, and (v) sensory perception of exogenous signals released by the plant host and neighboring agrobacterial cells at the site of infection (see reference 5). Genes encoding each of these functions are generally clustered on the Ti plasmid, with the exception of two spatially distinct regions, the virulence or vir region and the transfer-DNA or T-DNA required for infection of plants, and the tra and trb regions required for conjugative plasmid transfer (Fig. 1). The following sections summarize our current understanding of the Ti plasmid-encoded functions.

**Ti PLASMID MAINTENANCE**

*repABC*

The Ti plasmids belong to the repABC family of replicons, whose members are widely distributed among many species of *Alphaproteobacteria* (1). This replicon family is composed mainly of extrachromosomal plasmids and some secondary chromosomes. The repABC cassette was identified nearly 25 years ago as essential for replication and partitioning of an octopine-type Ti plasmid (6). The cassette is composed of three genes (*repA, repB, repC*), a cis-acting partitioning site (*parS*), and an origin of replication *oriV* (Fig. 2). *repA* and *repB* code for a partitioning system, and *repC* encodes the replication initiator protein (1, 7). The RepA/RepB partitioning system closely resembles the ParA/ParB systems harbored by many plasmids, phages, and chromosomes in diverse bacterial species (8, 9). These proteins function together with *parS* sequence to ensure faithful plasmid partitioning during cell division (119). The RepC initiator protein is uniquely associated with repABC cassettes in *Alphaproteobacteria* and is unrelated in sequence to other well-characterized Rep proteins (1). RepC binds a putative *oriV* sequence in vitro (10), and cloned *repC* genes support autonomous replication of associated replicons (10, 11). The in vivo
findings strongly indicate that RepC functions as the initiator protein and further suggest that the oriV replication origin is located within the repC gene.

Plasmid Partitioning

The RepA/RepB partitioning system resembles that of the well-characterized ParA/ParB systems and likely has a similar mechanism of action (Fig. 2) (1, 119). In brief, these partitioning systems consist of three components: a weak NTPase (ParA), a DNA-binding protein (ParB), and the cis-acting parS sequence to which ParB binds. RepA is a member of the ParA/MinD/Soj superfamily of ATPases, and, as shown for many ParA homologs, the RepA proteins of pTiR10 and other Ti plasmids negatively autoregulate rep gene expression by binding an upstream promoter (12, 13). In the Par systems, ADP-bound ParA is active in autorepression via a DNA-binding activity of its N-terminal domain. In contrast, the ATP-bound form of ParA is dimeric and fulfills another function through cooperative but nonspecific binding to DNA. ParA-ATP binds DNA dynamically, oscillating over the nucleoid. This oscillatory behavior is thought to influence the positioning of associated DNA in the predivisional cell or DNA translocation during cell segregation (14, 15, 16). This dynamic activity and the capacity of ParA proteins to form filaments or spiral structures are reminiscent of eukaryotic cytoskeletal or motor proteins and suggest that the partitioning system resembles a form of bacterial mitosis (17). ParA also binds ParB. ParB specifically binds the centromere-like parS sequence via a helix-turn-helix motif located in its C terminus. Thus, ParA foci or filamentous structures are postulated to bind cognate ParB-parS partitioning complexes to provide a pulling or ratcheting force necessary for spatial localization of the associated plasmids in predivisional cells.

Although most characterized parAB loci are auto-regulated, they usually function separately from the replication and copy number control system of the hosting replicon. In contrast, the repABC genes associated with the alphaproteobacterial megaplasmids, e.g., symbiotic plasmid p42d and octopine-type Ti plasmids, are co-transcribed and constitute a single operon (Fig. 2) (1). This genetic context places expression of the partitioning genes and the replication initiator protein under the same regulatory control. Therefore, RepA-ADP-mediated negative autoregulation is partly responsible for the low copy number of these plasmids. Additionally, in the repABC cassette of Ti plasmids, a small gene located between repB and repC encodes a small nontranslocated countertranscribed RNA that downregulates expression
of repC (18). This short transcript maintains the Ti plasmid at a low copy number and also functions to inhibit replication of coresident replicons bearing the same repABC cassette in a mechanism termed incompatibility (1). The repABC cassettes have two other noteworthy features (Fig. 2). First, the DNA sequence GANTC is over-represented in the putative replication origin and in the promoter of the countertranscribed repE RNA. These sequences are potential substrates for the DNA methylase CcrM, which in Caulobacter crescentus contributes to cell cycle timing (1). Second, a binding site for another C. crescentus cell cycle factor, the two-component response regulator CtrA, was identified upstream of the repABC promoter region of plasmid pTiR10 (1). These findings warrant further studies defining the contribution of DNA methylation, CrtA, and possibly other factors in temporal control of Ti plasmid replication during the cell cycle.

Plasmid Replication
As mentioned above, the RepC-type proteins have been found only in Alphaproteobacteria, and the presumptive RepC-binding target, oriV, resides within repC. In contrast to other plasmid replication systems in which the initiator protein binds directly repeated DNA sequences called iterons to target the replication machinery to the origin of replication, repABC origins lack such repeat sequences. The repC genes do, however, contain AT-rich sequences of ∼150 nucleotides near the middle of their sequences, which is commonly encountered with other replication origins (Fig. 2) (1). The RepC protein from pTiR10 was purified and shown to bind a region of imperfect dyad symmetry within the AT-rich segment, lending support to the proposal that this sequence conforms to the origin of replication (10). RepC has two domains, an N-terminal domain (NTD) that exhibits DNA-binding activity and a C-terminal domain (CTD) whose function is currently unknown. The NTD has structural similarity to members of the DnaD family of replication proteins found in low GC-content Gram-positive bacteria as well as to members of the MarR family of transcriptional factors (1). While the molecular details are unknown, it is reasonable to predict that RepC-oriV binding serves to recruit other replication proteins to oriV to build the replisome.

Another interesting feature of the RepC proteins encoded by plasmids pTiR10 and p42D is that they appear to function only in cis (10, 11). That is, they act only on the origin of replication embedded within the repC gene itself and not on origins located within repABC cassettes of coresident plasmids. Consistent with this activity, overproduction of RepC results in an increase in the copy number of the plasmid encoding the protein but has no effect on the copy numbers of co-resident repABC replicons. Some alphaproteobacterial species can have as many as six repABC family replicons (1). In such cells, the cis-acting function of RepC proteins could serve to ensure the fidelity of replication and copy number control of cognate replicons without interfering effects on heterologous replicons.

Ti Plasmid-Encoded Type IV Secretion Systems
Ti plasmids carry genes for elaboration of two DNA conjugation systems, one (Tra/Trb) responsible for conjugal transfer of the Ti plasmid and the second (VirB/ VirD4) dedicated to the delivery of a segment of the Ti plasmid called the T-DNA as well as several effector proteins to plant cells during the infection process (Fig. 1) (5). Studies of the Ti plasmid transfer system have focused mainly on defining the functions of regulatory factors in controlling tra/trb gene expression (for example, see reference 19), whereas investigations of the T-DNA transfer system have explored the biogenesis, mechanism of action, and architecture of the conjugation channel and associated T pilus (see references 20, 21). All bacterial conjugation systems are now grouped together with a set of ancestrally related translocation systems in pathogenic bacteria that are dedicated to the delivery of effector proteins into eukaryotic cells during the course of infection. Collectively, these translocation systems are called the type IV secretion systems (T4SSs) (22). The VirB/VirD4 system has emerged as a paradigm for this T4SS superfamily (20), and, here, we will use the nomenclature associated with this system when discussing the functions of the individual subunits. The following sections will summarize general features of T4SSs, as well as available mechanistic and structural information about the Tra/Trb and VirB/VirD4 systems. The overall mechanism of type IV secretion can be viewed as three biochemically distinct but spatially and temporally coupled reactions: (i) substrate processing as a translocation-competent transfer intermediate, (ii) substrate docking with a protein termed the type IV coupling protein (T4CP), and (iii) substrate transfer through the envelope-spanning translocation channel (Fig. 3).

DNA and Effector Protein Substrate Processing
The tra and vir genes, respectively, code for proteins responsible for processing the Ti plasmid and T-DNA. The overall conjugative DNA-processing reaction is as follows. The relaxase is the principal enzyme required
FIGURE 3  Schematic showing steps of type IV secretion, as presented for the Ti-encoded VirB/VirD4 transfer system. Step I: the DNA transfer and replication (Dtr) proteins bind the oriT-like right border repeat sequence (Ti plasmid, red squares flanking T-DNA) to form the relaxosome. VirD2 relaxase nicks the T strand, which is then unwound from the template strand of the pTi plasmid. Step II: ParA-like VirC1 and VirD2, and probably other factors, mediate binding of the VirD2-T-strand transfer intermediate with the VirD4 substrate receptor or type IV coupling protein (T4CP). Step III: The transfer intermediate is translocated across the cell envelope through a secretion channel composed of the VirD4 T4CP and the VirB mating pair formation (Mpf) proteins. Effector proteins, e.g., VirE2, VirE3, VirF, also dock with VirD4 and then are delivered independently of the T-DNA through the secretion channel. Independently of VirD4, the VirB proteins also assemble into a conjugative pilus, which is used to establish contact with a susceptible target cell. IM, inner membrane; P, periplasm; OM, outer membrane. doi:10.1128/microbiolspec.PLAS-0010-2013.f3
for DNA processing. This enzyme binds a cognate origin of transfer (oriT) sequence and nicks the DNA strand (T strand) destined for transfer. Upon nicking, the relaxase remains covalently bound to the 5’ end of the T strand. Additional auxiliary or accessory factors, termed DNA transfer and replication (Dtr) proteins, also bind at the oriT sequence to form the relaxosome. The Dtr factors enhance relaxase binding and cleavage at oriT, and they can also participate in docking of the DNA substrate with the substrate receptor for the cognate T4SS channel. Upon nicking, the T strand is unwound from the template strand, the Dtr factors are dissociated from the relaxase-T-strand particle, and the transfer intermediate is delivered to and through the transfer channel (23, 24).

Recent phylogenetic studies resulted in classification of the relaxases into 8 different mobilization (MOB) groups (25). The Ti-encoded TraA relaxase is closely related to the RSF1010 MobA relaxase and therefore grouped in the MOBQ family. Accordingly, the oriT-binding target of TraA is closely related to that of RSF1010 (5). By contrast, VirD2 is grouped in the MOBP family and its binding targets, the oriT-like border sequences that flank the T-DNA, resemble the oriT sequences of IncP plasmids (26). Although TraA and VirD2 cleave their DNA substrates by similar catalytic mechanisms, differences in their primary sequences and with the cognate Dtr factors likely confer specificity of the Ti plasmid as a substrate for the Tra/Trb T4SS and T-DNA as a substrate for the VirB/VirD4 system.

VirD2 binding and nicking at T-DNA border sequences is enhanced by the accessory factors VirD1, VirC1, and VirC2 (Fig. 3). VirD1 is important for VirD2 nicking on the supercoiled, double-stranded plasmid. VirC1 and VirC2 bind a sequence termed overdrive located immediately adjacent to the right border repeat sequences of octopine-type Ti plasmids (27, 28). This binding reaction stimulates T-DNA processing and results in accumulation of many copies of free VirD2-T-strand transfer intermediates in a cell (29). Interestingly, VirC1 is a member of the ParA family of ATPases that, as discussed above, mediate partitioning of chromosomes and plasmids during cell division (5). VirC1 was shown to localize at A. tumefaciens cell poles, to recruit the VirD2-T-strand complex to the cell poles, and to interact with VirD4, the substrate receptor for the VirB/ D4 T4SS (29). Taken together, these findings prompted a model that A. tumefaciens adapted an ancestral Par-like function for the novel purposes of (i) stimulating a conjugal DNA-processing reaction and (ii) promoting DNA substrate docking with a cognate T4SS receptor. Both activities potentially mediate transfer of many copies of T-DNA to susceptible plant host cells, presumably for enhanced probability of infection (29).

How the T-DNA and Ti plasmid substrates engage with their respective T4SSs is presently not known at a molecular level, but some general features of this interaction have been defined. As can be surmised from the above, the protein components associated with the translocated DNA carry the recognition signals for the substrate-T4SS docking reaction. The VirC accessory factors facilitate contact between the VirD2 relaxase and the VirD4 receptor (29), but the VirD2 relaxase also possesses a translocation sequence that contributes to this interaction. VirD2 was shown to carry a translocation sequence at its C terminus by use of the Cre recombinase reporter assay for translocation (CRAfT) (30, 31). In this assay, full-length or fragments of protein substrates, e.g., VirD2, are fused at their N termini to Cre recombinase and translocation is monitored to a reporter bacterial or plant cell carrying a lox cassette whereby Cre-mediated excision confers a reporter activity, e.g., antibiotic resistance. Studies of the VirD2 translocation signal showed that a cluster of positively charged Arg residues was important for Cre transfer, leading to a proposal that C-terminally mediated ionic interactions are important for docking with the VirD4 receptor (31).

The VirB/VirD4 T4SS also translocates several effector proteins, including VirE2, VirE3, and VirF to target cells (Fig. 3) (30, 32). These proteins do not interact with the VirD2-T-strand intermediate in agrobacteria, but instead they are independently translocated through the T4SS into the plant cell. Functions of these translocated proteins are discussed briefly below. These effectors also carry positively charged C-terminal domains that are required for substrate-VirD4 engagement (33). Interestingly, the VirB/VirD4 T4SS is also capable of translocating plasmid RSF1010, a non-self-transmissible plasmid of the MOBQ family, to target cells (34, 35). In contrast to VirD2, the MobA relaxase carries two internal motifs designated as translocation signals 1 and 2 (TS1 and TS2), each of which can mediate transfer of the relaxase-T-strand complex through the T4SS (36). RSF1010 transfer also requires the accessory factor MobB, which is thought to function analogously to VirC1 and VirC2 in promoting DNA substrate-receptor docking. Thus, at least two distinct types of translocation signals, a positively charged C-terminal motif or internal motifs of unspecified sequence composition, can mediate transfer of DNA and protein substrates through the VirB/VirD4 T4SS.
The T4CP Receptor

All conjugation systems and nearly all T4SS effector translocation systems have a substrate receptor that is ancestrally related to the Ti plasmid-encoded VirD4 subunit (37). Receptor activities of these proteins have been demonstrated genetically (38, 39), through demonstration of relaxase-receptor interactions in vitro (40, 41), and with a ChIP-based, UV-cross-linking assay (42). By use of the latter assay, designated as transfer DNA immunoprecipitation (TrIP), cross-linkable interactions were identified between the translocating T-DNA substrate and components of the A. tumefaciens VirB/VirD4 T4SS. Confirming its role as the T-DNA receptor, VirD4 was shown to form a cross-linkable contact with the T-DNA substrate even in a strain lacking the VirB channel subunits (42). VirD4-like receptors are also termed type IV coupling proteins (T4CPs) because they functionally couple the DNA-processing and transfer reactions (43). T4CPs contain Walker A and B nucleotide-binding motifs, which are essential for nucleotide binding and hydrolysis, and mutations in these motifs abolish translocation indicating that one or more stages of transfer are energized by NTP hydrolysis. T4CPs are tethered to the inner membrane by an N-terminal membrane anchor sequence (37). An X-ray structure of the soluble, ∼50-kDa cytoplasmic domain of the TrwB T4CP encoded by the conjugative plasmid R388 revealed a globular hexameric assembly in which each subunit is composed of two distinct domains, a nucleotide-binding domain (NBD) and a 7-helix motif called the all-α-domain (AAD) that faces the cytoplasm (44, 45). The six TrwB protomers assemble to form a globular ring that is ∼110 Å in diameter and 90 Å in height, with a ∼20-Å-wide channel in the center that constrains to 8 Å at the cytoplasmic pole. The N terminus of the TrwB hexamer spans the inner membrane. TrwB undergoes conformational changes in the central channel upon substrate binding and hydrolysis (46), suggesting that T4CPs might act as motor proteins during secretion, but precisely how the T4CP interacts with substrates and energizes substrate transfer is not yet defined.

The T4SS Channel

The TraG T4CP interacts with a translocation channel composed of the Trb proteins to mediate conjugative transfer of the Ti plasmid to agrobacterial recipient cells. Similarly, VirD4 interacts with the VirB channel to coordinate T-DNA and protein substrate transfer to plant cells. Both channels are assembled from at least 11 subunits whose stoichiometries for the most part are unknown (Fig. 4). The Ti plasmid-encoded Trb proteins were assigned the same names as their closest homologs in the database, the Trb proteins encoded by plasmid FIGURE 4 Genetic organization of the A. tumefaciens Ti plasmid-encoded virB and trb operons. The virB genes and some of the known functions of the encoded products are presented at the top. This T4SS is closely related in operon organization and subunit composition to a T4SS encoded by the E. coli conjugative plasmid pKM101. The Trb system is closely related in operon organization and subunit composition to a T4SS encoded by the E. coli conjugative plasmid RP4. Genes encoding protein homologs are identically color-coded. doi:10.1128/microbiolspec.PLAS-0010-2013.f4

A. tumefaciens pTi virB

pKM101 tra

korA

A. tumefaciens pTi trb

RP4 trb
20, 47). Phylogenetic analyses established that the VirB4 ATPase family is ancestrally related to the VirD4 superfamily (25), leading to a proposal that these subunits also function as homohexamers. VirB4-like TrwK was found to assemble as a homohexamer (52), but other VirB4 homologs exist in solution as monomers and dimers (53, 54), raising the possibility that the oligomeric state of these ATPases might vary depending on membrane association, ATP binding or hydrolysis, or interactions with other T4S machine subunits. VirB11 is a member of a family of ATPases termed the “traffic ATPases” (21). These ATPases are associated with Gram-negative bacterial type II, type III, type IV, and type VI secretion systems (20, 55). VirB11-like ATPases are peripheral inner membrane proteins, but soluble forms also exist, suggesting that these ATPases might exist in a dynamic equilibrium with the membrane. Electron microscopy (EM) visualization of VirB11 homologs showed hexameric rings of ∼100 to 120 Å in diameter (56, 57). The N- and C-terminal halves of the 6 protomers each form rings, giving rise to a double-stacked structure wherein the nucleotide-binding site is at the interface between the two domains. The EM studies have shown that the VirB11 hexamers undergo dynamic structural changes upon ATP binding and hydrolysis, although the functional importance of these transitions is not defined at this time (56, 57). The T4SS ATPases coordinate early steps of substrate transfer. As mentioned above, the A. tumefaciens VirD4 T4CP engages with the T-DNA substrate, as shown by TrIP (42). Further studies showed that VirD4 then delivers the DNA substrate to the VirB11 ATPase by a mechanism that does not require ATP hydrolysis by either subunit (42, 58). In the early studies, no evidence was obtained for interaction of the DNA substrate with the A. tumefaciens VirB4 ATPase (42), but the role of VirB4 subunits in substrate transfer might need to be revisited in light of more recent findings that VirB4 homologs associated with other T4SSs bind DNA in vitro (52, 53) and in vivo (53).

2. The inner membrane channel translocase: At this time, there is no structural information available for the inner membrane portion of the T4SS channel. In the VirB/VirD4 system, three integral...
membrane proteins, including VirB3, VirB6, and VirB8, are postulated to form this translocase on the basis of experimentally derived inner membrane topology models, protein-protein interaction data, and results of TrIP studies (Fig. 3 and 4) (20, 21, 42). VirB3 interacts with VirB4 (59) and might play a role in coordinating a biologically important interaction between this ATPase and other components of the translocase. VirB6 is unique among the VirB subunits in spanning the inner membrane 5 to 6 times, a topology common among subunits of inner membrane translocases and transporters (60). The TrIP studies placed VirB6 and VirB8 at an intermediate point in the postulated T-DNA translocation pathway, dispensable for DNA transfer to VirD4 and VirB11, but necessary for transfer to VirB2 and VirB9 (42). VirB6 and VirB8 functionally interact, as evidenced by the finding that null mutations in virB6 or virB8, respectively, block substrate transfer to VirB8 or VirB6. However, certain mutations in VirB6 were shown to permit T-DNA contacts with VirB6 but block contacts with VirB8 (60), giving rise to a proposal that the T-DNA substrate forms close contacts sequentially with VirB6 and then VirB8. Walker A mutations in each of the energetic components—VirD4, VirB4, and VirB11—block substrate transfer from VirB11 to VirB6 and VirB8, suggesting that ATP energy utilization is required for driving the DNA substrate through the inner membrane portion of the channel (42, 58). Additionally, other subunits including VirB7, VirB9, and VirB10 are required for this transfer step (58). These subunits form a stabilizing “core” complex that is envisaged to house and form critical contacts with the inner membrane translocase as summarized below.

3. The envelope-spanning core complex: In A. tumefaciens, the outer membrane lipoprotein VirB7, outer membrane-associated VirB9, and bitopic VirB10 stabilize each other as well as other VirB channel subunits (20), and evidence has been presented for assembly of these proteins as a ring-shaped complex now called the core complex (Fig. 3 and 4) (61). An analogous core complex composed of the corresponding VirB homologs encoded by the conjugative plasmid pKM101 was structurally resolved by cryo-electron microscopy (CryoEM) and a portion was further solved by X-ray crystallography (62, 63). The pKM101 core complex, composed of 14 copies each of VirB7-like TraN, VirB9-like TraO, and VirB10-like TraF, is a 1.05-MDa structure of 185 Å in width and height (62). It is composed of two layers (I and O layers) that form a double-walled ringlike structure. The I layer is composed of the N-terminal domains of TraO and TraF and forms a 55-Å-diameter ring at the inner membrane. The O layer is composed of TraN and the C-terminal domains of TraO and TraF and forms a main body and narrower cap with a central hole of 10 Å that is presumed to span the outer membrane. A crystal structure of the entire O layer further showed that TraF/VirB10 forms the outer membrane channel (63). Specifically, 14 copies of an α-helical domain termed “the antennae projection” or “AP” is thought to form the channel. In the assembled core complex, therefore, TraF/VirB10 subunits are predicted to span the entire cell envelope such that 14 N-terminal transmembrane helices form the ~55-Å inner membrane ring, a proline-rich region and β-barrel domain span the periplasm, and the AP forms the outer membrane pore (63, 64). A current model depicts this core complex as a structural scaffold for the translocation channel, wherein the ATPases are positioned at the base of the channel and VirB3, VirB6, and VirB8 are within the inner membrane ring. Other channel subunits, including the pilin subunit and a portion of VirB9, are postulated to form the distal portion of the channel within the core’s central chamber (63). Finally, a VirB2 pilus structure or the TraF/VirB10 AP forms the channel through which substrates pass across the outer membrane (21).

4. The conjugative pilus: T4SSs elaborate conjugative pili in addition to the translocation channel (Fig. 3). Although these two organelles likely assemble as a single supramolecular structure, we depict them as physically distinct to convey the idea that they fulfill distinct functions. For example, the extended pilus initiates contact with target cells, but it is completely dispensable for intercellular translocation (20). This was demonstrated genetically through the isolation of “uncoupling” mutations. Such mutations block pilus biogenesis without affecting substrate transfer or, conversely, block substrate transfer without affecting pilus biogenesis. The isolation of such mutations strongly indicates that pili extending from the cell surface function mainly or exclusively to initiate the donor-target cell contact. The conjugative pili are distinguished by width, length, and flexibility. For
example, F pili encoded by the *Escherichia coli* F plasmid are ~9 nm and flexible, and range in length up to 1 μm. By contrast, the VirB/VirD4 and Tra/Trb pili are both classified as P type. These pili are thicker (9 to 11 nm), more rigid, and shorter than F pili although length measurements are complicated by the fact that isolated pili are typically broken (66, 67). Plasmid RP4-encoded pili and probably those of the closely related Ti plasmid-encoded Tra/Trb system are thick and straight, whereas those elaborated by the *A. tumefaciens* VirB/VirD4 system are more flexible (68, 69). P pili are abundantly present in the extracellular milieu, often as bundles, and are rarely found associated with cells. Whether this is an artifact accompanying the preparation of cells for electron microscopy owing to their fragility or is a normally occurring process is not known (69, 70).

Pili are composed predominantly of a single pilin subunit. Both VirB2 and RP4-encoded TrbC are small (~7-kDa) proteins with hydrophilic N and C termini and two hydrophobic stretches of ~20 to 22 residues separated by a small central hydrophilic loop (71). These pilins are synthesized as pro-proteins with unusually long (~30 to 50 residues) leader peptides that are cleaved upon insertion into the inner membrane (71, 72). They are then processed further to yield a membrane pool of mature pilin subunits. Interestingly, VirB2 and TrbC from plasmid RP4 and probably also Ti plasmid-encoded TrbC are processed in part by a head-to-tail cyclization reaction whereby the N- and C-terminal residues of the pro-protein are covalently joined (23). Cyclization stabilizes the P-type pilins in the membrane and appears to be essential for pilus assembly.

Mature pilin monomers are thought to assemble as a pool in the inner membrane for use upon receipt of an unknown signal in building the conjugative pilus. Evidence has been presented for a role by the *A. tumefaciens* VirB4 ATPase in dislocation of VirB2 monomers from the inner membrane, and a pilus assembly pathway involving VirB4, VirB5, and VirB8 has been postulated (74, 75). VirB5 likely plays a critical role in pilus polymerization, as deduced from evidence that the conjugative pilus assemblies from its base (76) and that VirB5 subunits are located at the tip of the polymerized pilus (77). The X-ray structure of VirB5-like TraC from pKM101 revealed a 3-helix bundle flanked by a smaller globular part (78).

Further mutational analyses identified residues important for DNA transfer and binding of pilus-specific bacteriophages (78). Precisely how VirB5 interacts with VirB2 in the assembled pilus is unknown, but the present findings point to a role for the TraC/VirB5-like subunits in pilus nucleation and establishment of donor-recipient cell contacts.

**T-DNA Transfer to the Plant Cell**

In nature, the Ti plasmid-encoded Tra/Trb and VirB/VirD4 transfer systems are repressed in the absence of plant-derived inducing signals. However, when *A. tumefaciens* cells encounter wounded plant tissue, sensory perception of plant-derived molecules serves to activate the expression of the Ti plasmid-encoded *vir* regulon. An important outcome of *vir* gene induction is the translocation of T-DNA and protein substrates to susceptible plant cells and the resulting formation of crown gall tumors.

**Role of Cotransported Proteins in T-DNA Transfer and Plasmid Conjugation**

As mentioned above, *A. tumefaciens* cells use the VirB/VirD4 T4SS to deliver the VirD2-T-strand and effector proteins to the plant target cell. One of these effectors is VirE2, a single-stranded DNA-binding protein (SSB). Upon translocation of VirE2, the SSB binds cooperatively along the length of the T strand to generate a VirD2-T-strand-VirE2 particle termed the T complex (79). VirD2 and VirE2 in turn contribute in various ways to successful translocation of the T complex to the plant nucleus. Importantly, both proteins carry nuclear localization sequences (NLSs) that help guide the T complex to the nucleus through specific interactions with plant proteins. For example, both VirD2 and VirE2 were shown to interact with one or more members of the importin-α family, which function as adaptor molecules by interacting with NLS motifs in cargo proteins and with the nuclear shuttle protein importin-β to promote nuclear uptake (80, 81). VirD2 is also a phosphoprotein, and several additional plant proteins that interact with VirD2 may play roles in phosphorylation/dephosphorylation (82). For example, VirD2 interacts with and is phosphorylated by the cyclin-dependent kinase-activating kinase CAK2Ms (83). CAK2Ms also phosphorylate RNA polymerase II large subunits, which in turn recruits a TATA-box binding protein important for transcription initiation. The binding of VirD2 to a TATA-box binding protein led to a suggestion that
VirD2 phosphorylation may play an additional role in targeting T complexes to chromatin. VirD2 also interacts with plant cyclophilins, which contribute to the maintenance of protein folding, although the biological role of this interaction is not clear at this time (83, 84). Finally, intriguingly, VirE2 has been shown to form gated channels in black lipid membranes, leading to a proposal that VirE2 might promote entrance of the T complex across the plant cytoplasmic membrane through a channel-forming activity (85). While it is not immediately obvious how a protein could functionally as a channel for the T complex and an SSB that coats the length of the T strand, the findings are intriguing and warrant further study.

Once in the nucleus, T-DNA integrates into the plant nuclear genome by nonhomologous or “illegitimate” recombination. The T-DNA invades at nick gaps in the plant genome possibly generated as a consequence of active DNA replication. The invading ends of the single-stranded T-DNA are proposed to anneal via short regions of homology to the unnicked strand of the plant DNA. Once the ends of T-DNA are ligated to the target ends of plant DNA, the second strand of the T-DNA is replicated and annealed to the opposite strand of the plant DNA. VirD2 and VirE2 have been implicated in contributing to T-DNA integration but precisely how remains unclear (see reference 79).

**T-DNA Genes Expressed in Plant Cells**

Octopine-type Ti plasmids carry two T-DNA fragments designated T<sub>L</sub>-DNA and T<sub>R</sub>-DNA of 13 and 7.8 kb in length, respectively (see Fig. 1) (5). Once integrated into the plant nuclear genome, these T-DNAs encode 13 proteins with two main functions. One set of enzymes promotes the synthesis of two plant growth regulators, auxins and cytokinin zeatin. Production of these plant hormones results in a stimulation of cell division and a loss of cell growth control, ultimately leading to the formation of characteristic crown gall tumors. The second group of enzymes promotes the synthesis of novel amino acid and sugar derivatives called opines. Octopine-type T-DNA’s code for synthesis of octopine which is a reductive condensation product of pyruvate with arginine. Octopine synthases also promote condensation of pyruvate with other amino acids to produce lysopine, histopine, or octopinic acid. Octopine-type T-DNAs also code for synthesis of other opines, including mannopine, mannopinic acid, agropine and agropinic acid, whereas T-DNAs carried by other Ti plasmids code for nopalines, which are derived from α-ketoglutarate and arginine as well as other classes of opines (5).

**OPINE CATABOLISM AND THE “OPINE CONCEPT”**

Plants cannot metabolize opines and instead release them into the extracellular milieu. By contrast, the Ti plasmid carries opine catabolism genes that are responsible for the active transport of opines across the agrobacterial envelope and their degradation in the cytoplasm. Over 40 genes coding for at least 6 ATP-binding cassette-type permeases and 12 opine catabolic enzymes are responsible for opine uptake and degradation (5). The capacity of infecting agrobacteria to degrade opines for use as carbon and energy is central to the evolution of these bacteria as phytopathogens. The “opine concept” was developed to rationalize the finding that *A. tumefaciens* evolved as a pathogen by acquiring the ability to transfer DNA to plant cells. According to this concept, *A. tumefaciens* adapted an ancestral DNA conjugation system for interkingdom DNA transport specifically to incite synthesis of opines by the plant host. The cotransfer of oncogenes ensures that transformed plant cells proliferate, resulting in enhanced opine synthesis. The environment of the tumor thus is a rich chemical environment favorable for growth and propagation of the infecting *A. tumefaciens* (86, 87). It is noteworthy that a given *A. tumefaciens* strain catabolizes only those opines that it incites plant cells to synthesize. This is thought to ensure a selective advantage of the infecting bacterium over other *A. tumefaciens* strains that are present in the vicinity of the tumor.

**SIGNALING NETWORKS CONTROLLING Ti PLASMID FUNCTIONS**

Most of the Ti plasmid-encoded functions summarized above are activated only in response to sensory perception of extracellular signals. These signals originate from wounded and transformed plant tissue, as well as other *A. tumefaciens* cells at the infection site. Upon encountering wounded plant tissues, *A. tumefaciens* integrates these signals into a regulatory cascade that culminates in the following sequence of events: (i) T-DNA transfer and oncogenesis, (ii) production and utilization of opine nutrients, (iii) elevated Ti plasmid copy number, and (iv) dissemination of the Ti plasmid. This regulatory network has been extensively characterized and many of its features are known in molecular detail, as summarized below.

**Plant Wound Signals and VirA/VirG-Mediated Sensory Perception**

*A. tumefaciens* induces its virulence regulon upon sensory perception of various plant-derived signals. These signals
are present at the plant wound site and include specific classes of phenolic compounds, monosaccharides, low phosphorus levels, and acidic pH (88). Sensory perception is achieved through the Ti plasmid-encoded VirA/VirG two-component regulatory system (89, 90). VirA is a histidine sensor kinase that autophosphorylates at a conserved histidine residue and then transfers the phosphate group to a conserved aspartate residue on VirG. Phosphorylated VirG in turn activates transcription of the operons comprising the vir regulon that are responsible for T-DNA processing and translocation to plant cells. VirA senses the plant-derived signals, the most important of which are phenolic compounds that carry an ortho-methoxy group (91). The type of substitution at the para position distinguishes strong inducers such as acetylsyringone from weaker inducers such as ferulic acid and acetovanillone (89). A variety of monosaccharides, including glucose, galactose, arabinose, and the acidic sugars D-galacturonic acid and D-glucuronic acid, strongly enhance vir gene induction. The presence of these compounds is a general feature of most plant wounds and likely contributes to the extremely broad host range of A. tumefaciens. VirA functions as a homodimer with an N-terminal transmembrane domain and large C-terminal cytoplasmic domain. Recent studies suggest that a cytoplasmic linker domain of VirA interacts directly with phenolics and imparts specificity for phenolics (92). The periplasmic sugar-binding protein ChvE binds the monosaccharide sugars and the ChvE-sugar complexes interact with the periplasmic domain of VirA, thereby inducing a conformational change that increases VirA’s sensitivity to phenolic inducer molecules (90, 93, 94). The periplasmic domain of VirA also senses acidic pH, which is required for maximal induction of the vir genes, but the underlying mechanism is unknown (95). Phospho-VirG activates transcription of the vir genes by interacting with a cis-acting regulatory sequence (TNCAATTTGAAAPy) called the vir box located upstream of each of the vir promoters (88).

Opine Signals and Regulation of Opine Catabolism

Upon delivery and incorporation of T-DNA into the plant nuclear genome, expression of the T-DNA genes results in synthesis and release of opines into the milieu. The opines are taken up by agrobacterial cells for use not only as nutrients, but also as activators of gene expression (Fig. 5). Opines interact with opine-responsive transcriptional regulators that control the Ti plasmid-encoded opine catabolic functions (96). Opines derived from amino acids activate gene expression through binding a LysR-type transcriptional activator. For example, octopine binds LysR-like OccR (96). OccR positively regulates expression of the occ genes involved in octopine uptake and catabolism by inducing a bend in the DNA at the OccR-binding site. Octopine binding alters both the affinity of OccR for its target site and the angle of the DNA bend, establishing that octopine modulates OccR regulatory activity inside the bacterium (97, 98). Opines derived from sugars, e.g., agarcinopines and mannopine, function through binding a LacI-type repressor (99). These repressors limit catabolic gene function in the absence of opines but are inactivated in their presence allowing transcription and opine utilization.

N-Acylhomoserine Lactone Signals and Regulation of Ti Plasmid Conjugative Transfer

Besides regulating expression of opine catabolism genes, opines serve another important regulatory function. A regulatory cascade has been shown to activate Ti plasmid transfer under conditions of high cell density (100, 101, 102). This cascade initiates when A. tumefaciens cells import opines released from plant cells (Fig. 5). For example, binding of the octopine-OccR to the promoter located upstream of the occ operon induces expression of the octopine catabolism genes and a gene located at the 3′ end of the operon. This gene codes for the TraR transcriptional regulator, a protein related to LuxR shown over 20 years ago to regulate synthesis of quorum signals (QS) known as N-acylhomoserine lactones (AHLs) (5, 103). At low cell densities, quorum signals are present at a low concentration, whereas at high cell densities they accumulate in the surrounding environment and then passively diffuse back into the bacterial cell to activate transcription of a defined set of genes. In the case of A. tumefaciens, TraR responds to N-3-(oxo octonoyl)-t-homoserine lactone (3-oxo-C8-HSL), which is produced by the Ti plasmid-encoded TraI AHL synthase. This AHL acts in conjunction with TraR to activate transcription of the Ti plasmid-encoded tra genes as well as tral. Synthesis of TraR under conditions of high cell density thus creates a positive-feedback loop whereby the TraR-AHL complex induces synthesis of TraI, which in turn synthesizes more AHL (104). Ultimately, the regulatory cascade involving opine-mediated expression of traR and TraR-AHL-mediated expression of Ti plasmid transfer genes at high cell densities, results in enhanced Ti plasmid conjugative transfer to neighboring agrobacterial cells in the environment of the plant tumor. This signaling network is subject to
additional levels of feedback regulation, for example, through synthesis of a TraR antiactivator (TraM) (105), a quorum-quenching AHL lactonase (AiiB) (106), and a truncated form of TraR (TrlR) that dimerizes with and poisons TraR activity (107). It is generally considered that *A. tumefaciens* evolved this complex regulatory system to maximize the number of Ti plasmid-carrying bacterial cells, and hence the potential for interkingdom transmission of opine-encoding T-DNA, in the vicinity of susceptible host tissue (100, 101, 102).

**Plant Phenol, Opine, and AHL Regulatory Control of Ti Plasmid Replication**

Most Ti plasmids exist in single copy within a cell as a result of tight regulatory control of the repABC gene cluster. For plasmid pTiR10, this tight control is achieved by autorepression at one of the four upstream promoters (the P4 promoter) by RepA and RepB, and by transcriptional and posttranscriptional inhibition of repC expression by the countertranscript RNA RepE (see Fig. 2) (1). However, expression of the repABC operon, and therefore the Ti plasmid copy number, is also subject to regulation by two diffusible chemical signals (Fig. 5). First, as discussed above, the VirA/VirG two-component system senses specific classes of plant phenolic compounds to activate vir gene expression through binding of phospho-VirG to upstream vir boxes. Phospho-VirG also has been shown to bind a vir box centered 71-bp upstream of the promoter P4 controlling repABC expression. Sensing of plant phenolics by the VirA/VirG system thus also activates expression of the repABC operon, resulting in a 3- to 4-fold increase in Ti plasmid copy number (108). Second, repABC operon is subject to regulatory control by the opine-activated QS system (109). In response to opine activation of TraR
synthesis and sensory perception of AHL, TraR induces expression not only of the conjugation genes, but also of the repABC operon, resulting in an 8-fold increase in Ti plasmid copy number (109). Sensory perception of these two exogenous signals, plant phenolics and bacterial AHLs, thus results in enhanced repABC gene expression and an appreciable increase in Ti plasmid copy number at the site of infection. The resulting increase in Ti gene dosage correlates with enhanced Ti plasmid dissemination and virulence potential by invading agrobacterial cells (100).

**FITNESS COST TO MAINTENANCE OF TI PLASMIDS**

From the above discussion, it is evident that Ti plasmids code for a myriad of biological functions in response to sensory perception of a complex array of signals of plant and bacterial origin. A recent study assessed the fitness cost accompanying the metabolic load associated with carriage of Ti plasmids (110). As might be expected, the findings indicated that under conditions of nutrient abundance, the low-copy Ti plasmids exert only a modest cost to the agrobacterial host. Under such nutrient replete conditions, Ti plasmidless strains do not significantly outcompete Ti plasmid-carrying strains in batch culture, suggesting that carriage of the Ti plasmid does not impose a large metabolic burden. The complex regulatory network described above operates to minimize expression of the Ti plasmid genes under such conditions. Indeed, the only Ti plasmid locus required under these conditions is the repABC cassette, which is tightly controlled for maintenance of the Ti plasmid at single copy.

Surprisingly, however, there is a significant fitness cost associated with the Ti plasmid under conditions of nutrient limitation, even without induction of the virulence genes (110). Under these conditions, the expression profile of Ti plasmid genes is not expected to vary appreciably from growth in a rich environment, establishing that even the maintenance of this large plasmid at low copy and its vertical transmission during cell division imposes an appreciable metabolic burden. Nutrient limitation is likely a common stress for agrobacteria in nature, but despite the carriage cost of Ti plasmids, Ti-plasmidless strains are uncommon. In part, this can be attributed to the coupled expression of replication and partitioning genes, the latter ensuring vertical transmission of the plasmid during cell division. Additionally, a Ti plasmid toxin-antitoxin system also was recently described. Cells that lose the plasmid during cell division are killed owing to antitoxin instability and toxin stability (111).

Under conditions of nutrient limitation but favorable for vir gene expression, a cascade of events culminates in elaboration of the VirB/VirD4 T4SS, processing and transfer of T-DNA and effector proteins, expression of opine catabolism genes, upregulation of the repABC cassette, synthesis of the Ti plasmid Trb/Tra T4SS, and processing and translocation of the Ti plasmid. This burst of biological activity has the effect of imposing a major metabolic burden, as demonstrated through competition assays between plasmid-carrying and plasmid-free strains under vir-inducing conditions (110). Tight regulation of Ti plasmid genes under these conditions minimizes this metabolic load, yet the fitness cost accompanying vir gene expression begs the question of whether strains of agrobacteria called “cheaters” arise in nature. Such strains, for example, would carry variant forms of the Ti plasmid lacking the virulence genes while retaining the plasmid maintenance functions and, importantly, opine catabolism genes. Indeed, recent studies have shown that opine-catabolizing, avirulent strains of agrobacterial “cheaters” are widespread in nature, as are opine-catabolizing nonagrobacterial soil microbes (110, 112). Thus, virulent strains of agrobacteria face competition at the infection site with a spectrum of microbial “cheaters” or opportunists. Perhaps horizontal transmission of the Ti plasmid, activated by the cascade of plant, opine, and QS signals at the infection site, has evolved not just for dissemination of the Ti plasmid among coresident agrobacteria, but also as a means of forcing cooperation by agrobacterial cheaters (110). Clearly, there is a strong selective pressure in the framework of an intricate signaling network that serves to couple pathogenesis with Ti plasmid maintenance and dissemination at the infection site.

**BIOTECNOLOGICAL APPLICATIONS**

One of the most interesting features of Ti plasmids is their capacity to mediate conjugative transfer of T-DNA across kingdom boundaries. This property is of central importance for *A. tumefaciens* pathogenesis, but the discovery of transkingdom sex also spawned the multi-billion dollar industry of plant genetic engineering. It is now recognized that *A. tumefaciens* is capable of delivering any DNA flanked by T-DNA borders to an extremely broad range of plant hosts. Susceptible plant species include a wide range of gymnosperms and dicotyledonous species of agricultural importance. Early problems encountered in transformation of monocots such as rice, corn, and wheat were overcome with the use of actively
dividing cells such as immature embryos, preinduction of *A. tumefaciens* with phenolic inducers prior to infection, and screens for optimal plant genotype, type and age of plant tissue, bacterial strains, and T-DNA border-containing vectors. For rice and corn and many other species initially found to be recalcitrant to *A. tumefaciens* infection, most of these parameters have been optimized such that “agrotransformation” is now a routine technique (113).

For plant biologists, *A. tumefaciens*-mediated T-DNA transfer also has broader, more basic applications. For example, it is now possible to isolate novel plant genes by T-DNA tagging. Several variations to this methodology exist depending on the desired goals. For example, because insertions are generally randomly distributed throughout the plant genome, T-DNA is widely used as a mutagen for isolating plant genes with novel phenotypes. If the mutagenic T-DNA carries a bacterial origin of replication, the mutated gene of interest can easily be recovered in bacteria by suitable molecular techniques. Furthermore, if the T-DNA is engineered to carry a selectable or scorable gene near one of its ends, insertion downstream of a plant promoter will permit characterization of promoter activity. Conversely, if the T-DNA is engineered to carry an outward reading promoter, insertion can result in a modulation of gene expression with potentially interesting phenotypic consequences. Although random T-DNA insertion is a boon to investigators interested in characterizing plant genes, it is an undesired event for plant genetic engineering. In addition to the potential result that T-DNA will insert into an essential gene, insertion often is accompanied by rearrangements of flanking sequences, thus further increasing the chances that the insertion will have undesired consequences. Ideally, T-DNA could be delivered to a restricted number of sites in the plant genome. Progress toward this goal has involved the use of the bacteriophage P1 Cre/lox system for site-specific integration in the plant genome (114). The Cre site-specific recombinase catalyzes strand exchange between two lox sites, which for P1 results in circularization of the P1 genome upon infection of bacterial cells. For directed T-DNA insertion, both the plant and the T-DNA are engineered to carry lox sequences and the plant is also engineered to express the Cre protein. Upon entry of T-DNA into the plant cell, Cre was shown to catalyze the site-specific integration of T-DNA at the plant lox site. The frequency of directed insertion events is low in comparison with random insertion events, but additional manipulation of this system should enhance its general applicability.

Also of importance, the host range of *A. tumefaciens* now extends beyond the plant kingdom to include budding and fission yeast, many species of filamentous fungi, and even human cells (115, 116). The transformation of filamentous fungi with *A. tumefaciens* was an exciting advancement. *A. tumefaciens* was shown to efficiently deliver DNA to fungal protoplasts and fungal conidia and hyphal tissue. This DNA transfer system is especially valuable for species that are recalcitrant to transformation by other methods. Its overall simplicity and high efficiency make this gene delivery system an extremely useful tool for the genetic manipulation and characterization of fungi, and essentially all the current methodologies developed for basic studies of plants can now be applied to fungi. The discovery that *A. tumefaciens* can transform HeLa cells in laboratory culture expands the host range of this bacterium even further (116), enticing one to ask whether T-DNA transfer might form the basis of a viable gene delivery system for humans.

Finally, although the primary substrate of interest for the infecting bacterium and plant biotechnology is the T-DNA, Ti plasmid-encoded T4SSs as well as many other bacterial T4SSs also translocate protein substrates (22, 37). These translocation systems thus might be adaptable for the targeted delivery of therapeutic proteins to eukaryotic cell types, including human cells, of interest. The potential for protein therapy becomes more tangible as we learn more about type IV effector translocation signals, substrate-T4SS docking mechanisms, and T4SS-eukaryotic cell interactions.

ACKNOWLEDGMENTS

We thank members of the Christie laboratory for helpful discussions.

This work was supported by National Institutes of Health grant GM48746 to P.J.C.

Conflict of interest: We declare that we have no conflicts.

REFERENCES


Gordon and Christie


The Agrobacterium Ti Plasmids


