The *tra* Region of the Nopaline-Type Ti Plasmid Is a Chimera with Elements Related to the Transfer Systems of RSF1010, RP4, and F

STEPHEN K. FARRAND,^{1,2*} INGYU HWANG,^{1†} AND DAVID M. COOK^{1‡}

Departments of Crop Sciences¹ and Microbiology,² University of Illinois at Urbana-Champaign, Urbana, Illinois 61801

Received 25 January 1996/Accepted 6 May 1996

The Ti plasmids of *Agrobacterium tumefaciens* encode two transfer systems. One mediates the translocation of the T-DNA from the bacterium to a plant cell, while the other is responsible for the conjugal transfer of the entire Ti plasmid from one bacterium to another. The determinants responsible for conjugal transfer map to two regions, *tra* and *trb*, of the nopaline-type Ti plasmid pTiC58. By using transposon mutagenesis with Tn3HoHo1, we localized the *tra* determinants to an 8.5-kb region that also contains the *oriT* region. Fusions to *lacZ* formed by transposon insertions indicated that this region is expressed as two divergently transcribed units. We determined the complete nucleotide sequence of an 8,755-bp region of the Ti plasmid encompassing the transposon insertions defining *tra*. The region contains six identifiable genes organized as two units divergently transcribable from a 258-bp intergenic region that contains the *oriT* site. One unit encodes *traA*, *traF*, and *traB*, while the second encodes *traC*, *traD*, and *traG*. Reporter insertions located downstream of both sets of genes did not affect conjugation but were expressed, suggesting that the two units encode additional genes that are not involved in transfer under the conditions tested. Proteins of the predicted sizes were expressible from *traA*, *traC*, *traD*, and *traG*. The products of several Ti plasmid *tra* genes are related to those of other conjugation systems. The 127-kDa protein expressed from *traA* contains domains related to MobA of RSF1010 and to the helicase domain of TraI of plasmid F. The translation product of *traF* is related to TraF of RP4, and that of *traG* is related to TraG of RP4 and to VirD4 of the Ti plasmid T-DNA transfer system. Genetic analysis indicated that at least *traG* and *traF* are essential for conjugal transfer, while sequence analysis predicts that *traA* also encodes an essential function. *traB*, while not essential, is required for maximum frequency of transfer. Patterns of sequence relatedness indicate that the *oriT* and the predicted cognate site-specific endonuclease encoded by *traA* share lineage with those of the transfer systems of RSF1010 and plasmid F, while genes of the Ti plasmid encoding other essential *tra* functions share common ancestry with genes of the RP4 conjugation system.

The nopaline- and octopine-type Ti plasmids of *Agrobacterium tumefaciens* encode two conjugal transfer systems. One, called *vir*, is responsible for transferring a portion of the Ti plasmid, the T-region, from the bacterium to a plant cell (for reviews, see references 8, 32, and 75). This system has been examined in considerable detail. Like other conjugation systems, it is composed of *cis* elements and *trans*-acting factors. The *cis* elements, two 25-bp imperfect direct repeats called borders, define the T-region and are the functional equivalents of *oriT* sites of conventional conjugation systems (74). Both borders are nicked on the bottom strand by a site-specific single-strand DNA endonuclease, with the right-border sequence defining the polarity of transfer of the T-strand to the plant cell (67). *trans*-Acting functions involved in the transfer process are encoded within the *vir* region, which also is located on the Ti plasmid (8, 32, 52, 75). This regulon, which is com-

posed of four operons encoding structural components of the transfer system, is responsible for two sets of functions: the enzymes and other factors required for nicking the DNA at the border repeats, and the surface-associated complex, the mating bridge, which facilitates and directs the translocation of the DNA transfer intermediate from the bacterium to the plant. These functions correspond to the DNA transfer and replication (Dtr) functions and the mating pair formation (Mpf) functions associated with all conjugal transfer systems.

Despite the fact that this system recognizes a eukaryotic cell as the recipient, analysis of the derived amino acid sequences of proteins encoded by *vir* clearly shows that the T-DNA transfer system is phylogenetically related to other conjugal transfer systems. The T-region border sequences and the cognate VirD2 protein responsible for nicking at the T-region borders are related to *oriT* and to TraI, the *oriT*-specific endonuclease of the broad-host-range plasmid RP4 (44). Moreover, the 11 proteins encoded by the *virB* operon, which are believed to form the Mpf complex through which the T-DNA intermediate is transferred from the bacterium to the plant cell, are related at the amino acid sequence level to a set of proteins that is essential for the conjugal transfer of the IncN plasmid pKM101 (50). Intriguingly, this relatedness is conserved at the level of gene order; the 11 *virB* genes are organized in a manner essentially identical to that of the corresponding genes of pKM101. Recently, the conjugal nature of this Ti plasmid

* Corresponding author. Mailing address: Department of Crop Sciences, University of Illinois at Urbana-Champaign, 240 Edward R. Madigan Laboratory, 1201 West Gregory Dr., Urbana, IL 61801. Phone: (217) 333-1524. Fax: (217) 244-7830. Electronic mail address: stephenf@uiuc.edu.

† Present address: Department of Plant Biology, University of California at Berkeley, Berkeley, CA 94720.

‡ Present address: Division of Bacterial Products, Center for Biologicals Research, Food and Drug Administration, Bethesda, MD 20892.

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics ^a	Source or reference
<i>A. tumefaciens</i>		
NT1(pTiC58Δ <i>accR</i>)	Derivative of C58 harboring a Tra ^c mutant of pTiC58	5
C58C1RS	Ti plasmid-cured C58; Rif ^r Str ^r	64
<i>E. coli</i>		
DH5α	<i>supE44 ΔlacU169(φ80lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i>	54
S17-1(pHoHo1, pSSHe)	<i>pro res⁻ mod⁺ mob⁺ Amp^r Cm^r Tp^r Str^r; Tn3HoHo1 delivery strain</i>	This study
BL21(DE3)	B strain; F ⁻ <i>ompT</i> r _B ⁻ m _B ⁻ <i>hsdS gal(λDE3 cts857 int-1 Sam7 nin5 lacUV5-T7 gene 1)</i>	61
2174(pPH1JI)	<i>met pro</i> Gm ^r Sp ^r ; used to force homogenotes	7
Plasmids		
pUC19	Cloning vector; Amp ^r	54
pBluescript SK(+)	Cloning vector; Amp ^r	Stratagene
pET-14b	T7 promoter expression vector; Amp ^r	Novagen
pRK415	IncP1α-based broad-host-range cloning vector; Tet ^r	35
pCP13/B	IncP1α-based broad-host-range cloning vector; Tet ^r	26
pTHB58	Cosmid clone of a <i>Bam</i> HI partial digest of pTiC58 carrying the <i>tra</i> region in pCP13/B; Tet ^r	26
pRKH4	<i>Hind</i> III fragment 4 of pTiC58 cloned in pRK415	This study
pMP1	<i>Hind</i> III fragment 3 of pTiC58 cloned in pCP13/B	This study
ptrAC1-DG	<i>traCDG</i> of pTiC58 cloned in pET-14b	This study
ptrA-F	<i>traAF</i> of pTiC58 cloned in pET-14b	This study

^a Tra^c, constitutive for conjugal transfer; mob⁺, mobilizes transfer of other plasmids because a derivative of RP4 is integrated in the host chromosome (59); Amp^r, resistance to ampicillin; Cm^r, resistance to chloramphenicol; Gm^r, resistance to gentamicin; Sp^r, resistance to spectinomycin; Str^r, resistance to streptomycin; Tet^r, resistance to tetracycline; Tp^r, resistance to trimethoprim.

transfer system was established at the functional level; *vir* can mediate the transfer of plasmid DNA from an *Agrobacterium* donor to bacterial or yeast recipients (6, 10).

Although the self-conjugal nature of Ti plasmids was shown almost 30 years ago (for a review, see reference 21), considerably less is known about the conjugation system of these plasmids. Early studies indicated that at least two regions of both the octopine- and the nopaline-type Ti plasmids contain genes that are essential for conjugal transfer (15, 31). Mutations defining these two regions do not map to known *vir* loci and have no effect on tumorigenesis. Coupled with the observations that mutations abolishing tumorigenesis do not affect conjugal transfer, these results indicate that the Ti plasmid *tra* system is genetically independent of the *vir* system. More recently, by using Tn5 mutagenesis, Beck von Bodman et al. (5) showed that insertions in three genetically distinct regions of the nopaline-type Ti plasmid pTiC58 abolished conjugal transfer. One of these regions, TraI, subsequently was shown to encode a regulatory gene, *traR*, that is essential for the initiation of transcription of the *tra* genes (49). This region does not encode any other genes required for conjugation. A second region, originally called TraIII, encodes, in addition to a second regulatory gene (23, 24), the Mpf system of the Ti plasmid conjugation apparatus (2, 38). TraIII, now called *trb* (2, 38), maps about 100 kb from the other two Tra regions.

The third region, originally called TraII (5), maps adjacent to TraI but is genetically distinct from it. This region contains the *cis*-acting *oriT* site (14). Interestingly, the sequence of the Ti plasmid *oriT* bears a strong resemblance to those of a family of mobilizable but non-self-conjugal plasmids exemplified by the IncQ element RSF1010 (14, 68). We predicted that TraII also must encode *trans*-acting functions since Tn5 insertions in this region are complementable by a cosmid clone overlapping the locus (5). In this report we present the DNA sequence of the TraII region, which we rename *tra*. We show that this region contains, in addition to *oriT*, six identifiable genes that are organized in two divergently oriented transcriptional units. Our genetic studies and predictions based on relatedness of

the translational products of the Ti plasmid *tra* genes and other proteins in the databases indicate that most of the *tra* genes encode functions essential to conjugation. Furthermore, we show that the products of several of these genes are related to proteins involved in the transfer systems of other plasmids, including RSF1010, RP4, and F.

MATERIALS AND METHODS

Bacterial strains, growth conditions, and plasmids. The strains of *A. tumefaciens* and *Escherichia coli* and plasmids used in this study are listed in Table 1. *A. tumefaciens* strains were grown at 28°C in L broth (LB) (54), in ATN (48) or AB minimal medium (11), or on nutrient agar plates (Difco Laboratories, Detroit, Mich.). *E. coli* strains were grown at 28 or 37°C in LB, in M9 minimal medium (54), or on L agar plates. Mannitol or glucose, at a final concentration of 0.2%, was used as the sole carbon source in the minimal medium. When required, antibiotics were added at the following concentrations: for *A. tumefaciens*, kanamycin at 50 μg/ml, tetracycline at 2 μg/ml, carbenicillin at 100 μg/ml, gentamicin at 30 μg/ml, rifampin at 100 μg/ml, and streptomycin at 200 μg/ml; and for *E. coli*, kanamycin at 50 μg/ml, tetracycline at 10 μg/ml, and ampicillin at 100 μg/ml. X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside; Sigma Chemical Co., St. Louis, Mo.) was included in media at 40 μg/ml to assess the production of β-galactosidase.

DNA manipulations and plasmid constructions. Plasmids up to 50 kb in size were isolated by the alkaline lysis method (54). Ti plasmids were isolated as described by Hayman and Farrand (27). Standard recombinant DNA techniques were used as described by Sambrook et al. (54). Digestions with restriction endonucleases were done according to the instructions of the manufacturers, and digestion products were separated by electrophoresis in agarose gels with Tris-borate-EDTA buffer (54).

Plasmid pRKH4 was constructed by subcloning *Hind*III fragment 4 of pTiC58 from pTHB58 (Table 1) into pRK415 (35). Plasmid pMP1 was constructed by subcloning *Hind*III fragment 3 from pTHB58 into pCP13/B (26). The *traC*, *traD*, and *traG* genes were cloned as a unit into the T7 expression vector pET-14b (Table 1) as follows. A 310-bp *Nde*I-*Bam*HI fragment, generated by PCR, which contains the entire *traC* open reading frame (ORF) was cloned into pET-14b, resulting in ptrAC1. The insert was sequenced to confirm that no errors created by PCR were present in the clone. This cloning produced a translational fusion of the 60-bp His-Tag leader sequence, provided by the vector, to the N-terminal end of TraC. The operon was reconstituted by cloning the remaining ORFs, *traD*, *traG*, and 322 bp of *orfX*, as a 2,564-bp *Kpn*I-*Eco*RI fragment into ptrAC1, utilizing the unique *Kpn*I site located at the 3' end of *traC* and a unique *Eco*RI site in the vector portion of the plasmid. This clone was named ptrAC1-DG (Table 1). The *traA* and *traF* genes were cloned into pET-14b as follows. A 160-bp fragment containing the 5' end of *traA* was modified by PCR such that the amplified fragment contained a change in the start codon of *traA*, from GTG to

ATG, and an *EcoRI* site at its 3' end corresponding to the unique *EcoRI* site within the *traA* ORF. Changing the start codon to ATG allowed us to clone this fragment as an *NdeI-EcoRI* fragment into pUC19 to produce pUC19NE160. The modified fragment was verified by DNA sequence analysis. The *traAF* region was reconstituted by inserting a 4,549-bp *EcoRI-BamHI* fragment containing the remaining sequences into pUC19NE160 to form pUC19AF. The entire *traAF* region was recloned from pUC19AF into pET-14b as an *NdeI-BamHI* fragment to create *ptrA*-AF (Table 1). This cloning produced a translational fusion of the *traA* ATG initiation codon to the nucleotides of the vector encoding the His-Tag leader sequence.

Plasmid DNA was introduced into *A. tumefaciens* and *E. coli* hosts by transformation as previously described (30, 54).

Tn3HoHo1 mutagenesis and homogenization. Recombinant clones were mutagenized with Tn3HoHo1 as previously described (60) with the following modification. Plasmids pHoHo1 and pSShe were introduced by transformation into *E. coli* S17-1. This strain harbors a derivative of RP4 integrated into the bacterial chromosome (59). When a target plasmid is introduced into S17-1 (pHoHo1, pSShe), transposon-induced mutant plasmids can be isolated by directly mating out to a suitable recipient. The sites and orientations of the insertions were determined by restriction endonuclease analysis. Insertion mutations of interest were homogenized into pTiC58 Δ *accR* (Table 1) as previously described (33, 53), with pPH1J1 as the eviction plasmid (7). The validities of all marker-exchanged Ti plasmids were confirmed by restriction endonuclease analysis.

DNA sequence analysis. Both strands of appropriate DNA fragments were sequenced by the dideoxy method (55) using the Sequenase kit (version 2; United States Biochemical Co., Cleveland, Ohio). Nucleotide sequences were assembled and analyzed, and ORFs were identified and translated by using the DNASTAR (DNASTAR, Inc., Madison, Wis.) and DNA Strider (40) programs. DNA and deduced amino acid sequences were compared with those in the databases by using the BLAST search protocol (3). The PILEUP, GAP, and BESTFIT subroutines of the GCG program (Genetics Computer Group, Madison, Wis.) were used to compare sequences and to identify conserved regions among several protein sequences.

The sites of insertions of several Tn3HoHo1 mutations were determined by DNA sequence analysis with the primer Tn3LEAD as described by Hwang et al. (33).

Conjugation assays. Conjugal transfer of pTiC58 Δ *accR* and its mutant derivatives to the recipient strain, *A. tumefaciens* C58C1RS (Table 1), was assayed by the spot plate mating technique as described by Beck von Bodman et al. (5). This technique measures only initial mating events. Transfer frequencies are expressed as numbers of transconjugants per input donor.

Protein expression. Proteins expressed from the inserts in *ptrA*-AF and *ptrA*C1-DG were visualized as described by Tabor and Richardson (62). Each of the plasmids was transformed into *E. coli* BL21(DE3) (Table 1). Three cultures of each strain were grown in M9 glucose medium to an optical density at 600 nm of 0.5 to 0.6. Isopropyl- β -D-thiogalactopyranoside (IPTG; GIBCO BRL, Gaithersburg, Md.) was added to a final concentration of 1 mM to two of the cultures, and all three were incubated with shaking for 30 min. Rifampin was added to one of the induced cultures, and all of them were incubated for an additional 45 min. L-[³⁵S]methionine (specific activity, >1,000 Ci/mmol; Amersham, Arlington Heights, Ill.) was added to each culture to a final concentration of 10 μ Ci/ml, followed by incubation for an additional 15 min. The cells were collected by centrifugation, washed once with cold 5 mM Tris-HCl-2 mM EDTA (pH 8.2), and resuspended in cold double-distilled water. An equal volume of 2 $\times$ sodium dodecyl sulfate (SDS) gel loading buffer (54) was added, the samples were boiled for 5 min, and insoluble cell debris was removed by centrifugation.

SDS-polyacrylamide gel electrophoresis. Radiolabeled proteins were separated by electrophoresis on denaturing polyacrylamide gels containing 0.1% SDS (36). Following electrophoresis, the gels were dried under vacuum and autoradiographed by exposure to Kodak XRP5 film.

Nucleotide sequence accession numbers. The corrected sequence of the *oriT* region of pTiC58 was deposited in the GenBank database under accession number M95646. The entire 8,755-bp nucleotide sequence of the *tra* region was deposited in the GenBank database under accession number U40389.

RESULTS

Mapping the *tra* locus on pTiC58. By using Tn5 insertion mutagenesis, Beck von Bodman et al. (5) defined a ≥ 2 -kb region on pTiC58 called TraII that is essential for conjugation. This region maps about 10 kb counterclockwise from *acc* (27) at 7 o'clock on the Ti plasmid. Each of these insertion mutations was complementable by pTHB58, a cosmid clone of pTiC58 to which we have localized the origin of conjugal transfer (*oriT*) (14). pTHB58 contains an insert of approximately 30 kb of Ti plasmid DNA, with the *oriT* region mapping about 10 kb from one end (Fig. 1). To more precisely map the TraII locus, we mutagenized pTHB58 with Tn3HoHo1 as described in

Materials and Methods. We concentrated on insertions mapping to the contiguous fragments *Bam*HI-10 and *Bam*HI-4, since this region contains *oriT* and also is the site in which the Tn5 insertions defining TraII are located. A total of 20 independent insertions of Tn3HoHo1 mapping to this region were identified (Fig. 1; Table 2). Insertions also were identified in pRKH4, which contains *Hind*III fragment 4, and identified in pMP1, which contains *Hind*III fragment 3. Four of the insertions, V16, V7, V1, and IV2, identified in pRKH4, fill gaps between insertions I46 and I16 and between I16 and III25 on pTHB58, while insertions 5, 17, and 6 in pMP1 map between insertions III33 and III1 identified in pTHB58 (Fig. 1; Table 2). Three additional insertions, 43, 57 and 14, in pMP1 map to the right of III1.

The 30 selected Tn3HoHo1 insertions were recombined into pTiC58 Δ *accR*, a conjugation-constitutive (*tra*^c) derivative of pTiC58 (5), and each of the resulting homogenotes was tested for conjugal transfer. Fifteen of the mutations, bounded by and including I16 and III33, abolish transfer (Fig. 1; Table 2). Four adjacent insertions, 5, 17, 6, and III1, which map to the far right side (Fig. 1), do not abolish conjugation but significantly lower the transfer frequencies (Table 2). Together these insertions demarcate a region of approximately 8.5 kb. Mutations to the right of this region, starting with insertion 43, and those to the left, starting with insertion V1, have no detectable effect on conjugation (Fig. 1, Table 2). We rename this region *tra*, which is consistent with the nomenclature adopted for F and for RP4 (22, 45).

Insertions of Tn3HoHo1 can lead to fusions between the disrupted gene and *lacZ* present on the transposon (60). We used these *lacZ* fusions, marker exchanged into pTiC58 Δ *accR*, as reporters to determine the direction of transcription of *tra*. Of the insertions mapping to the right of *oriT*, only those reporting transcription away from the *oriT* site yielded cells that produced β -galactosidase (Fig. 1). Similarly, of the insertions mapping to the left of *oriT*, only those oriented away from this *cis*-active site reported transcriptional activity. Thus, the *tra* region of pTiC58 appears to be composed of two units divergently transcribed from the *oriT* region (Fig. 1). Three insertions mapping to the left side, V7, III47, and III19, and two insertions, mapping to the right side, 43 and 14, are transcriptionally active but have no effect on conjugation. This suggests that there is at least one gene within each of the transcriptional groups that is not essential for conjugation.

Nucleotide sequence of the essential *tra* locus. We determined the nucleotide sequences of both strands of an 8,755-bp region of pTiC58 from the *Pst*I site just to the left of insertion V1 to the *Bam*HI site at the junction between fragments 35 and 24 (Fig. 2). The sequence encompasses the entire *tra* region as defined by our genetic analysis. The nucleotide sequence includes corrections in the sequence of the 657-bp *Kpn*I-*Eco*RI fragment to which we had previously mapped the Ti plasmid *oriT* (14). As we resequenced this region, we identified several errors in the original sequence, and the correct sequence of the *oriT* region has been deposited separately in GenBank under accession number M95646.

We identified seven complete ORFs in the *tra* region (Fig. 2 and 3). All seven are oriented in a fashion consistent with our analysis of the *lacZ* reporter fusions, and most contain a reasonable consensus ribosome binding site (RBS). Moreover, all of the insertions affecting transfer map to the region encompassing these ORFs. On the basis of our sequence and genetic analyses, we designate the ORFs to the left of *oriT* as *traC*, *traD*, and *traG* and those to the right of *oriT* as *traA*, *traF*, and *traB* (Fig. 3). A complete ORF, designated *orfX*, is located downstream of *traG*. However, insertions mapping to this ORF

FIG. 1. Mutational analysis of the *tra* region of pTiC58. The restriction map of the *tra* region of pTiC58 and the locations of the independent Tn3HoHo1 insertions, identified in the cosmid clone pTHB58, in pRKH4 (which contains *Hind*III fragment 4), and in pMP1 (which contains *Hind*III fragment 3) are shown. Each insertion was marker exchanged (mx) into pTiC58 Δ *accR*, and the conjugal transfer ability of these mutants was assessed as described in Materials and Methods. mx transfer symbols: +, constitutive conjugal transfer at wild-type frequencies; +/- and -/+, transfer frequency reduced but still detectable; -, conjugation frequency of $<10^{-8}$ per input donor. The horizontal crossbar capping each insertion site indicates the direction of transcription reported by the *lacZ* gene of the Tn3HoHo1 element as determined in strains harboring the marker-exchanged Ti plasmid mutants. mx β -galactosidase (β -gal) activity symbols: +, colonies are blue on medium containing X-Gal; -, colonies are white on medium containing X-Gal.

do not affect conjugal transfer. The sizes of these genes as well as those of their predicted translational products are given in Table 3. An incomplete ORF is located immediately downstream of *traB* (data not shown). However, insertion mutations in this region have no effect on conjugal transfer (Fig. 1; Table 2). Potential -10 and -35 sequences are located upstream of the predicted start codons of *traC* and *traA* within the 258-bp region between the two *tra* operons (Fig. 2). The core *oriT*-active sequence lies within this intergenic region (14). We also identified an 18-bp inverted repeat sequence, which we call the *tra* box, mapping between coordinates 3055 and 3072 within the intergenic region (Fig. 2). This sequence is almost identical to two inverted repeats located in the 5' untranslated region of the *traI* gene, which encodes a protein required for the synthesis of the *Agrobacterium* *N*-acyl homoserine lactone autoinducer (23, 34). In turn, these regions of dyad symmetry are related to that of an inverted repeat called the *lux* box, located in the promoter region of the *lux* operon of *Vibrio fischeri* (58). These sequences are believed to define the binding sites for members of the LuxR family of transcriptional activators.

We determined the precise locations of three of the Tn3HoHo1 insertions by sequence analysis as described in Materials and Methods (Fig. 2). Insertion I41 is located at coordinate 2432 in *traG* and produces a transcriptional fusion to the *lacZ* reporter gene. Insertion II24 maps to coordinate

4201 and produces a transcriptional fusion in *traA*. Insertion 749, located at coordinate 6179, produces a translational fusion of LacZ to the carboxyl end of TraA. No insertions were isolated within the 258-bp intergenic region. Restriction analysis indicated that insertions III6 and III33 are located in *traF* and that insertions 5, 17, 6, and III1 are located in *traB*. Similarly, insertions II28, III25, IV2, and I16 all are located in *traG*, while insertion I20 is located in *traC*.

Expression of the *tra* genes in *E. coli*. Several of the ORFs overlap substantially with adjacent ORFs (Fig. 2). In addition, the *traA* ORF is very large (Table 3). We considered it important to verify the sequence analysis by expressing the products of these ORFs. Attempts to visualize products of the *tra* genes expressed from the native intergenic promoter region in *E. coli* or *A. tumefaciens* were unsuccessful. Therefore, we engineered the two *tra* operons into *E. coli* expression vectors as described in Materials and Methods. In each case, the first *tra* gene of the operon was translationally fused to the His-Tag leader sequence of the vector. Plasmid ptraC1-DG, which carries *traCDG*, expressed insert-specific proteins with sizes of 9.5, 15, and 70 kDa (Fig. 4B). Expression of these proteins required induction with IPTG. These sizes correspond reasonably well with those of TraD, His-tagged TraC, and TraG. Similarly, extracts of an *E. coli* strain harboring ptra-AF contained a protein that migrated with an electrophoretic mobility corre-

TABLE 2. Transfer frequencies of mutant derivatives of pTiC58 Δ accR

Ti plasmid ^a	Gene and allele ^b	Relative transfer frequency ^c
pTiC58 Δ accR	Wild type	1.0
pDCKIII34	—	1.0
pDCKI22	—	1.0
pDCKIII19	—	2.0
pDCKIII47	—	1.6
pDCKI46	—	2.0
pDCKV16	—	0.9
pDCKV7	—	1.8
pDCKV1	<i>orfX::Tn3lac1</i>	2.0
pDCKI16	<i>traG::Tn3lac5</i>	<10 ⁻⁴
pDCKIV2	<i>traG::Tn3lac4</i>	<10 ⁻⁴
pDCKIII25	<i>traG::Tn3lac3</i>	<10 ⁻⁴
pDCKII28	<i>traG::Tn3lac2</i>	<10 ⁻⁴
pDCKI41	<i>traG::Tn3lac1</i>	<10 ⁻⁴
pDCKI20	<i>traC::Tn3lac1</i>	<10 ⁻⁴
pDCKIII45	<i>traA::Tn3lac1</i>	<10 ⁻⁴
pDCK794	<i>traA::Tn3lac2</i>	<10 ⁻⁴
pDCKII24	<i>traA::Tn3lac3</i>	<10 ⁻⁴
pDCKII15	<i>traA::Tn3lac4</i>	<10 ⁻⁴
pDCKI18	<i>traA::Tn3lac5</i>	<10 ⁻⁴
pDCK727	<i>traA::Tn3lac6</i>	<10 ⁻⁴
pTiK749	<i>traA::Tn3lac7</i>	<10 ⁻⁴
pDCKIII6	<i>traF::Tn3lac1</i>	<10 ⁻⁴
pDCKIII33	<i>traF::Tn3lac2</i>	<10 ⁻⁴
pDCK17	<i>traB::Tn3lac1</i>	0.010
pDCK5	<i>traB::Tn3lac2</i>	0.016
pDCK6	<i>traB::Tn3lac3</i>	0.02
pDCKII1	<i>traB::Tn3lac4</i>	0.1
pDCK43	—	1.1
DCK57	—	1.2
pDCK14	—	0.8

^a Derivatives of pTiC58 Δ accR produced by homogenotizing the Tn3HoHo1 insertions shown in Fig. 1 into the Ti plasmid.

^b —, an insertion is present in an unsequenced portion of the Ti plasmid.

^c Relative transfer frequencies were calculated by dividing the transfer frequency of the mutant Ti plasmid by that of pTiC58 Δ accR. The frequency of transfer of pTiC58 Δ accR ranged from 3.3×10^{-4} to 4.2×10^{-3} transconjugants per input donor depending upon the experiment. Values shown for each mutant represent a typical experiment, and they were calculated relative to that of the wild-type plasmid used as the control in that series of matings.

sponding to a size of approximately 130 kDa (Fig. 4A). This is similar to the size predicted for the His-tagged TraA protein. We were unable to visualize proteins whose sizes corresponded with those of TraF and TraB.

Comparisons of *tra* genes from pTiC58 with those of other transfer systems. The first 398 amino acids of the predicted sequence of TraA (TraA398) are similar to and align with the amino acid sequences of MobA of RSF1010 (16, 57) and MobL of pTF1 (19). TraA398 and MobA showed 27% identity and 47% similarity when conservative changes were included. TraA398 and MobL are 30% identical and 51% similar, while MobA and MobL aligned with 32% identity and 53% similarity (Fig. 5).

The deduced amino acid sequence of TraA between residues 388 and 880 (TraA388–880) shows similarity to the helix-case domain of TraI from plasmid F (9) and to the product of the *trwC* gene of the IncW plasmid R388 (39). Alignment of TraA388–880 and TraI974–1450 showed 25% identity and 48% similarity, while that of TraA388–880 and TrwC471–966 showed 30% identity and 51% similarity (Fig. 6). An alignment of TraI974–1450 and TrwC471–996 showed 31% identity and 51% similarity (data not shown). Five strongly conserved domains, designated motifs I, II, III, V, and VI, are present

within the regions in which TraA, TraI, and TrwC show high sequence similarity (Fig. 6). These motifs are present in members of a superfamily of proteins involved in DNA replication and recombination (28, 29). A predicted nucleoside triphosphate binding site (66) with the consensus sequence RxxxVxGx AGxGK(T/S) is present within the amino acid sequences defining motif I of TraA388–880, TraI974–1450, and TrwC471–966 (Fig. 6). Motif IV, which also is characteristic of this superfamily, was not identified in the sequence of TraA.

TraB, TraC, and TraD are not significantly related at the amino acid sequence level to any other protein sequences in the databases. However, the 658-amino-acid (658-aa) TraG protein of pTiC58 is similar to the 635-aa TraG protein of RP4 (76) and also to the 665-aa VirD4 protein of pTiC58 (52) (Fig. 7). Pairwise comparisons showed that TraG of pTiC58 and TraG of RP4 are 34% identical and 47% similar. TraG of the Ti plasmid and VirD4 are 23% identical and 49% similar, while TraG of RP4 and VirD4 are 29% identical and 53% similar.

Comparisons of the predicted amino acid sequences showed that TraF of pTiC58 is 44 and 36% identical to and 66 and 62% similar to the TraF proteins of RP4 and R751 (76), respectively (Fig. 8). TraF of pTiC58 also is 65% identical and 79% similar to the amino acid sequence translatable from an ORF (*orf2*) which is located upstream of the *virF* gene in the *vir* region of the octopine-type Ti plasmid, pTi15955 (41). This homology between TraF of pTiC58 and a protein that is translatable from an ORF located in the *vir* region of an octopine-type Ti plasmid prompted us to compare sequence similarities between ORFs located in these regions of the two types of Ti plasmid. The region upstream of *virF* on pTi15955 contains three ORFs, *orf1*, *orf2*, and *orf3* (Fig. 9) (41). The sequence of the C-terminal 141 residues of TraA is 85% identical and 94% similar to the sequence of the predicted 141-aa protein that is translatable from *orf1*. The predicted N-terminal 115 residues of the 421-aa TraB protein share 63% identity and 71% similarity with the deduced protein product of *orf3*. At the nucleotide sequence level, the region of *tra* from coordinate 6117 to coordinate 7337 (Fig. 2), containing the 3' 424 bp of *traA*, all of *traF*, and the first 345 bp of *traB*, is 63% identical to the corresponding region of *vir*, which includes the incomplete *orf1*, all of *orf2* and *orf3*, and a 230-bp region spanning from the end of *orf3* up to and including the ATG start codon of the *virF* gene (data not shown). However, certain regions have greater degrees of identity than do others. The sequence of the last 424 nucleotides of *traA* is 80% identical to the DNA sequence of the incomplete *orf1* region of pTi15955, *traF* and *orf2* are 68% identical, and the 5' 347 bp of *traB* and the entire sequence of *orf3* are 70% identical. Nucleotide sequence identity between the remaining 919 bp of *traB* and the *vir* region fell to about 33% (data not shown).

DISCUSSION

The conjugal transfer system of pTiC58 is composed of at least three sets of genes (5). By a combination of genetics, DNA sequence analysis, and protein expression, we have fully defined one of these regions, which we call *tra*. This region is composed of three elements, two divergently transcribed gene sets that flank and define a 258-bp intergenic region. This intergenic region contains the *cis*-acting *oriT* site in addition to the promoters that drive expression of the two *tra* gene clusters. Although not formally proven, it is likely that the two gene sets each comprise single operons. On the left side, *traC* is promoter proximal. The two distal genes, *traD* and *traG*, are preceded by properly spaced sequences that match the con-

FIG. 2. Complete nucleotide sequence of the *tra* region of pTiC58 and deduced amino acid sequences of predicted ORFs. While the sequences of both strands were determined, only that of the 5' to 3' top strand is shown. Relevant restriction sites within the sequence are underlined. The start site of each ORF is preceded by its name, and a short dashed arrow shows the direction of expression. The stop codon of each ORF is marked with an asterisk. Lines above the nucleotide sequence indicate the positions of potential RBSS. Inverted triangles above the sequence mark the sites of the Tn³HoHo1 insertions that were determined by sequence analysis. In each case, the direction of transcription of the transposon *lacZ* gene is indicated by an arrowhead (< or >). Shaded boxes above the nucleotide sequence mark the locations of sequences that are homologous with those of the *oriT* sites of RP4 and RSF1010 and with T-region border repeats as specified. The 64-bp *Bam*HI-*Apa*I fragment, which is the smallest fragment that retains *oriT* activity (14), is delimited by the long dashed arrows. The converging arrows at coordinates 3055 to 3072 define the 18-bp *tra* box inverted repeat sequence.

2901 GCCTCGCGTGTTCGGCTGTTTTCAGCTGGTCTGCAATCTGGCAATTTCTTCCCTGATCTTCGATGATGGTTTCTTCATTCGGCTCGCATCTCCCTGGA
A E R T E A Q K L Q D Q L R A I E E R I K S S P K K M <--- *traC* start

3001 AATCTCGCGCTGTCTTCCGCTGCAAGATTCTCTCAAAGCACTTCGGAAGGAATGTGCGATCTGCACGTTCGGCAAGCCGACACTTTGGAGGATGAT
RSF1010 *nic* site homology *Bgl*II

3101 CCCGCCCTCGACGAGCGGATCCAAAGGGCGCAATTATACGTGCTGACGCGACGCTTGCCTAGGGGGCCAAACAGGGGGCCACTGTGGCTCACCGC
BamHI *oriT* region *Apa*I

3201 TCCCGACGAACGAGCTTCAAACGGGAGCTTTTACCGCGTGGCCATCGCCCACTTCTCAGCCAGCATCGTCAGCCGGCGACGGCCGACGCTGGTGTCT
traA start ---> M A I A H F S A S I V S R G D G R S V V L

3301 GTCTCGGCTTACGACACTGCGCGAAGATGGAATACGAGCGGAGGCCCGACCATCGACTACACCCGCAAGCAAGGCTGGTGCATCAGGAATTCATA
S A A Y Q H C A K M E Y E R E A R T I D Y T R K Q G L V H Q E F I *Eco*RI

3401 CTCCCGCGGATGCCCGAATGGGTCCGCGCTTGATTGACACTGTCTGTGCGCGGGCTCTGAGGCTTCTGGAACAAGTCGAGGCTTTCGAAA
L P A D A P K W V R A L I A D C S V A G A S E A F W N K V E A F E K

3501 AACGCTCCGACGCGAGCTGGCCCGCGACTGACGATTGCACTGCGCGGGAGCTTACATCCGAGCAGAATCGCCCTGGTCCGTGATTTCGTGGAGAA
R S D A Q L A R D L T I A L P R E L T S E Q N I A L V R F V E K

3601 ACATATCTGGGAGGGCATGGTCCGCGACTGGTCTATACGACCAATCCCGCAATCCGCACATCCACCTGATGACCACTTGGCGGCGCTAACCGAG
H I L G K G M V A D W Y H D N P G N P H I H L M T T L R P L T E

3701 GACGGTTCGGGCGCAAGAGGTAGCGGTCTTGGTGGAGCGCGAGCTGGTTCGCACGAAATCCGGCAAACTCTTACGAACTTTGGGCGCGTTCGA
D G F G A K K V A V I G E D G Q L V R T K S G K I L Y E L W A G S T

3801 CGGACGATTTCAACGCTGCTGCGCGATGGATGGTTCGAACGCTGAACCATCACCTGACGCTTGGCGGCATCGATCTCAAGATCGGATGGCGCTCTACGA
D D F N V V R D G W F E R L N H L T L G G I D L K I D G R S Y E *Cla*I *Cla*I

3901 GAAGCAGGAATCGACCTCGAACCGACGATCCATCTCGGAGTTGGCGCCAAAGCAATCTCGCGAAAGCGGAACAACAGGGTGTTCGGCGGAGCTCGAA
K Q G I D L E P T I H L G V G A K A I S R K A E Q Q G V R P E L E *Bcl*II

4001 AGGATCGAATCGAACGAAGAACCGGCTCGGAAACACCCGCGCATCTCAAATAACGACCATCGTCTCGACCTGATCGGGAAGAGAGCGTCT
R I E L N E E R R S E N T R I L K N P A I V L D L I M R E K S V F

4101 TCGATGAACGCGACGTCGCGAAAGTGCTGCACCGCTATGTCGACGATCCGCGCTCTTCCAGCAGCTGATGTTGCGCATCATCTGAATCCGAGGTGCT
D E R D V A K V L H R Y V D D P A V F Q Q L M L R I I L N P E V L *Sal*I

4201 ACGGTTCGACGCGACGATCGAGTTCGCGACCGGAGAGAAAGTGGCGGCGCTATTGACACGGGCGATGATCCGGCTGGAAGCGAGATGGCGCGA
V *Pst*I
R L Q R D T I E F A T D E K V P A R Y S T R A M I R L E A T E

4301 CAGGCGATGTCGCTGCGACAAGGAACGATCGCGTTTCCAGCGCGTGTCTGCTCGACGACTTTTGGGCGCATGGGCGGTGTCGGAGGAACAGA
Q A M W L S D K E T H A V S T A V L L A A T F G R H G R L S E E Q K

4401 AAGCAGCATCGAGTGCATCGCGGTCGCGCACGATCGCGCGGTCTGCGGACGCGGGTGTGGCAAGACCATGATGAAGCGCGCGCGGAGGC
A A I E C I A G P A R I A A V V G R A G A G K T T M M K A A R E A

4501 GTGGAGCTGCGCGGATATCGTGTGCTGCGTGGCGCGCTCGCGGCAAGCGCTCGGAGGCTTGGATAAGGAAGCGCGATCGAGGCCGACGCTCTCC
W E L A G Y R V V G G A L A G K A S E G L D K E A G I E S T L S *Eco*RV

4601 TCCTGGAGTTCGCTGGAACCGCGTGTGACGTGCTCGACAATAAACCCTCTTCTGTCATGGACGAGCGCGCATGGTGGCTCGAAACAGATGGCGG
S W E L R W N R G R D V L D N K T V F V M D E A G M V A S K Q M A G

4701 GCTTTGTCGATGCCGTCGTCAGGCGAGGGGCAAGATCGTGTAGTCGCGGATCCGAGCAGCTTCAGCCGATCGAAGCGGGGCGCTTTCGCGGCCAT
F V D A V V R A G A K I V L V G D P E Q L Q P I E A G A F R A I

4801 TGTGATCGGATGTTTACGCTGAATCGAATCGATCTATCGCCAGCGGAGSACTGGAAGCGGCTGCTGATCGATCGCGCGCGCAATTCGAG
V D R I G Y A E L E T I Y R Q R E D W M R K A S L D L A R G N V E

4901 AAGCGTTGGCGCTCTACAATGCGAATGCGAGAATGTTGGCGAAGCTGTGAAGCGGAAGCGGTTGAGCGCTGATCGCGACTGGAACCGTGATTACG
K A L A L Y N A N A R I V G E R L K A E A V E R L I A D W N R D Y D

5001 ATCAGCAGAAGACGATGATCTCTCGCCATCTGCGCGCGACGTGCGCATGCTCAACGTCATGGCGCGGAGAAGCTGGTTCGAGCGCGGATCTGTCG
Q T K T T L I L A H L R R D V R M L N V M A R E K L V R G I V G

5101 CGAGGCGCATGCTTTCAGAACGCTGACGGCGAAGCGGTTTCCAGCAGCGATCAGATCGTCTTCTGAAGAATGAGACCTTACTTGGGTCAGAAAC
E G H V F R T A D G E R R F H A G D Q I V F L K N E T L L G V K N

5201 GGCATGATCGGCCATGTCGTTGAGGCTGTGCCAAACCGCATCGTTCGCGTGGTGGGATCGGGATCATCGCGCGCATGTTGTGGTGAACAACGCTTCT
G M I G H V V E A V P N R I V A V V G D R D H R R H V V E Q R F Y

5301 ACAGCAATCTCGATCAGGCTATCGGACGACCATCCACAAGTCTCAGGCGCCACTGTCGACAGGGTCAAGGTGCTGGCTCTCTCTGCTCGATCGGCA
S N L D H G Y A T T I H K S Q G A T V D R V K V L A S L S L D R H *Sal*I

5401 TCTGACCTATGTCGCGATGACCGCCATCGCGAGGATCTCCAGCTTATTACGGATGCGGTTCTTCTGCGCTTCAACGGTGGCTAGGCAAGGTTCTTTCC
L T Y V A M T R H R E D L Q L Y Y G C R S F A F N G G L A K V L S

5501 CGAAAGAACGCCAAGGAGACGACGCTGATTACGAGCGCGCAAACTCTACCGCGAGGCGCTGCGCTTTGCCGAAACCGTGGCTGCACATCATGCAGG
R K N A K E T T L D Y E R K L Y R E A L R F A E N R G L H I M Q V

5601 TGGCTCGAAACAATGCTGCGGACCGGCTGCACTGGACGCTGCGCAAAAACAAAGGTTTCCGATCTGCTTCATAGGCTTCTGTCGCTGCTGGTGGCGCT
A R T M L R D R L D W T L R Q K T K V S D L V H R L R A L G E R L

5701 CGGCTCGCAATCCCAAAACCAACGATGAAGGAGCGCGCCCATGGTCAGAGCATCAAGACATTCTCCGCTCTGTTGGCGATACGGTGGCTGGT
G L D Q S P K T Q T M K E A A P M V T G I K T F S G S V A D T V G

FIG. 2—Continued.

sensus sequence for an RBS. Moreover, the start codon of *traD* is within 4 bp of the stop codon of *traC*, and the *traG* ORF overlaps that of *traD* by four codons. Neither *traD* nor *traG* contains identifiable -10 or -35 sequences in its upstream

regions. Similarly, to the right of *oriT*, the start codon of *traF* overlaps the translational stop codon of the preceding, promoter-proximal *traA* gene. However, *traF* is not preceded by any sequences showing a good match to the consensus RBS,

5801 GACAAGCTCGGCGCCGATCCACGCTGAAGCAGCAATGGGAGGAGGTCTCGGCGCGCTTCGCTACGTCTTCGCCGATCCCGAAACGCCTTCGCCGCCG
D K L G A D P T L K Q Q W E E V S A R F R Y V F A D P E T A F R A V

5901 TCAACTTCGATACGGTTCGGCAGATAAGGAGGAGCGGAAAGCGGTCTTCGAGAGCTTGAGGCGGAGCGCGCATCGATCGGGCGCGTGAAGGCAAGAC
N F D T V L A D K E A A K A V L Q K L E A E P A S I G A L K G K T

6001 CGGTATTCTGGCCAGCAAGACGGAGCGCGAAGCGCGCGGTGCGCAGAGTCAATGTTCCGCCCTGAAGCGCGATCTGGAGCAATATCTCGGATGCGC
G I L A S K T E R E A R R V A E V N V P A L K R D L E Q Y L R M R

6101 GAGACCGCGACGTCGGGATAGTGACCGAAGCAAGCCCTGCGCGAGCGGTCTCGATCGACATCCCGCGCTCTCGCGCGCGCAGCGTGGTGGT
E T A T L R I V T E E Q A L R Q R V S I D I P A L S P A A R V V L E

6201 AGCGCGTGCAGCGGATCGACCGGAACGATCTGCGCGCGCCCTTGGATGCGCTCAGCAATCGCGAGACCAAGCTCGAGATCGACGGGTCAACCA
R V R D A I D R N D L P A A L G Y A L S N R E T K L E I D G P F N Q

6301 GGCTGTTACCGAACGCTTCGGGAGCGGACCTGCTGAGCAACGTGGCGCGAGCGCTCGGAAAGCTGTACGAAAACTCTCGATGATGCGCGCG
A V T E R F G E R T L L S N V A R E P S G K L Y E K L S D G M R P

6401 GAGCAGAAAGAGCAGTTGAAGCAAGCTGGCCGATCATGCGCAGCGCCAGCAGCTCGCGCGCCACGAGCGCAGCGTGCAGTGCCTGAAACAGGCGCGG
E Q K E Q L K Q A W P I M R T A Q Q L A A H E R T V Q S L K Q A E E

6501 AATTGCGCGAGCTTGGCGAGGACCGGTGCTGAAACATGAGCAACGCGCTGTATCCGGTTTCTCGGTGTCGAGCGCTCGTCTTGAAGCGCGC
L R Q T L R Q A P V L K Q *

6601 ACCGTCTATGGGCTCATCGTGGTATCGAGTGAACGTCACGCCGAGCGGACCCCTTGGGCTCTGGCGTGTGGAAGCTCGAGCGCGCGGTTCACTCGC
T V M G L I G G Y R V N V T P S P L G L W R V E K L R R P V Q S G

6701 GCGATCTGGTGTGTTCTGCGCCCGCACAAACGCTTCTTCGAGGAAGCGTGGCTACGCGGTATCTCGTTCGGGCTCTGCCAGGTGTTGTTGCGCGC
D L V F V C P P H N A F F E E A W L R G Y L R S G L C P G W F A P

6801 GCTCATAAATCCGTTCTGCGCGTTCGCCGTCAGCGCGTGAATCGCGATCGGATGTCATCGAGCGTCACTCGGTCCTTCCATGTCAGCGCG
L I K S V L A L P G Q R V E I A D R I V I D G H P V S A S T V S A

6901 ACGGACAGCGAGGCGAGGCGATCGCGCCCTTGGCGCGCGCTGCTGCCCGAGGATTCTATTCTCATTTCTTTCGAGTTCCTATGATTCAC
T D S E G R A I A P F A G G V V P P G F L F L H S S F A S S Y D S R

7001 GATACTTCGCGCCGATCCCGACAGCGCTTCTCGGCTGGCGCGCGGTGTTCAGGTTTACCGGTACCGCTGGCGTGGCGCTCTGATCCTGC
Y F G P I P D S G L L G L A R P V F T F D P *

7101 TTTCGATCGCGCTCGGAGCGGTGGGCGGAGGCGCCCTCTTCTCCGACGCCATGCTCTTCCGTTGCTATGGCGCAATCTCATCAGCACT
S I A V G A V G E A L F P T A M L F P L L W A L R L

7201 GGTTCGGGCGCGGTGTCTCGCGGTACTTCTCACCGCTCGCGCGCGTGGCGAGGTTGTGGCGAACTTCTACCGCGCGATTTCTGGCAGCGCTC
V A G A V S A G Y F L T A S R G L P Q G V A N F Y A A D F W H G L

7301 CTGCTATGGTGGCGGCTCGCGCGCTTGTGCGCGTGCATGCGCGCTTGGCGCGCGCTTGCAGGAGGCTTCGGGAGGCGCGCTCTGGGT
L L V A A S A G F V A H A F W P A R L Q K R L P L G G L W

7401 GGGGAAAGCGGTGCGCTACCTTGGCGCGCGGTGCTGATGGGCTGCGCGCTTGGCGATCACGGGTGGGCGCATCACTAACCGAGCGGGCATCT
G K P V R Y L A A A V L M G L P P F G I T G W A H P L T A A G I L

7501 GTTTCGCGCTTTGGATGTTGGGCTGGTGGCGACACAGCGCGCTCGCGATGATGACATCGCGATACTGGCGGCTGTGCCATTGCCCTGGGAGGC
F P G F G W W G L G A T T A G L A M M T S R Y W P A A A I A L G G

7601 TTCTGGTCTGGTCCGCGCTACGTGGACGCAACCTGTCTTACCGGATGGATGGAAGGCGTCCGACCTCGAACAGGCTCAACGCTTGGACGTGATGGT
F W F W S A A T W T Q P V L P D G W K G V D L E Q G Q T L G R D G S

7701 CGCTTGACCATCACGCTGATCGCAACGCTGCGCGCGCGGAGCGGAAACCGTGTATCGTCTTCCGAAAGCGCGCTCGGCTTATGGAC
L D H H R D L I A T V R A A A G A E T R V I V L P E S A L G L W T

7801 CCCGACTGTGCGACGGCTTTGGCAGGCGGCTTGGGAGGCGGACGCTACCGTGTATCGCGCGCGCGCTCATCGATCGGGCGGCTACGACACGCTC
P T V A R L W Q A G L R G A D V T V I A G A A V I D P G G Y D N V

7901 ATGGTGACGCTCTCGGAGGCTGAACACGATCTCTATCGGAGCGGATGCCGATCCCGGTATCGATGTGGCAGCATGGTGAATGGACCGCGGAG
M V T V S E G E T R I L Y R E R M P I P V S M W Q P W L Q T L G R D G G

8001 GCGGTGGCGCTCAAGCGATTCTTTCGCGCAACCGCGCGGTGATCTCGCTGGCAGGAGTGGCGCGCTGATCTGCTACGAGCAGCTCATCGTCTGGC
G G A Q A H F F A N P A V D L A G T R I A P L I C Y E Q L I V W P

8101 CATACTGCACTCGATGCTGTTTCCCGCGCGCATCGTGCACGCGCAACGGATGGTGGACCGAGGAAACGTCGATCGTCCGATCCAAACAGCAGG
I L H S M L F S P A A I V A T G N G W W T E G T S I V A G I Q A A T

8201 GTGATCGCTGGGCGAACTCTTCGGCGACCGCTGTCACGGCTTTCAACACATAAGGAGACTTTCGCGATGGTGGATTCCGCCCTCATCAAGGATGC
V I A W A K L F G R P V V T A F N T *

8301 AGCGATCCCGCTGAAACCGCGATCTGCGAACGTTCAATGAAAGCGGAGGCGCGGACCGCTCGCGGTCACTGTCCGCTCCGGCAACCGCATCG
8401 TGCTTGTGCGAGGCGGACCAACCGGACGAGCGCTCGCATGTATACGCCAGAACCTCGGTGCAATATCGTTCGGTTCGGCTTACGCGATATCCCGC
8501 CGGGCTCGGTATCGTCAAGCGCGGCGAGTTGAAGCGGACCTGATCGAGCTTGGGAAACATCAAAATGGGACCGGCGCTTTCGCCAAAGTCTATCGT
8601 ATCGTACCAAAATGGTACGGTAACCTACCGAAGCGAGTCTTCCGCGGTTTTTGAAGACGCGATCATCGATGGCAACGGGGTATTTTCGAGGGGA
8701 CAGCGGTGTTCGGGACGCGATCCGGGCGACGTGAAGTGCAGAGCGGATCC 8755

FIG. 2.—Continued.

suggesting that translation of this gene may be coupled to that of *traA*. *traB*, which overlaps *traF* by 25 codons, is preceded by a reasonable RBS sequence. As with the leftward operon, there are no identifiable -10 or -35 sequences upstream of *traF* or *traB*.

This gene organization differs from that of plasmid F, in which the key *tra* genes are colinear and oriented in a single transcriptional direction (22). However, the organization of *tra*

on pTiC58 closely resembles that of the *TraI* region of the IncP1 α plasmid RP4 (45, 76). In both of these plasmids, the *tra* genes are organized as two divergently transcribed units. Moreover, in both of them, the *oriT* is located in the intergenic region between the two groups of genes. The organization of the Ti plasmid *tra* genes also resembles that of the *mob* region of RSF1010. This IncQ plasmid is not self-conjugal but can be mobilized by other conjugal plasmids coresident in the donor

FIG. 3. Physicogenetic map of the *tra* region of pTiC58. The map presents the region between the *Pst*I site at base pair coordinate 1 and the *Bam*HI site at base pair coordinate 8755 corresponding to the nucleotide sequence presented in Fig. 2. Genes identified by mutations, sequence analysis, and protein expression are presented as arrows positioned under the restriction map. The numbers in parentheses give the coordinates of the start and stop codons of each of the genes. The location and coordinates of the *oriT* site, as defined in reference 14, are shown above the map. The positions of the Tn3HoHo1 insertions V1 and 5, which delimit the ends of the essential *tra* region, also are shown. Abbreviations: A, *Apa*I; B, *Bam*HI; E, *Eco*RI; Ev, *Eco*RV; H, *Hind*III; K, *Kpn*I; P, *Pst*I; S, *Sal*I; X, *Xho*I.

cell (73). RSF1010 contains its own *oriT* and a set of three genes, *mobA*, *mobB*, and *mobC*. These three genes, each of which is required for mobilization of this plasmid by a conjugal element, are organized as two sets, *mobAB* and *mobC*, which are divergently transcribed from a 199-bp intergenic region that contains the RSF1010 *oriT* (16, 17, 57).

From our genetic analysis, we conclude that *traG*, on the left, and *traF*, on the right, are absolutely required for conjugation. Insertions in either gene completely abolish conjugation (Fig. 1). *orfX*, which is downstream of *traG*, appears to be nonessential. Transposon insertions in the region defining this ORF have no detectable effect on conjugal transfer of the Ti plasmid to *Agrobacterium* recipients (Fig. 1; Table 2). However, analysis of *lacZ* fusions located in this ORF indicates that it is expressed as part of the left-hand *tra* operon. In a similar fashion, our analysis suggests that *traB*, which is located downstream of *traF*, is not essential for conjugation per se but is required for maximum frequency of transfer. Insertions 17, 5, and 6, mapping to the 5' end of *traB* (Fig. 1), reduce transfer frequencies by 2 to 3 orders of magnitude, while insertion III1, which maps to the far 3' end of the gene, reduces frequencies by only a factor of 10. We cannot conclude from our data whether *traC* or *traD* is essential for conjugal transfer since insertion mutations in either of these genes most likely have a polar effect on the expression of *traG*. Similarly, we cannot conclusively state that *traA* is an essential gene since insertions in this ORF most likely have a polar effect on *traF*. However, as discussed below,

analysis of protein sequences suggests that *traA* does encode functions that are required for conjugation.

While *traG*, on the left, and *traF*, on the right, define the boundaries of the essential core *tra* region, it is likely that there are additional genes associated with both operons. On the left side, the eight Tn3HoHo1 insertions bounded by and including III34 and V1 have no effect on conjugal transfer but express β -galactosidase activity if oriented in the right-to-left direction (Fig. 1). Similarly, the three insertions mapping to the right of *traB* in the right-hand operon have no effect on conjugal transfer but confer expression of β -galactosidase activity if oriented in the left-to-right direction (Fig. 1). This is reminiscent of the Tra1 regulon of RP4. The two divergently transcribed units making up Tra1 encode 12 polypeptides, only 5 of which are essential for conjugal transfer (25, 37). Others influence transfer frequencies but are not essential, while the remainder, as defined by mutational analysis, play no detectable role in conjugation (37, 45).

On the basis of amino acid sequence relatedness, we can predict functions for several of the Ti plasmid *tra* proteins. The sequence of the *traA* gene translates to a very large protein of about 124 kDa, a prediction confirmed by visualization of a protein of about this size in extracts of *E. coli* expressing *ptrA*-AF (Fig. 4). The first one-third of TraA is related to MobL of pTF1 and to MobA of RSF1010, both of which function as site-specific single-strand endonucleases that cleave at their respective plasmid *oriT* sites (18, 56). This re-

TABLE 3. Characteristics of the *tra* genes of pTiC58 and their encoded products

<i>tra</i> gene	Gene size (bp)	Product size		Representative homologs
		No. of amino acid residues	Molecular mass (Da)	
<i>traA</i>	3,306	1,101	123,482	aa 1–398, MobA of RSF1010 and MobL of pTF1; aa 388–880, TraI of F (aa 974–1450) and TrwC of R388 (aa 471–966)
<i>traF</i>	531	176	18,842	TraF of RP4
<i>traB</i>	1,266	421	45,166	None
<i>traC</i>	297	98	10,350	None
<i>traD</i>	216	71	7,837	None
<i>traG</i>	1,977	658	71,061	TraG of RP4, VirD4 of pTiC58 and pTi15955, TrwB of R388

FIG. 4. Proteins expressed from the *tra* genes of pTiC58 in *E. coli*. Total proteins, radiolabeled with [³⁵S]methionine, were prepared from *E. coli* BL21(DE) harboring the appropriate expression plasmids and were separated by SDS-polyacrylamide gel electrophoresis as described in Materials and Methods. (A) Proteins expressed from *ptrA*-AF and the vector pET-14b. (B) Proteins expressed from *ptrA*1-DG and the vector pET-14b. Lanes 1 and 4, total protein profiles of uninduced cells not treated with rifampin; lanes 2 and 5, total protein profiles of cells induced with IPTG but not treated with rifampin; lanes 3 and 6, total protein profiles of cells induced with IPTG and treated with rifampin. Labeled arrowheads mark the positions of proteins as follows: in panel A, the His-tagged TraA protein (A) and in panel B, the TraG (G), the His-tagged TraC (C), and the TraD (D) proteins. The mobilities of molecular mass standards (MW Std) and their sizes (in kilodaltons) are shown to the right of the autoradiograms.

gion of TraA contains two tyrosine residues, Y-25 and Y-42, which are conserved in MobA and MobL. Tyrosine residues in the amino termini of TraI of RP4 and VirD2 of the Ti plasmid T-DNA transfer system are critical for the endonucleolytic activity of these proteins, as well as for their ability to form

covalent linkages at the 5' ends of the DNA transfer intermediates (4, 47, 65). In addition, a central region of TraA is related to the helicase domain of TraI from plasmid F. This domain of TraA contains several of the motifs, including a nucleoside triphosphate binding site, that are conserved in helicases associated with DNA replication in general and with plasmid transfer in particular (28, 29). This suggests that TraA is a multifunctional protein that catalyzes the site-specific single-strand nicking reaction at the Ti plasmid *oriT*, as well as the helicase activity required for displacement replication associated with strand transfer during conjugation (22, 71, 72). On the basis of its size and its putative multifunctional structure, TraA most closely resembles TraI of plasmid F, which is known to possess nicking and helicase activities (1, 63), and TrwC of R388 (39).

TraF of pTiC58 is related to a gene product of the same name from RP4. This protein, although encoded by a gene in the Tra1 regulon of RP4, actually forms part of the membrane-associated mating bridge assembly (37, 70). TraF of RP4 contains an N-terminal signal sequence, and a recent study showed that this protein is localized to the periplasmic and membrane fractions of *E. coli* (69). Similarly, the amino terminus of the Ti plasmid TraF protein contains a potential secretion signal se-

FIG. 5. The amino terminus of TraA is related to MobA of RSF1010 and MobL of pTiC58. The predicted amino-terminal 398 amino acids of TraA were aligned with the amino acid sequences of MobA and MobL by using the PILEUP algorithm from the GCG package as described in Materials and Methods. Three regions, labeled A to C, in which the sequence relatedness is highest are overlined. Amino acids which are identical or conserved (on the basis of the groups L, I, V, and M; G, A, and S; H, K, and R; D and E; N and Q; Y and F; and S and T) among at least two of the proteins are shown as white letters on a black background.

FIG. 6. The middle region of TraA is related to the helicase domain of TraI from plasmid F and a domain of TrwC from plasmid R388. The predicted amino acid sequence from residue 388 to residue 880 of TraA was aligned with the amino acid sequences of TraI and TrwC as described in the legend to Fig. 5 and in Materials and Methods. Five domains, labeled Motifs I, II, III, V, and VI, that are conserved in a family of proteins with helicase activity are overlined. A type A nucleoside triphosphate (NTP)-binding domain is located within domain I as indicated. Identical and conserved residues are displayed as described in the legend to Fig. 5.

quence (data not shown). The role of TraF in conjugation is unknown, but the protein is required for sensitivity to male-specific bacteriophages conferred by the Mpf complex of RP4 (70). Given the amino acid sequence resemblance, we predict that the Ti plasmid TraF protein plays a role similar to that of the RP4 homolog in conjugal transfer of the Ti plasmid.

TraC and TraD are not related at the amino acid sequence level to any protein sequences in the data banks. However, the *tra* system of RP4 and the *mob* system of RSF1010 both encode small proteins that are associated with the nucleoprotein relaxosome at the *oriT* site (24, 57). These proteins are specific to their transfer systems; neither TraJ, TraK, nor TraH of RP4 is related to MobB or MobC of RSF1010. These proteins are believed to assist the nicking enzyme in recognizing and interacting with the cognate *oriT* (24, 43, 56). We speculate that

TraC and TraD may play similar roles in guiding TraA to the Ti plasmid *oriT* site.

TraG is related to a family of proteins that are common to conjugal transfer systems of many plasmids from gram-negative bacteria. They include TraG of RP4, TrwB of R388, and VirD4 of the Ti plasmid *vir* transfer system. TraG of RP4 and VirD4 of pTiA6 each contains an N-terminal secretion signal sequence (51, 76), and both proteins are located in the bacterial membrane system (42, 69). The other members of this family, including TraG of the Ti plasmid, contain similar amino-terminal signal sequences, predicting that they also localize to the membranes. Lanka and colleagues have proposed that TraG and its homologs function to interface the nucleoprotein

FIG. 7. TraG of pTiC58 is related to TraG of RP4 and to VirD4. The predicted amino acid sequence of TraG from pTiC58 was aligned with the sequences of TraG from RP4 and VirD4 from pTiC58 as described in the legend to Fig. 5 and in Materials and Methods. All three proteins contain sequences similar to type A (designated as Motif A) and type B (designated as Motif B) nucleoside triphosphate-binding domains. Identical and conserved residues are displayed as described in the legend to Fig. 5.

C58TraF	LSNRRA	VTRFLGAGL	VLGGATVMGL	ICGYAVVWLP	SEPLGLRWVF
RP4TraF	MSP.FQRITK	YVAIGGGAAI	LLAG...EAY	LACARVNTTP	IPVGLYWKS
R751TraF	MSILKRIIAF	GVVIAVGAAL	LLAA...GSY	AAGARVNTTP	IPVGLYWTS
15955orf2	VPRFR	ALLPLTGAAY	FVSATLAAAT	TEGYRLNLP	SFPLGLWRIF
C58TraF	KLREPVQSD	LVEVCPHNA	FFEPALRGY	LRSGLCFGL	APLTKSVLHL
RP4TraF	NS...PVEKGA	YVVFQPPQVG	VFSDAKERGY	IASGFCPCGY	YVVKRVLAH
R751TraF	SP...PVEKGA	YVVFQPPQVG	VFSDAKERGY	IASGFCPCGY	YVVKRVLAH
15955orf2	QLRQVATSD	LVEVCPHNA	VFSDAKERGY	LRSGLCFGL	APLTKSVLHL
C58TraF	EGGRVEIADR	LIVIDGHPVS	ASTYSATDSE	GRALAEFASG	...VPTGFI
RP4TraF	KGEVVAIDCA	GVRVNSCLLF	MSALIKADPS	GRPLERYQSD	SYTLGTASVI
R751TraF	KDAVAVADD	GVRVNSCLLF	MSALIKADPS	GRPLERYQSD	SYTLGTASVI
15955orf2	EGGRVEIADR	LIVIDGHPVS	ASTYSATDSE	GRALAEFASG	...VPTGFI
C58TraF	FLSSSFASSY	DSRYFGPVPD	SGLLSGLRPV	ETEDH	
RP4TraF	IMSVDSDTSF	DGRYFGPVINR	SCNITVIRPV	LFW...	
R751TraF	IMSVDSDTSF	DGRYFGPVINR	SCNITVIRPV	LFW...	
15955orf2	FLSSSFASSY	DSRYFGPVPD	SGLLSGLRPV	ETEDH	

FIG. 8. TraF of pTiC58 is related to TraF of RP4 and R751 and to a protein translatable from an ORF in the *vir* region of pTi15955. The predicted amino acid sequence of TraF from pTiC58 was aligned with the sequences of TraF from RP4 and R751 and the translation product of *orf2* from pTi15955 as described in the legend to Fig. 5 and in Materials and Methods. Identical and conserved residues are displayed as described in the legend to Fig. 5.

relaxosome of the DNA transfer intermediate with the mating bridge complex (37, 45). The protein is essential for conjugation and also for mobilization of non-self-conjugal plasmids such as RSF1010, which themselves do not code for TraG-like proteins. Considering that TraG is essential for conjugal transfer of pTiC58, we propose a similar role for this protein in the Ti plasmid transfer system.

The *tra* region of pTiC58 is composed of elements related to transfer systems of several plasmid types. The *oriT* site clearly derives from the same lineage as does that of RSF1010 (14). This is a widespread *oriT*; the core sequence has been identified in other plasmids of gram-negative origin as well as in two plasmids, pIP501 and pG01, of gram-positive bacteria (12, 14, 68). Interestingly, while the core *oriT* sequence is strongly conserved among these plasmids, the inverted repeat sequences that are characteristic of *oriT* domains are very different. Consistent with the homology at *oriT*, the TraA protein, which we propose nicks at the Ti plasmid *oriT*, contains a domain that is closely related to MobA of RSF1010 (Fig. 5). However, TraA is substantially larger than MobA, and a second domain of this Ti plasmid-encoded protein bears a resemblance to the helicase domain of TraI from plasmid F. Thus, the two domains of TraA may have derived from different transfer systems.

While the *oriT*-nickase system is related to that of RSF1010, the Ti plasmid *tra* region also contains genes clearly related to the *tra* system of RP4. TraF and TraG are closely related to their RP4 homologs (Fig. 7 and 8). Interestingly, *tra* of pTiC58 and TraI of RP4 both are organized as two units divergently transcribed from the *oriT*, and both encode the same mixed sets of functions. The TraI regulon of RP4 encodes the *cis*- and *trans*-acting Dtr functions (13, 37, 70). However, TraF of RP4 is a component of the Mpf complex and is not required for processing at *oriT*. TraG of RP4 also is not part of the Dtr system. The Ti plasmid *tra* region certainly encodes DNA metabolism functions. Like RP4, this region also encodes the

FIG. 9. The *traAFB* region of pTiC58 shares nucleotide sequence identity with a region upstream of *virF* on the octopine-type Ti plasmid pTi15955. The figure presents a schematic alignment, based on nucleotide sequence identities, of ORFs (presented as arrows) from the *traAFB* region of pTiC58 and the region preceding *virF* on pTi15955. The 2.8-kb region of pTiC58 contains the 3' 324 bp of *traA* and all of *traF* and *traB*. The region of pTi15955 that aligns with the *tra* region sequence from pTiC58 contains three ORFs of unknown function (41). The region of nucleotide sequence relatedness between the sequences of the two Ti plasmids is denoted by an open box delimited by vertical dashed lines. The fill pattern in each of the arrows corresponds to the region of nucleotide (and derived amino acid) sequence relatedness.

TraF and TraG homologs. Taken together, the *tra* region of pTiC58 bears organizational resemblance to TraI of RP4 but at the gene level contains elements related to those of RSF1010, RP4, and F.

Like RP4, the Dtr and Mpf genes of the conjugal transfer system of pTiC58 are segregated into two regions. In addition to *tra*, a second region, called *trb*, is essential for conjugation (5). Our recent analyses indicate that *trb* encodes the membrane-associated Mpf complex of the Ti plasmid conjugation system (2, 38). Moreover, proteins encoded by genes located within this region of the Ti plasmid are strongly related to those encoded by Tra2 of RP4, which also encodes the Mpf complex. Given that in both plasmids *tra* is separated from *trb* but that *tra* contains two genes, *traF* and *traG*, whose products localize to the mating bridge, it seems clear that the structural components of the Mpf systems of pTiC58 and RP4 derived from a common ancestor.

Electron micrographic heteroduplex studies showed that pTiC58 contains four regions that are strongly conserved in the octopine-type Ti plasmid pTi15955 (20). Two of these regions, B and C, correspond to *tra* and *trb*, and sequence analysis indicates that the *tra* and *trb* regions of pTi15955 and pTiC58 are virtually identical (2, 38). Moreover, pTi15955 will mobilize transfer of a recombinant plasmid containing the core *oriT* of pTiC58 (14). Given that the two plasmid types are incompatible and that heteroduplex region B extends into the plasmid replication regions, it is clear that extensive regions of the nopaline- and octopine-type Ti plasmids share a relatively recent common origin.

The *vir* system, which mediates transfer of a segment of the Ti plasmid to plant cells, also shows phylogenetic relatedness to other conjugal systems (44, 50, 68). While VirD4 is related to the TraG proteins of pTiC58 and RP4, the remaining Vir proteins that are believed to be involved in T-DNA processing are not strongly related to other Tra proteins of the Ti plasmid. VirD2, which functions to produce single-strand, site-specific nicks at the T-region borders (46), shows little sequence similarity to TraA but is more closely related to TraI, the *oriT*-specific endonuclease of RP4 (44). Unlike TraA, neither VirD2 nor TraI contains a domain predicted to confer helicase activity. Presumably this function is provided by a host-encoded protein in these two transfer systems (72). Significantly, the *oriT* of RP4 is related to the Ti plasmid T-region borders, and VirD2 will cleave the RP4 *oriT* sequence at the *nic* site (44, 46, 47). Thus, the Dtr functions of the Ti plasmid *vir* system apparently share a common ancestral origin with those of the RP4 *tra* system but not with those of the Ti plasmid *tra* system. The *vir* mating bridge of the Ti plasmid is composed of proteins that are most similar to those of the Mpf of the IncN plasmid pKM101 (50), while those of the Ti plasmid *tra* system resemble the Mpf components of RP4 (2, 38).

The nucleotide sequence homology between the *traAF* region of pTiC58 and the region just upstream of *virF* of the octopine-type Ti plasmids is intriguing. A complete copy of a *traF* homolog is present in this region of the octopine-type Ti plasmid, as is a portion of *traB* (Fig. 9). These two sets of ORFs are approximately 70% identical at the nucleotide sequence level. Moreover, the incomplete 3' end of the *orf1* sequence is 80% identical to the sequence at the 3' end of *traA* of pTiC58. No additional sequence is available for the *virF* region of the octopine-type Ti plasmid, so we do not know if this sequence homology extends further into the 5' region of *traA*. However, it is intriguing to note that the carboxy-terminal one-third of TraA is not significantly related to any protein sequence in the databases. At the other end, there is no significant nucleotide sequence homology between the last two-thirds of *traB* and the

region of the octopine-type Ti plasmid immediately upstream of *virF* (Fig. 9), nor is there any homology between *virF* and any portion of the *tra* region at the nucleotide or derived amino acid sequence level. *virF* is a host range determinant and is required for tumorigenicity in certain but not all plant species (41). It is not known if *orf1*, *orf2*, or *orf3* expresses protein or what the roles of such proteins might be. However, genetic analysis (41) and electron micrographic heteroduplex studies (20) indicate that these three ORFs, as well as a functional copy of *virF*, are not present in pTiC58. Most likely, these ORFs represent scars of past recombination events.

Our results are consistent with the notion that both conjugation systems of the Ti plasmid are chimeras. The *vir* system is composed of elements common to pKM101 and RP4, while the *tra* system contains elements similar to those of RP4, RSF1010, and plasmid F. We suggest that these two transfer systems arose independently. Moreover, on the basis of analysis of mutations in both *tra* and *vir* (5, 15, 31), we suggest that these two Ti plasmid conjugal transfer systems function independently. This strongly suggests that pTiC58 may well have evolved from the fusion of two plasmids, each encoding a different conjugal transfer system. One system retained its function of transfer to bacterial hosts, while the other became modified to transfer DNA to plant cells.

ACKNOWLEDGMENTS

We thank Susanne Beck von Bodman and Pei-Li Li for discussions concerning conjugation of pTiC58 and Pei-Li Li for expert assistance with the computer analyses. We also thank Stephen C. Winans for sharing results and for discussions concerning the transfer system of octopine-type Ti plasmids.

Portions of this work were supported by grants R01-CA44051 and R01-GM52465 from the National Institutes of Health and grants 89-37263-4754 and 93-37301-8943 from the USDA to S.K.F.

REFERENCES

1. Abdel-Monem, M., G. Taucher-Scholz, and M. Q. Klinkert. 1983. Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. *Proc. Natl. Acad. Sci. USA* **80**:4659-4663.
2. Alt-Mörbe, J., J. L. Stryker, C. Fuqua, P.-L. Li, S. K. Farrand, and S. C. Winans. 1996. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid *vir* genes. *J. Bacteriol.* **178**:4248-4257.
3. Altschul, S. F., W. Gish, W. Miller, E. W. Meyers, and D. J. Lipman. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403-410.
4. Balzer, D., W. Pansegrau, and E. Lanka. 1994. Essential motifs of relaxase (TraI) and TraG proteins involved in conjugative transfer of plasmid RP4. *J. Bacteriol.* **176**:4285-4295.
5. Beck von Bodman, S., J. E. McCutchan, and S. K. Farrand. 1989. Characterization of the conjugal transfer functions of *Agrobacterium tumefaciens* Ti plasmid pTiC58. *J. Bacteriol.* **171**:5281-5289.
6. Beijersbergen, A., A. den Dulk-Ras, R. A. Schilperoord, and P. J. J. Hooykaas. 1992. Conjugative transfer by the virulence system of *Agrobacterium tumefaciens*. *Science* **256**:1324-1327.
7. Beringer, J. E., J. L. Beynon, A. V. Buchanan-Wollaston, and A. W. B. Johnston. 1978. Transfer of the drug-resistance transposon Tn5 to *Rhizobium*. *Nature (London)* **276**:633-634.
8. Binns, A. N., and M. F. Thomashow. 1988. Cell biology of *Agrobacterium* infection and transformation of plants. *Annu. Rev. Microbiol.* **42**:575-606.
9. Bradshaw, H. D., Jr., B. A. Traxler, E. G. Minkley, Jr., E. W. Nester, and M. P. Gordon. 1990. Nucleotide sequence of the *traI* (helicase) gene from the sex factor F. *J. Bacteriol.* **172**:4127-4131.
10. Bundock, P., A. den Dulk-Ras, A. Beijersbergen, and P. J. J. Hooykaas. 1995. Trans-kingdom T-DNA transfer from *Agrobacterium tumefaciens* to *Saccharomyces cerevisiae*. *EMBO J.* **14**:3206-3214.
11. Chilton, M.-D., T. C. Currier, S. K. Farrand, A. J. Bendich, M. P. Gordon, and E. W. Nester. 1974. *Agrobacterium tumefaciens* DNA and P58 bacteriophage DNA not detected in crown gall tumors. *Proc. Natl. Acad. Sci. USA* **71**:3672-3676.
12. Climo, M. W., V. K. Sharma, and G. L. Archer. 1995. Localization of *oriT* of *Staphylococcus aureus* conjugative plasmid pGO1, abstr. H-122, p. 513. In Abstracts of the 95th General Meeting of the American Society for Microbiology 1995. American Society for Microbiology, Washington, D.C.

13. Cole, S. P., E. Lanka, and D. G. Guiney. 1993. Site-directed mutations in the relaxase operon of RP4. *J. Bacteriol.* **175**:4911–4916.
14. Cook, D. M., and S. K. Farrand. 1992. The *oriT* region of the *Agrobacterium tumefaciens* Ti plasmid pTiC58 shares DNA sequence identity with the transfer origins of RSF1010 and RK2/RP4 and with the T-region borders. *J. Bacteriol.* **174**:6238–6246.
15. de Greve, H., H. Decraemer, J. Seurinck, M. van Montagu, and J. Schell. 1981. The functional organization of the octopine *Agrobacterium tumefaciens* plasmid pTiB6S3. *Plasmid* **6**:235–248.
16. Derbyshire, K. M., G. Hatfull, and N. Willetts. 1987. Mobilization of the nonconjugative plasmid RSF1010: a genetic and DNA sequence analysis of the mobilization region. *Mol. Gen. Genet.* **206**:161–168.
17. Derbyshire, K. M., and N. S. Willetts. 1987. Mobilization of the non-conjugative plasmid RSF1010: a genetic analysis of its origin of transfer. *Mol. Gen. Genet.* **206**:154–160.
18. Drolet, M., and P. C. K. Lau. 1992. Mobilization protein-DNA binding and divergent transcription at the transfer origin of the *Thiobacillus ferrooxidans* pTF1 plasmid. *Mol. Microbiol.* **6**:1061–1071.
19. Drolet, M., P. Zanga, and P. C. K. Lau. 1990. The mobilization and origin of transfer regions of a *Thiobacillus ferrooxidans* plasmid: relatedness to plasmids RSF1010 and pSC101. *Mol. Microbiol.* **4**:1381–1391.
20. Engler, G., A. Depicker, R. Maenhaut, R. Villarroel, M. van Montagu, and J. Schell. 1981. Physical mapping of DNA base sequence homologies between an octopine and a nopaline Ti plasmid of *Agrobacterium tumefaciens*. *J. Mol. Biol.* **152**:183–208.
21. Farrand, S. K. 1993. Conjugal transfer of *Agrobacterium* plasmids, p. 255–291. In D. B. Clewell (ed.), *Bacterial conjugation*. Plenum Press, New York.
22. Frost, L. S., K. Ippen-Ihler, and R. A. Skurray. 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. *Microbiol. Rev.* **58**:162–210.
23. Fuqua, W. C., and S. C. Winans. 1994. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. *J. Bacteriol.* **176**:2796–2806.
24. Fürste, J. P., W. Pansegrau, G. Ziegler, M. Kröger, and E. Lanka. 1989. Conjugative transfer of promiscuous IncP plasmids: interaction of plasmid-encoded products with the transfer origin. *Proc. Natl. Acad. Sci. USA* **86**:1771–1775.
25. Guiney, D. G., C. Deiss, V. Simnad, L. Yee, W. Pansegrau, and E. Lanka. 1989. Mutagenesis of the TraI core region of RK2 by using Tn5: identification of the plasmid-specific transfer genes. *J. Bacteriol.* **171**:4100–4103.
26. Hayman, G. T., and S. K. Farrand. 1988. Characterization and mapping of the agrocinopine-agrocin 84 locus on the nopaline Ti plasmid pTiC58. *J. Bacteriol.* **170**:1759–1767.
27. Hayman, G. T., and S. K. Farrand. 1990. *Agrobacterium* plasmids encode structurally and functionally different loci for catabolism of agrocinopine-type opines. *Mol. Gen. Genet.* **223**:465–473.
28. Hodgman, T. C. 1988. A new superfamily of replicative proteins. *Nature (London)* **333**:22–23.
29. Hodgman, T. C. 1988. Erratum: a new superfamily of replicative proteins. *Nature (London)* **333**:578.
30. Holsters, M., D. de Waele, A. Depicker, E. Messens, M. van Montagu, and J. Schell. 1978. Transfection and transformation of *Agrobacterium tumefaciens*. *Mol. Gen. Genet.* **163**:181–187.
31. Holsters, M., B. Silva, F. van Vliet, C. Genetello, M. de Block, P. Dhaese, A. Depicker, D. Inzé, G. Engler, R. Villarroel, M. van Montagu, and J. Schell. 1980. The functional organization of the nopaline *A. tumefaciens* plasmid pTiC58. *Plasmid* **3**:212–230.
32. Hooykaas, P. J. J., and A. G. M. Beijersbergen. 1994. The virulence system of *Agrobacterium tumefaciens*. *Annu. Rev. Phytopathol.* **32**:157–179.
33. Hwang, I., D. M. Cook, and S. K. Farrand. 1995. A new regulatory element modulates homoserine lactone-mediated autoinduction of Ti plasmid conjugal transfer. *J. Bacteriol.* **177**:449–458.
34. Hwang, I., P.-L. Li, L. Zhang, K. R. Piper, D. M. Cook, M. E. Tate, and S. K. Farrand. 1994. TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid *N*-acylhomoserine lactone autoinducer. *Proc. Natl. Acad. Sci. USA* **91**:4639–4643.
35. Keen, N. T., S. Tamaki, D. Kobayashi, and D. Trollinger. 1988. Improved broad-host-range plasmids for DNA cloning in Gram-negative bacteria. *Gene* **70**:191–197.
36. Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**:680–685.
37. Lessl, M., D. Balzer, K. Weyrauch, and E. Lanka. 1993. The mating pair formation system of plasmid RP4 defined by RSF1010 mobilization and donor-specific phage propagation. *J. Bacteriol.* **175**:6415–6425.
38. Li, P.-L., and S. K. Farrand. 1996. Characterization of the *trb* region on the nopaline-type Ti plasmid pTiC58 of *Agrobacterium tumefaciens*, abstr. H-76, p. 496. In Abstracts of the 96th General Meeting of the American Society for Microbiology 1996. American Society for Microbiology, Washington, D.C.
39. Llosa, M., S. Bolland, and F. de la Cruz. 1994. Genetic organization of the conjugal DNA processing region of the IncW plasmid R388. *J. Mol. Biol.* **235**:448–464.
40. Mark, C. 1988. "DNA Strider": a "C" program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers. *Nucleic Acids Res.* **16**:1829–1836.
41. Melchers, L. S., M. J. Maroney, A. den Dulk-Ras, D. V. Thompson, H. A. J. van Vuuren, R. A. Schilperoort, and P. J. J. Hooykaas. 1990. Octopine and nopaline strains of *Agrobacterium tumefaciens* differ in virulence: molecular characterization of the *virF* locus. *Plant Mol. Biol.* **14**:249–259.
42. Okamoto, S., A. Toyoda-Yamamoto, K. Ito, I. Takebe, and Y. Machida. 1991. Localization and orientation of the VirD4 protein of *Agrobacterium tumefaciens* in the cell membrane. *Mol. Gen. Genet.* **228**:24–32.
43. Pansegrau, W., D. Balzer, V. Kruff, R. Lurz, and E. Lanka. 1990. *In vitro* assembly of relaxosomes at the transfer origin of plasmid RP4. *Proc. Natl. Acad. Sci. USA* **87**:6555–6559.
44. Pansegrau, W., and E. Lanka. 1991. Common sequence motifs in DNA relaxases and nick regions from a variety of DNA transfer systems. *Nucleic Acids Res.* **19**:3455.
45. Pansegrau, W., E. Lanka, P. T. Barth, D. H. Figurski, D. G. Guiney, D. Haas, D. R. Helinski, H. Schwab, V. A. Stanisich, and C. M. Thomas. 1994. Complete nucleotide sequence of Birmingham IncPα plasmids. Compilation and comparative analysis. *J. Mol. Biol.* **239**:623–663.
46. Pansegrau, W., F. Schoumacher, B. Hohn, and E. Lanka. 1993. Site-specific cleavage and joining of single-stranded DNA by VirD2 protein of *Agrobacterium tumefaciens* Ti plasmids: analogy to bacterial conjugation. *Proc. Natl. Acad. Sci. USA* **90**:11538–11542.
47. Pansegrau, W., W. Schröder, and E. Lanka. 1993. Relaxase (TraI) of IncPα plasmid RP4 catalyzes a site-specific cleaving-joining reaction of single-stranded DNA. *Proc. Natl. Acad. Sci. USA* **90**:2925–2929.
48. Petit, A., and J. Tempé. 1978. Isolation of *Agrobacterium* Ti plasmid regulatory mutants. *Mol. Gen. Genet.* **167**:147–155.
49. Piper, K. R., S. Beck von Bodman, and S. K. Farrand. 1993. Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. *Nature (London)* **362**:448–450.
50. Pohlman, R. F., H. D. Genetti, and S. C. Winans. 1994. Common ancestry between IncN conjugal transfer genes and macromolecular export systems of plant and animal pathogens. *Mol. Microbiol.* **14**:655–668.
51. Porter, S. G., M. F. Yanofsky, and E. W. Nester. 1987. Molecular characterization of the *virD* operon from *Agrobacterium tumefaciens*. *Nucleic Acids Res.* **15**:7503–7517.
52. Rogowsky, P. M., B. S. Powell, K. Shirasu, T.-S. Lin, P. Morel, E. M. Zyprian, T. R. Steck, and C. I. Kado. 1990. Molecular characterization of the *vir* regulon of *Agrobacterium tumefaciens*: complete nucleotide sequence and gene organization of the 28.63-kbp regulon cloned as a single unit. *Plasmid* **23**:85–106.
53. Ruvkun, G., and F. Ausubel. 1981. A general method for site-directed mutagenesis in prokaryotes. *Nature (London)* **289**:85–88.
54. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
55. Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463–5467.
56. Scherzinger, E., R. Lurz, S. Otto, and B. Dobrinski. 1992. *In vitro* cleavage of double- and single-stranded DNA by plasmid RSF1010-encoded mobilization proteins. *Nucleic Acids Res.* **20**:41–48.
57. Scholz, P., V. Haring, B. Wittmann-Liebold, K. Ashman, M. Bagdasarian, and E. Scherzinger. 1989. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. *Gene* **75**:271–288.
58. Shadel, G. S., and T. O. Baldwin. 1992. Positive autoregulation of the *Vibrio fischeri luxR* gene. *J. Biol. Chem.* **267**:7696–7702.
59. Simon, R., U. Priefer, and A. Pühler. 1983. Vector plasmids for *in vivo* and *in vitro* manipulations of gram-negative bacteria, p. 98–106. In A. Pühler (ed.), *Molecular genetics of the bacteria-plant interaction*. Springer-Verlag KG, Berlin.
60. Stachel, S. E., G. An, C. Flores, and E. W. Nester. 1985. A *Tn3lacZ* transposon for the random generation of β-galactosidase gene fusions: application to the analysis of gene expression in *Agrobacterium*. *EMBO J.* **4**:891–898.
61. Studier, F. W., A. H. Rosenberg, J. J. Dunn, and J. W. Dubendorff. 1987. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **154**:60–89.
62. Tabor, S., and C. C. Richardson. 1985. A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. USA* **82**:1074–1078.
63. Traxler, B. A., and E. G. Minkley, Jr. 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. *J. Mol. Biol.* **204**:205–209.
64. van Larebeke, N., C. Genetello, J. Schell, R. A. Schilperoort, A. K. Hermans, J. P. Hernalsteens, and M. van Montagu. 1975. Acquisition of tumor-inducing ability by non-oncogenic agrobacteria as a result of plasmid transfer. *Nature (London)* **255**:742–743.
65. Vogel, A. M., and A. Das. 1992. Mutational analysis of *Agrobacterium tumefaciens virD2*: tyrosine 29 is essential for endonuclease activity. *J. Bacteriol.* **174**:303–308.
66. Walker, J. E., M. Sarste, M. J. Runswick, and N. J. Gay. 1982. Distantly related sequences in the α- and β-subunits of ATP synthase, myosin, kinases

- and other ATP-requiring enzymes and a common nucleotide binding fold. *EMBO J.* **1**:945–951.
67. **Wang, K., L. Herrera-Estrella, M. van Montagu, and P. Zambryski.** 1984. Right 25 bp terminus sequence of the nopaline T-DNA is essential for and determines direction of DNA transfer from *Agrobacterium* to the plant genome. *Cell* **38**:455–462.
 68. **Waters, V. L., and D. G. Guiney.** 1993. Processes at the nick region link conjugation, T-DNA transfer and rolling circle replication. *Mol. Microbiol.* **9**:1123–1130.
 69. **Waters, V. L., E. Lanka, and D. G. Guiney.** 1995. RK2 plasmid transfer proteins TraF and TraG localize to the host cell inner membrane, abstr. H-123, p. 513. *In* Abstracts of the 95th General Meeting of the American Society for Microbiology 1995. American Society for Microbiology, Washington, D.C.
 70. **Waters, V. L., B. Stack, W. Pansegrau, E. Lanka, and D. G. Guiney.** 1992. Mutational analysis of essential IncP α plasmid transfer genes *traF* and *traG* and involvement of *traF* in phage sensitivity. *J. Bacteriol.* **174**:6666–6673.
 71. **Wilkins, B., and E. Lanka.** 1993. DNA processing and replication during plasmid transfer between Gram-negative bacteria, p. 105–136. *In* D. B. Clewell (ed.), *Bacterial conjugation*. Plenum Press, New York.
 72. **Wilkins, B. M.** 1995. Gene transfer by bacterial conjugation: diversity of systems and functional specializations, p. 59–88. *In* S. Baumberg, J. P. W. Young, S. R. Saunders, and E. M. H. Willington (ed.), *Population genetics in bacteria*. Cambridge University Press, Cambridge.
 73. **Willetts, N. S., and C. Crowther.** 1981. Mobilization of nonconjugative IncQ plasmid RSF1010. *Genet. Res.* **37**:311–316.
 74. **Yadav, N. S., J. Vanderleyden, D. R. Bennett, W. M. Barnes, and M.-D. Chilton.** 1982. Short direct repeats flank the T-DNA on a nopaline Ti plasmid. *Proc. Natl. Acad. Sci. USA* **79**:6322–6326.
 75. **Zambryski, P. C.** 1992. Chronicles from the *Agrobacterium*-plant cell DNA transfer story. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **43**:465–490.
 76. **Zieglin, G., W. Pansegrau, B. Strack, D. Balzer, M. Kröger, V. Druft, and E. Lanka.** 1991. Nucleotide sequence and organization of genes flanking the transfer origin of promiscuous plasmid RP4. *DNA Sequence Map* **1**:303–327.