

Sequence Determination and Characterization of the Replicator Region in the Tumor-Inducing Plasmid pTiB6S3

SATOSHI TABATA,¹ PAUL J. J. HOOYKAAS,² AND ATSUHIRO OKA^{3*}

Department of Biology, Faculty of Science, Nagoya University, Nagoya-shi, Aichi 464,¹ and Institute for Chemical Research, Kyoto University, Uji-shi, Kyoto 611,³ Japan, and Department of Plant Molecular Biology, Leiden University, 2333 AL Leiden, The Netherlands²

Received 2 August 1988/Accepted 6 December 1988

The replicator region of the 195-kilobase-pair (kb) tumor-inducing plasmid pTiB6S3 was previously identified by isolation of a 6.8-kb miniplasmid (B. P. Koekman, P. J. J. Hooykaas, and R. A. Schilperoort, *Plasmid* 7:119-132, 1982). This miniplasmid was joined to ColE1-based vectors and subjected to mutagenesis. The resulting mutant plasmids were examined for their ability to replicate autonomously in *Agrobacterium tumefaciens*. It was found that a 4.2-kb region was sufficient for displaying replication characteristics similar to those of the parental pTiB6S3. Nucleotide sequence analysis of this 4.2-kb region revealed the presence of three possible reading frames in the same direction (*repA*, *repB*, and *repC*). Proteins coded for by these frames were identified by *in vitro* synthesis in a coupled transcription-translation system. The replicating ability became attenuated by *repA* and *repB* mutations but was completely abolished by *repC* mutations. The size, arrangement, and mutational effects of the three *rep* genes were quite similar to those of the *rep* genes that were previously identified in the hairy root-inducing plasmid pRiA4b. However, defects caused by *rep* mutations in one plasmid were unable to be complemented by corresponding functions in the other plasmid.

Crown gall and the related disease hairy root are caused by *Agrobacterium tumefaciens* and *A. rhizogenes*, respectively. These bacteria possess large plasmids called tumor-inducing and root-inducing plasmids, respectively (designated pTi and pRi, respectively). Portions of pTi and pRi DNAs are transferred into plant cells and then integrated into their nuclear genomes. The transferred DNA (T-DNA) carries genes which control tumor morphology. The T-DNA also contains genes which direct the synthesis of unique amino acid derivatives called opines. Imperfect 25-base-pair (bp) direct repeats flanking the T-DNA are requisite in *cis* for transfer. A set of plasmid genes (*vir*) essential for virulence is located outside the T-DNA. Despite the resemblance in these characteristics, the similarity of the nucleotide sequences between these two plasmids is limited to specific areas. In addition, pTi and pRi belong to different incompatibility groups and can be maintained stably together in a single bacterium (for reviews, see references 9 and 16). By isolation of miniplasmids the replication and incompatibility regions (replicator regions) of the octopine-type plasmid pTiB6S3 and the agropine-type plasmid pRiA4b have been localized (10-12, 17). To understand the molecular basis of pTi and pRi replication and to help in the construction of vectors for use in the *Agrobacterium* delivery system, we have introduced various mutations within the replicator region of pTiB6S3, determined the gene organization, possible functions, and nucleotide sequence of the pTiB6S3 replicator, and compared these characteristics with those previously published for the pRiA4b replicator (17). Complementation analysis was also performed between pTi and pRi replication mutants.

MATERIALS AND METHODS

General methods. Methods for transformation with *Escherichia coli* and *A. tumefaciens*, preparation of plasmid DNA from *E. coli* and *A. tumefaciens* cells, restriction endonucle-

ase digestion, repair synthesis, ligation, linker-mediated ligation, gel electrophoresis in agarose and polyacrylamide, extraction of DNA fragments from gels, rapid clone analysis with *E. coli*, rapid detection of plasmid DNA from *Agrobacterium* cells, and determination of the segregation frequency of plasmid-free cells were described previously (17).

Culture media, antibiotics, and reagents. L broth, L agar, YEB, YEB agar, and antibiotic concentrations used were described previously (17). Restriction endonucleases were purchased from Takara Shuzo or New England BioLabs, Inc. DNA polymerase I Klenow fragment, T4 DNA ligase, and linker oligonucleotides were obtained from Takara Shuzo. A low-molecular-weight protein mixture was used for the molecular size markers (Pharmacia, Inc.).

Plasmids, bacteria, and phages. The plasmids used and their relevant characteristics are listed in Table 1. The *E. coli* strains used were DH1 (F⁻ *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1*) and JM109 [*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 Δ(lac-proAB)* (F' *traD36 proA⁺B⁺ lacI^s lacZΔM15*)] (21). Pathogenic *A. tumefaciens* AT1005, AT1006, AT1013 and AT1018 carried pTiC58, pTiB6S3, pTiA6, and pTiBo542, respectively (7, 11, 13). The nonpathogenic *A. tumefaciens* strains used were pTi-free GV3101 (7) and GV3101 carrying various plasmids (Table 1). Single-stranded DNA phages M13mp18 and M13mp19 (21) were the vectors used for sequencing.

Determination of pTi incompatibility properties. Since two pTi plasmids incompatible with each other could not be introduced at all in a single *Agrobacterium* cell, the incompatibility characteristics of two plasmids were examined by transformation for one plasmid with the recipient bacteria carrying the other plasmid, and transformants were selected without an antibiotic for the resident plasmid. When resident plasmids were *ein⁺* (establishment inhibition) (12), transformation occurred either at a normal frequency or at a frequency no greater than 10⁻² times a normal frequency. Transformants obtained in the former case generally had retained the resident plasmid, and the two plasmids were

* Corresponding author.

TABLE 1. Plasmids used and their relevant characteristics

Plasmid	Selective marker	Relevant characteristics	Source or reference
pAO254 ^a	Km ^r	ColE1 <i>ori lacZ_α^b kan</i> of Tn5	This study
pAO260 ^a	Km ^r	ColE1 <i>ori lacZ_α^b kan</i> of Tn5	This study
pUC18	Cb ^r	ColE1 <i>ori lacZ_α^b bla</i>	21
pHSG298	Km ^r	ColE1 <i>ori lacZ_α^b kan</i> of Tn903	Takara Shuzo
pKK223-3	Cb ^r	ColE1 <i>ori bla tac</i> promoter	Pharmacia
pAL2821	Cb ^r	pTiB6S3 <i>repA⁺B⁺C⁺ bla</i>	11, 12
pAO224	Cb ^r	pUC18 carrying pRiA4b <i>repA⁺B⁺C⁺</i>	17
pAO269	Km ^r	pAO254 carrying pRiA4b <i>repA⁺B⁺C⁺</i>	17
pAO273	Km ^r	pAO254 carrying pRiA4b <i>repA⁺B⁺C⁺</i>	17
pAO274	Km ^r	pAO254 carrying pRiA4b <i>repA⁺BC⁺</i>	17
pAO385	Km ^r	pAO254 carrying pRiA4b <i>repA⁺BC</i>	17
pAO395	Km ^r	pAO254 carrying pRiA4b <i>repAB⁺C⁺</i>	17
pAO401	Km ^r	pAO254 carrying pRiA4b <i>repAB⁺C</i>	17
pAO403	Km ^r	pAO254 carrying pRiA4b <i>repABC⁺</i>	17

^a The transcriptional orientations of *lacZ_α* and *kan* are opposite in pAO254 but are the same in pAO260.

^b *lacZ_α* contains the multiple cloning restriction sites of pUC18.

thus judged to belong to different incompatibility groups. In the latter case, transformants had lost the resident plasmid or contained a cointegrate between the two plasmids, and the two plasmids were thus judged to belong to a single incompatibility group. Plasmids carried by transformants were determined by the rapid plasmid detection procedure with several colonies because all *ein⁺* pTi plasmids used carried no selective marker. When *ein* mutants (Km^r or Cb^r) were used as resident plasmids, transformation always occurred at a normal frequency. The unselected markers of transformants for resident plasmids were then tested. If a resident plasmid was retained in or was expelled from transformants, the two plasmids were judged as compatible or incompatible, respectively.

Complementation of *rep* genes between pTi and pRi. A pTi *rep* mutant (Km^r) was introduced by transformation into GV3101 and GV3101(pAO224), and transformants were selected on kanamycin-containing agar plates. When transformation occurred (or did not occur) with the two recipient strains almost identically with regard to the size and number of transformant colonies, the pTi *rep* mutation was judged to be noncomplemented by the pRi *rep* genes carried by pAO224. When transformation occurred only with GV3101(pAO224) or when transformant colonies with GV3101(pAO224) were larger than those with GV3101, the pTi *rep* mutation was judged to be complemented by the pRi *rep* genes. Complementation by reverse combinations was also carried out with pRi *rep* mutants (Km^r) instead of pTi *rep* mutants and with GV3101(pAO359) (see Results) instead of GV3101(pAO224).

Nucleotide sequence determination. Replicative-form DNA of M13mp18 or M13mp19 was joined to DNA fragments which were to be sequenced and into which deletions of various sizes were to be introduced by the unidirectional progressive deletion procedure (21). Each deletion clone was sequenced by the chain termination method (18) with an M13 7-DEAZA sequencing kit (Takara Shuzo). In some cases,

heptadeca-deoxyribonucleotides synthesized by a System 1 DNA synthesizer (Beckman Instruments, Inc.) were used as nonuniversal primers.

Protein analysis. Polypeptides were labeled with L-[³⁵S]methionine (Amersham Corp.) in an *E. coli*-coupled transcription-translation system in vitro (22) with a prokaryotic DNA-directed translation kit (Amersham) and then separated in 12.5% sodium dodecyl sulfate-polyacrylamide gels as described by Laemmli (14). Protein bands were visualized by fluorography with Enlightning (New England Nuclear Corp.).

RESULTS

Hybrid plasmids and their deletion and insertion derivatives replicating in both *A. tumefaciens* and *E. coli*. pAL2821 (Cb^r) (size, about 6.8 kilobase pairs [kb]) was previously isolated by in vitro manipulation from a derivative of pTiB6S3 with a copy of Tn1 integrated near the replicator region (11, 12). To facilitate plasmid DNA preparation, we joined ColE1-based vector pAO254 (Km^r) to pAL2821 after linearization at the unique *Sall* site located close to the *bla* gene derived from the unique *Sall* site located close to the *bla* gene derived from Tn1. The resulting recombinants were named pTi-1(I) and pTi-1(II) (Fig. 1); I and II indicate relative polarities between the pAL2821 and pAO254 moieties. Polarities I and II were defined as the direction in which the transcription of the *rep* genes (see below) of pAL2821 was, respectively, opposite to and the same as that of *lacZ_α* of pAO254. Both pTi-1(I) and pTi-1(II) but not the pAO254 vector itself transformed *Agrobacterium* GV3101 cells to Km^r as well as pAL2821 did. Since the number and size of transformant colonies were indistinguishable among pTi-1(I), pTi-1(II), and pAL2821 (about 2,000 transformants per 2 × 10⁹ recipients per μg of DNA under the conditions used), it was concluded that these three plasmids replicate almost identically in *Agrobacterium* cells. In addition, the results indicated that, regardless of its polarity, the vector moiety does not affect the ability of pAL2821 DNA to allow autonomous replication. This procedure provided a simple system for assaying the ability of pTi-1 and its derivatives (described below) to replicate autonomously in *Agrobacterium* cells (Rep function).

pTi-1 DNA was cut with various restriction endonucleases, and a cleavage map of the pAL2821 DNA was constructed. Representative cleavage sites are shown at the top of Fig. 1. By use of appropriate cleavage sites, various deletion derivatives from either pTi-1(I) or pTi-1(II) were constructed and named pTi-2 to pTi-8. In addition, a 4-bp insertion was introduced both at the *Mlu*I site (pTi-9) and at the *Bam*HI site (pTi-10) by filling in of the respective restriction cohesive ends followed by religation. Similarly, an *Xba*I linker of 8 bp was inserted at the *Stu*I site (pTi-11). The structures of the pTi moiety in all of these derivatives are illustrated in Fig. 1.

Replicating ability of pTi-1 to pTi-11. The possible Rep function of each plasmid described in Fig. 1 was examined by using the Km^r marker for selection. Replicating plasmids (Rep⁺) produced Km^r transformant colonies of a normal size, whereas those with defects in the Rep function (Rep⁻) produced no transformant colonies at all or transformant colonies at reduced growth rates that varied with each derivative. The number of transformants growing at reduced rates were almost the same as those with pAL2821. When these transformant colonies were observed at certain intervals, their sizes were markedly different in each mutant. On the basis of the supposition that the reduced rate of colony formation reflects the degree of defectiveness in the Rep

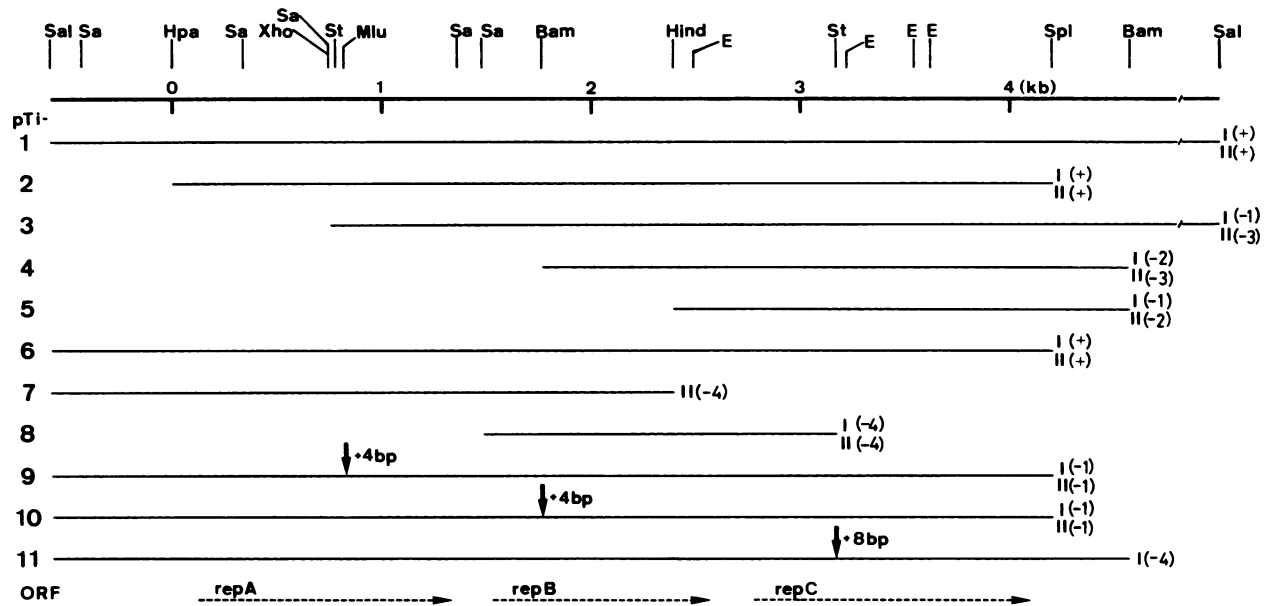


FIG. 1. Map of the pTi moiety in various hybrid plasmids. pTi-1 to pTi-11 carry the Km^r marker. The cleavage map is shown at the top; pAL2821 DNA was linearized at the unique *Sall* site (Bam, *Bam*HI; E, *Eco*RI; Hind, *Hind*III; Hpa, *Hpa*I; Mlu, *Mlu*I; Sa, *Sac*I; Sal, *Sall*; Spl, *Spl*I; St, *Stu*I; Xho, *Xho*I). The scale is graduated in kilobases from the left (the *Hpa*I site) to the right, corresponding to the nucleotide positions of the sequence in Fig. 3. The 1.6-kb *Bam*HI-*Sall* region at the right end is mostly occupied by the *Tn1 bla* gene of non-pTi DNA (11) and is drawn to a reduced scale. A DNA region carried by each pTi mutant is indicated by a horizontal line, and an insertion of an oligonucleotide is indicated by a vertical arrow. The Rep function is shown in parentheses at the right end of each map. + and -1 to -4 indicate Rep^+ and Rep^{-1} to Rep^{-4} (classified according to the degree of defects), respectively (see the text). I and II represent the polarities of the pTi moiety against the vector. The three broken arrows at the bottom represent open reading frames (ORF) deduced from the nucleotide sequence data.

function, the Rep^- phenotypes were classified for convenience into four groups from absolute Rep^- (Rep^{-4}) to modest Rep^- (Rep^{-1}) according to the rate of colony formation. The results of such assays are summarized in parentheses on the right side of Fig. 1. They indicate that deletions to the left of the *Hpa*I site or to the right of the *Spl*I site do not confer any defects on the Rep function (pTi-2). Deletions extending from the *Spl*I site inward to the *Stu*I or *Hind*III site and the 8-bp *Xba*I linker insertion at the *Stu*I site completely abolished the Rep function (Rep^{-4}). However, deletions from beyond the *Hpa*I site inwards to the *Xho*I, *Bam*HI, or *Hind*III site resulted only in an attenuation of the Rep function (Rep^{-1} , Rep^{-2} , or Rep^{-3}). A similar attenuation was observed with mutants containing 4-bp insertions at the *Mlu*I and *Bam*HI sites. The extent of attenuation observed with the deletion mutants differed, depending on the orientation of pTi DNA. Generally, mutants with polarity II exhibited a more intensive Rep^- phenotype than did the corresponding mutants with polarity I. In contrast, the 4-bp insertion mutants had no such polarity dependency and were less intensively Rep^- than were the corresponding deletion mutants (see Discussion).

The stabilities of pTi-1 to pTi-5, pTi-9, and pTi-10 were examined by growing GV3101 cells harboring each of the plasmids in the absence of kanamycin for selection. Host cells of Rep^+ plasmids generated plasmid-free derivatives at a frequency no greater than 0.2% during a growth period of 15 generations under the conditions used. However, host cells of Rep^{-1} , Rep^{-2} , and Rep^{-3} plasmids segregated plasmid-free cells at frequencies of 8 to 99.5%, depending on the class of Rep^- phenotype (8 to 65% for Rep^{-1} , 70 to 85%

for Rep^{-2} , and 95 to 99.5% for Rep^{-3}). The yield of covalently closed circular DNA of each plasmid was determined by both cesium chloride-ethidium bromide density gradient centrifugation and the rapid plasmid detection procedure. The results (data not shown) indicated that all of the Rep^{-1} and Rep^{-2} plasmids yielded fewer covalently closed circular molecules than did the Rep^+ plasmids of pTi-1 and pTi-2 after correcting for the number of plasmid-free cells. Those of the Rep^{-3} plasmids could not be determined because of the extremely high frequency of curing even under selective conditions. Therefore, the Rep^{-1} and Rep^{-2} phenotypes and probably also the Rep^{-3} phenotype of these mutants appeared to be due to the reduction of initiation frequency and not defects in partitioning. From these experimental results, it was concluded that the right half of the replicator region was indispensable for the Rep function and contained the replication origin, while the left half had an auxiliary role in enhancing the Rep function, confirming the previous work (12).

Incompatibility properties of pTi-2 and pTi-5. The incompatibility properties of pTi-2 were tested by introduction of the plasmid (Km^r selection) into *Agrobacterium* cells carrying pAO359 (a Cb^r Km^s plasmid equivalent to pTi-1), pAL2821, pTiC58, pTiB6S3, pTiA6, pTiBo542, or pAO224 (see Materials and Methods). pTi-2 exerted incompatibility toward all of the plasmids except for pTiBo542 and pAO224 by ejecting the resident plasmid or by forming a cointegrate with the resident plasmid. pTi-2 was completely compatible with pAO224, and both plasmids were stably maintained without selective antibiotics, as were the parental pTiB6S3 and pTiA4b (2, 20). Bacterial cells carrying pTi-2 and pTiBo542 lost pTi-2 but not pTiBo542 at a considerable frequency in the absence of kanamycin, and pTi-2 behaved

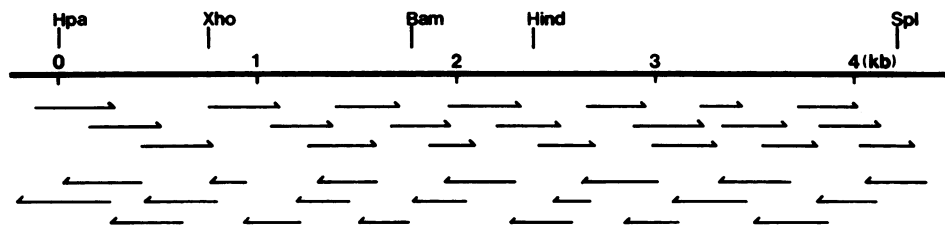


FIG. 2. Sequencing strategy. The top is the same as that of Fig. 1. Sequenced regions are shown by arrows in the 5'-to-3' direction.

as if it were a Rep⁻¹ mutant, suggesting that pTi-2 is partially incompatible with pTiBo542. A smaller plasmid, pTi-5, exhibited the same incompatibility characteristics as pTi-2. Therefore, the incompatibility functions of pTiB6S3 (or at least the most important one) are determined by the *HindIII-SpII* region of 1.8 kb. Conversely, when pAO359 or pAO224 was introduced by transformation (Cb^r selection) into *Agrobacterium* cells carrying pTi-2 or pTi-5, each resident plasmid was expelled by pAO359 but not by pAO224, indicating that this same 1.8-kb region also contains the target site for pTiB6S3 incompatibility.

Nucleotide sequence of the *HpaI-SpII* replicator region. To clarify the gene organization of the replicator region, we sequenced both strands of the 4.2-kb *HpaI-SpII* region in their entirety by the chain termination method (18). The sequencing strategy is shown in Fig. 2. By arranging overlapping sequences of both strands, we determined the nucleotide sequence of the *HpaI-SpII* region (Fig. 3). The nucleotide numbers used in Fig. 3 correspond to the scales in Fig. 1 and 2. Judging from the relative locations of initiation and termination codons for protein synthesis in all of the possible reading frames, the 4.2-kb region encodes three high-molecular-mass polypeptides of 44.5, 37.9, and 48.4 kilodaltons (kDa), all of which are transcribed in the same orientation, from left to right (Fig. 1). These proteins were named RepA, RepB, and RepC, respectively, and their amino acid sequences are shown under the nucleotide sequence in Fig. 3. The amino-terminal amino acid residues for the three Rep proteins were not directly determined, and the initiation codon for each protein was assumed to be the 5'-most one, which was accompanied by a potential ribosome-binding sequence (19). No other open reading frame longer than 400 bp was found in this 4.2-kb region. Furthermore, no significant direct and inverted repeats, which are frequently located at the origin regions of many replicons, were present.

Identification of the three Rep proteins. In an attempt to confirm the nucleotide sequence data and to visualize the three putative proteins, we synthesized polypeptides with an *E. coli*-coupled transcription-translation system (22). The exogenous templates used were covalently closed circular plasmid DNAs in which restriction fragments (nucleotide positions -577 to 2391, nucleotide positions 1489 to 3175, and nucleotide positions 2386 to 4568) containing each of the putative *repA*, *repB*, and *repC* genes were separately placed downstream of either the *lac* promoter on pAO254 and pHSG298 or the *tac* promoter on pKK223-3. The polypeptides synthesized were resolved on sodium dodecyl sulfate-polyacrylamide gels (Fig. 4). pAO254::*repC* and pAO254::*repA* directed the synthesis of proteins migrating at approximately 48 and 42 kDa, respectively, in addition to a protein of 26 kDa (Fig. 4, lanes 1 and 2). pHSG298::*repB* produced bands of 38 and 28 or 27 kDa (Fig. 4, lane 3). pKK223-3::*repC* and pKK223-3::*repB* produced, besides a

29-kDa band, the 48-kDa band and the 38-kDa band, respectively, but at severalfold-higher intensities than those observed with the *lac* promoter (Fig. 4, lanes 4 and 5). The 26-, 28-, or 27-, and 29-kDa bands presumably correspond to the gene products of Tn5 *kan*, Tn903 *kan*, and Tn3 *bla*, respectively, because they also appeared with the vector DNAs of pAO254, pHSG298, and pKK223-3, respectively (data not shown except for pKK223-3). The insertion of an oligonucleotide at the *MluI* site in *repA*, at the *BamHI* site in *repB*, or at the *StuI* site in *repC* led to the disappearance of the 42-, 38-, or 48-kDa band, respectively (data not shown). Furthermore, the observed molecular sizes were consistent with those calculated from the nucleotide sequence data. These results clearly indicate that the 42-, 38-, and 48-kDa bands correspond to the RepA, RepB, and RepC proteins, respectively. When the *rep* genes were improperly oriented in the expression vector, no Rep protein bands or only trace amounts of Rep protein bands were detected (data not shown), revealing that the majority of *rep* expression observed here occurred from the *lac* and *tac* promoters and not from the *rep* promoters.

Complementation between pTi and pRi *rep* genes. The size, arrangement, and mutational effects of the pTi *repABC* genes described above were quite similar to those of pRiA4b previously reported (17). The amino acid sequences of the RepA, RepB, and RepC proteins in these two plasmids were not, however, so highly conserved, as can be seen in the dot matrix analysis (Fig. 5). To examine whether each of these Rep proteins is interchangeable with the corresponding one in the other plasmid, we performed complementation tests as described in Materials and Methods. None of the *rep* mutations was complemented by the corresponding functions in the other plasmid. Since neither the oligonucleotide insertion mutants nor the deletion mutants could be complemented, the negative results were not due to additional defects in *cis*-acting functions carried by the *rep* deletion mutants (e.g., loss of the origin sequence). Thus, it was concluded that each of the Rep proteins of pTi fails to correct the defect caused by mutations in the corresponding genes of pRi and vice versa.

DISCUSSION

In this report, pAL2821, a previously constructed mini-pTiB6S3 plasmid (11), was joined to a ColE1-based vector, and the replicator region of pTiB6S3 was dissected by trimming the resulting chimeric plasmids or by introducing oligonucleotide insertion mutations. It was shown that the *HpaI-SpII* region of 4,203 bp was sufficient for the Rep function of pTiB6S3. Only three protein-coding frames longer than 400 bp were identified in this region from the nucleotide sequence data, as is the case for pRiA4b (17), and three polypeptides corresponding to these frames were identified by using an *E. coli*-coupled transcription-trans-

GTTAACCTCCATTAACCAAAAAACGGCTTGAACACGACCAACATTCGGTAATCGTCA	58	GCCAAAGGTATACCGCTGGATCTCGCTACACAAATAGGTCGCCGTCAAAAGCTGGCCGG	2219
CGGAAAAATGGCGTATCGCTAATTAACCGTGACAAAGAAATCTATCGTGAGCAAAAGCC	118	TCCGGTTGGTTGCACTTGCAGAAAGTCTCGGAACTCAAAGCGCGGACCGCAATTGGT	2279
RepA V S K A	4	S R W V A L A E G L A N S K A A D A I G	253
GCTGCCATATCCAGGAATGATCGCCCGTGGTGCATGTAACATCGCCGAGCATCGAGAA	178	GGGTGCTTGAATCTGAGCTATTCGGGACGTCGATAGCGACACCCGCTTTCTCTGTT	2339
A A I S R N D R P S V D V T I G E H A E	24	G V I E S E L F R T S D S D T R F S I V	273
CAGTAAGTTCGACGCTTCAAGCGATGAGTGAGGCGCTGTTCCGCCCACTTCGCAAAA	238	CTCAAGGCTGTTTTGAAACCTGCGGTAACACCTCAAAGAGCGTGAAGCTTGGAGCATG	2399
Q L S S Q L Q A M S E A I F P P T S H K	44	L K A V L K P A V K P S K S V K A W S M	293
ACGTTCCGAAATTTACGTCAGGTGAGGCGCAGCCCTCATGAAGATATCCGACTCGACT	298	CCCAAAGGTAAAAAGCGGCTGTAATCCAAAAGATACCGGACGCAAGCGCTTGTATTTT	2459
T I R K F T S G E A A R L N K I S D S T	64	P K G K K A A V I Q K D T G R T A L I F	313
CTTCGAAAAATGACGTTGGCGGGAAAGGCCGACGCTGAGCTGCCAGTAACGGAAG	358	GACGAAAACTGGTGCCTGATTTGGCGAATTCGTTGCTGACCACTGGACCGTCTGTAT	2519
I R K N T L A G E G P Q P E I A S N G R	84	D E K L V P A F G E F V A D Q L D R I Y	333
CGCTTTTACACGCTCGCCAGATAAATGAGATCGCGGATGCTGCCCTTCGACGCGG	418	GCCCAATTCATAGAAAACGCGATGGGGAAGCTCGCAATAACCAATTTCCATCGAT	2580
R F Y T L G Q I N E I R G M L A R S T R	104	A Q T T I E T T D G E K I D Q *	
GGACGTGAAAGCATCGAGTTTGTACTCATCGTAGAGTTCTGAGCATTGCAAGTCATC	478	CCGAGGTTCACTCAGATGAGATGAAATGAACCGCAAAAAGAAAAGGCCCGCAACGCGG	2640
G R E S I E F V I H R R G S L D P Q A S I	124		
CGAGTGACCAACTCAAAGGTGGATCTGGGAAGCAGCAGCTTCGGCTCACCTTGCACAG	538	TTCCGGAAGACCTTCTCTATAGTCTCGCAGCTAAGAGAATCGCATTTTCAGGAATCCCAG	2700
A V T N F K G G S G K T T T S A H I A Q	144		
TATCTCGGTTGCAAGGATATAGGTCCTAGCCGTTGACCTCGATCCGAGCCAGCTTGG	598	TCAAGAGTCCCGTGAGGAAGATCGTTTCGACGAGCTGATTTCTTTGCTAACAAAGG	2760
Y I A I Q G Y R V I A A R L N K I S D S T	164		
TCAGCACTCTCGCGTTCTGCCGGAACGGATGAGGCGGAATGAGACGCTTTCAGCC	658	TACAAGGAAATGCAGACGCTTTTCAACGACGCCCTTTGGCGCGGCGCGTACTCTC	2820
S A L I G V I P E T D V G A N E T L Y A	184	RepC M Q T H L S T T P F G R R P N T L	17
GCTATTCGTTATGACGATACTCGACGTCGTTGGGAGACGATCCGTCGCAATATTTT	718	GGCCAGATTTCAAGTCAGATGTCAGCAAGGCGGTGGCGTACGCTACCCGAAATAA	2880
A I R Y D D T R P I R P I R P T Y F	204	G Q I S S Q M S A K A V A P D A A N A N K	37
GATGGCTGCATCTGTACCGGAAATCTCGAGCTCATGGATTTGAGCAGCAGCACCA	778	TGGCATGTGTTCCAGAATATCCGGGAGCGGAGGAACTGCTCGGCCAACGGATCGCTCG	2940
D G I H L V P G N L R L M E F E H T T P	224	W H V F Q N I R E A R E L L G A T D R S	57
AAGCGTTGAGTGATCGAGGCGCGCGGCTGTTGTTCTTACACCGCTGGCTCAAGCC	838	CTTGGATCCCTCAATGCTTGTGACCTTTTCAATCCCGAGACAACTTACCGCGGATGGT	3000
K A I S D R G G R D G L F F T R V A Q A	244	L A I L N A L L T F H P E T T L T G D G	77
TTTGACGAGGTAGGCGACGATTACGATGTCGTTGTTATCGACTGCCCTCCTAGCTCGG	898	GAAATCTGTATGGCCATCAACGACGACTAGCGCACCGCCAAATGGCATGCGCGG	3060
F D E V G D D Y D V V V I D C P P Q L G	264	E I I V W P S N E Q L A A R A N G M P A	97
TTTCTGACTCTCAGCGGTTATGCTGCGCAGCAATGGTGGTTACGGTGCACCCGCAA	958	ACGACGTTGCTCGGCATCTTGCAGTTCTCGTGGAGTGGCGCTGCTCATCAGGCGGAT	3120
F I T I S G L C A A T A M V V T V H P Q	284	T T I R R H L A V L V E C G L V I R R D	117
ATGCTGATATTCGCTCAATGAGTCAGTTTCTGCTCATGACCCGTCGCTTCTTGGCGT	1018	AGCCCGAACCGTAAACGCTTCCGCGGTAAGGCGAGGGTGGCGAAATCGAGCAGCCCTAC	3180
I L D I A S M S Q F L I M T R D L I G V	304	N S K P D S I P E S E Y G F G N K P E A	137
GTGAGAGGGCGGAGGAACTTCTCAATACGACTTTATACGTTATCTTTGACCGCTAT	1078	GGCTTCGATCTGCTCCGATCGTGGCGCGCAAAAGAATTCAGAGATATGGCGAAGCG	3240
V R E A G G N L Q Y D P I R Y I L T R Y	324	G P D L S P I V A R A K E F R D M A E A	157
GAGCTCAAGATGCTCCGAAACGAAAGTCCGCTCGCTCCTCGAAACATGTTGAGGAC	1138	ATCCAGGCTGAGAAGAAAGCTTCCGCTGCGCAAGGAGCGGCTGACGCTGTTACCGCGT	3300
E P Q D A P Q T K V A A L I R N M F E D	344	I Q A E K K A F R V A K E R L L R R	177
CATGTTCTCAAAATCTATGGTCAAGTCCGCGCGTGTGATCGGGTTAACTAAG	1198	GATATCGTGAAGTTGATCGATCGGGATCGAAGAGGGCGTCCCGGCAACTGGTGGGA	3360
H V L T N P M V K S A A V S D A G L T K	364	D I V K L I D A G I E E G V P G N W C G	197
CAGACTCTGACGAGATCGGAGAGAAACCTGACCCGTTCAACTACGACCGGGCCATG	1258	GTGCAGCGGTTGATCAAGCGATTATCGGACGCTTACCGGCTCCGACCAAGACAGCTT	3420
Q T L Y E I G R E N L T R S T Y D R A M	384	V Q Q G V Y Q A I I G R L P R S A P R Q L	217
GAGTCGCTGATCGCGTGAACCGAGAGATCGAAGCGTTGATCAAGCGCGGTTGGGAAG	1318	GTGGAGGAGTCTGCATCGGTTCTCATGCGCTGTACATAGAAATCCGTCGCTTGGAA	3480
E S L D A V N A E I E A L I K A A W G R	404	V R E I C I G L H A L Y A I E I R D V L E	237
GCCTGATGACACTATCTCCCATCGTGGCGGTTGTTGGAGCTCCCAACCAAGTTTATT	1379	TCTTTCGCAAAAACAGATTACGACCGCAATGAGTCCCATTTTGGTCTGCATACAG	3540
A * *		S F A K T Q I Q D A N E S H F G R H I Q	257
CAGTGACCACTCGGGAAGCACTCTGGTATATCGCCGAAGATAGAACTCACTTGAGGA	1439	AATCAAACAGACTATACCTGAATCTGAATACCGCTTGGAAATAAACAGAGCGG	3600
AGGCCACGCTGGCTGGCCGGTTCGCGGTACATTTGCTGTTGGGAGCTCCCAACGGATC	1499	N S K P D S I P E S E Y G F G N K P E A	277
GGCCGTTGCTTCAAGGGCCATCGGAGCAACGCAAAACGGAAGAGACTTTTATCATGAGC	1559	GGCGGACGGTTGAGGAATTCGACAACGTCGGAAGCCTGCCAAGCGGGAATGCCATTA	3660
RepB N A E R Q T E E T F I M S	13	G G T V E E F D N V R S L P K R E L P I	297
CGTAAAGACGCAATGACTCGCTGTTCTCAAGAAGCAACCTTCGACGGAAGGCCATCC	1619	GAATGCTGCTGAATGCTGCGGAGCTGGAATGGCTGAGGTTGGCGAAATCCCG	3720
R K D A I D S L F L K K Q P S T E R P S	33	G I V L N A C P S V L E L A Q G G E I R	317
ATCGATAAGTCCGCGTTCGCGTTCGAGCTGGTGGATCTCTCGGATGGGTTCTTCGCTG	1679	CATTGGCGGATTTCTGGGACTTCGGAATGCTCGCCGATGCTGGGTATCAGCCAG	3780
I D K S A V R V R T G A I S A M G S S I	53	H W R D F L A T V E L A R P N L G I S Q	337
CAAGTATGGCCGAGGGAGCAAAAAGCCGATCTCGGTTGAGGATCAATGGCTCGCGGT	1739	AGCCGCTGGCGGGAAGCACTCGATGAGCTGGCGAGCAGCATGCGGCAATCAGCTTGG	3840
Q V M A E G A K A A S R I Q D Q I A A G	73	S A W R E A L D E L G E Q H A A I T L A	357
GAAACCGTCTGCTTTGGATCCGTCATGATGACGGTGCAGATACGAGCCGCTC	1799	CGCATCTATCAGAAGCGGACAGATTGGATCGGCGGCGGATCTTGGTAACTGACA	3900
E T V V S L D P S M I D G S Q I T D R L	93	A I Y Q K A D Q I G S A G G I L R N L T	377
CCCGTTGACGTCGATCCGAAGTTCGATCAACTGAAAGCGAGCATTTACAGGATGGGCG	1859	GATCGAGCCGCTGATGGTAAATTCACAGTGGCGGATGATCATGGCACTGCTGGGGCT	3960
P V D V D P K F D Q I E A S I S Q D G Q	113	D R A R D G K F S T W P M I M A L L R A	397
CAAGTCAACTCCTGTTGGCCGCACTCTGACACCACCGGTCGATATCAAACTCGTGTAT	1919	AAACTGTATGCGGAGAAAGGTGAAGGCTGAGCATCGGCTGCCCGCTTGAACGAGACGGCC	4020
Q V P I L V R P H P E T T G R Y Q I V Y	133	K I L D A Q K V K A D D R L P A L N E T A	417
GGCAGACGGCGGTTGCGCGCAGCCCAATACGAAAGAGAGTTTCCGCAATGTCGGA	1979	GACAATGGATCAGGGCTGGCGGATCCGATGCTGAGAACTGCTGGCAAGTCGAG	4080
G R R L R L R A A A N L R R E V S A I V R	153	D N G S G L R A S D A L L R T L G K S R	437
AATCTCACGGACCGCAACTGGTCTGGCAGCGGCGTGAAGTCTGATCGTGGCGGAC	2039	CCGAAATGAATATCCGGTTAGGCTGTTGCGGCTGCACGGGAGCGGCTGCTGATCTCG	4140
N L T D R E L V V A Q G R E N L D R A D	173	P K *	
CTCTCCTCATTGAGAAGCGGCTTTTCCGCTGCGCCTTGAAGAGGCGGGTTTATCGG	2099	CGCGGATGATCAGGTTGTTGACCGACGAAGGAGCAATTTGGTGCCACCAACCGCGCT	4200
L S F I E K A F A L R I E E A G F D R	193	ACG	
CCAACGGTCAATGCGGCACTATCGACCGCAAAAGCCGACTTACCGCTATATAGCTGTG	2159		
P T V I A A L S T D K A D L S R Y I A V	213		

FIG. 3. Nucleotide sequence of the replicator region. The sequence from the *HpaI* to *SplI* sites of the coding strand (5' to 3') is shown with the amino acid sequences of the three Rep proteins. An asterisk indicates a termination codon. A possible ribosome-binding sequence is overlined. The nucleotides are numbered from the *HpaI* site toward the right, and the nucleotide positions correspond to the scales in Fig. 1 and 2.

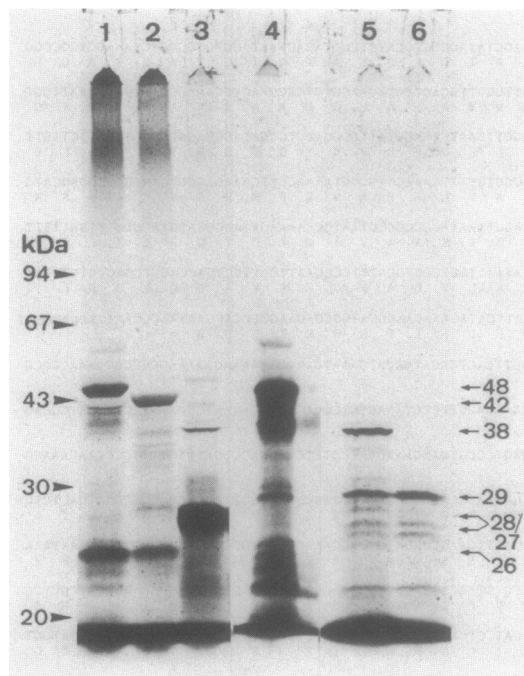


FIG. 4. Polypeptides synthesized in an in vitro-coupled transcription-translation system. Polypeptides were labeled with L-[³⁵S]methionine and analyzed by 12.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis. The band positions of marker proteins are shown on the left. On the right are shown the molecular masses for each band displayed by template DNA of pAO551 (pAO254 *lac* promoter::*repC*) (lane 1), pAO503 (pAO254 *lac* promoter::*repA*) (lane 2), pAO543 (pHSG298 *lac* promoter::*repB*) (lane 3), pAO517 (*tac* promoter::*repC*) (lane 4), pAO553 (*tac* promoter::*repB*) (lane 5), and pKK223-3 (lane 6).

lation system in vitro. Since there are several examples indicating that the signals for the initiation of transcription and translation in *E. coli* and *A. tumefaciens* overlap at least partly (e.g., see references 3, 5, and 19), these proteins are likely to be synthesized in *Agrobacterium* cells, too, and directly contribute to the Rep function. However, their actual expression in *Agrobacterium* cells would occur at a low efficiency, because the synthesis of three Rep proteins was well detected in vitro only when *rep* was placed

in the appropriate direction downstream of the *lac* or *tac* promoter.

Both of the deletions in *repC* and the 8-bp insertion in *repC* completely destroyed the Rep function. In addition, pTi-5 and pTi-4, each carrying only *repC* and truncated *repB*, were able to replicate, although at a low efficiency. Therefore, RepC is likely to be the principal initiation protein.

In contrast to the marked effects of *repC* mutations on the Rep function, disruption of *repA* or *repB* by an oligonucleotide insertion led to only a slight instability of the plasmids in *Agrobacterium* cells. Deletions in *repA* or *repB*, however, affected plasmid stability more seriously than did oligonucleotide insertions, although there was little correlation between the deletion size and the instability of the mutant plasmids. For instance, pTi-4, which contained a shorter deletion than pTi-5, was more intensively Rep⁻ than was pTi-5. Furthermore, a given *repAB* deletion mutant displayed a different class of Rep⁻ phenotype, depending on the polarity, indicating that the Rep function of such deletion mutants was influenced by the neighboring non-pTi DNA sequences. The direct effect caused by the absence of RepA and RepB (pTi-9 and pTi-10) was only feeble (Rep⁻¹) and, therefore, the strong defects caused by deletions in *repAB* (pTi-3, pTi-4, and pTi-5) seemed to be chiefly due to the lack of some *cis*-acting function(s) (see below) rather than to the lack of the RepA and RepB functions themselves. Generally, a deletion mutant with polarity II was more intensively Rep⁻ than was the same mutant with polarity I (Fig. 1). On the other hand, the corresponding deletion mutants constructed with the pAO260 vector instead of with the pAO254 vector showed the opposite polarity dependency (*kan* and *lacZ_α* were oriented head-to-head in pAO254 and head-to-tail in pAO260, and polarity II was always defined as the same direction of *rep* as *lacZ_α*), that is to say, a chimeric plasmid in which the *rep* genes were properly positioned downstream of *kan* exhibited a weaker Rep⁻ phenotype, regardless of the direction of *lacZ_α* (unpublished data). Thus, the transcription readthrough from *kan* seems partially to substitute for the *cis*-acting function(s) determined by the deleted portions in the *repAB* mutants. This might result from the activation of the replication origin, as was observed with the bacteriophage lambda and *E. coli* origins (4, 15), or from a higher expression of *repC* than that from its own promoter alone. The *cis*-acting function(s) might originally be provided by the promoters for *repA* and *repB*. As was suggested from the reduction of plasmid DNA yield, the instability of *repAB*

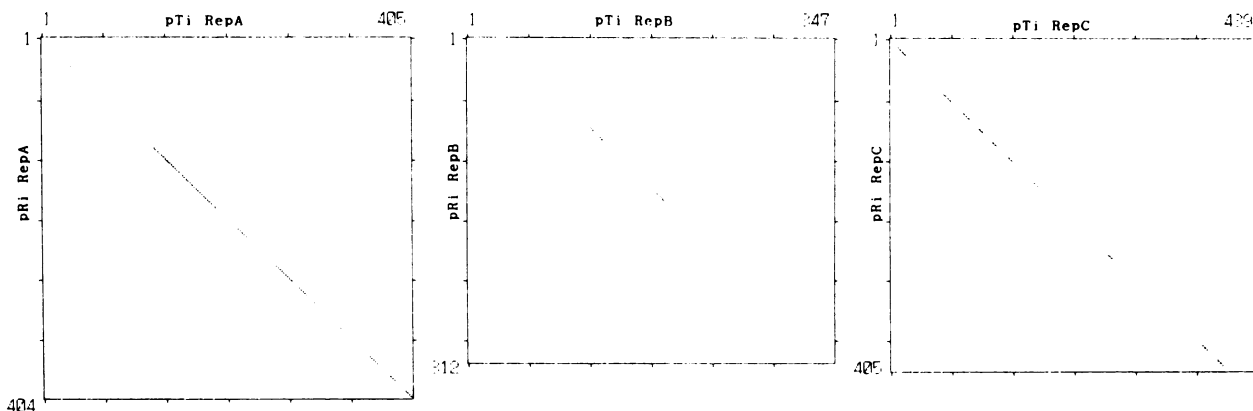


FIG. 5. Comparison of amino acid sequences of the RepA, RepB, and RepC proteins from pTiB6S3 and pRiA4b by dot matrix analysis. A dot represents the presence of 7 identical residues in each of 10 consecutive residues.

mutants was probably due to the reduction of initiation frequency and not to the defect at partitioning. Therefore, we speculate that RepA and RepB proteins control the unwinding of DNA strands to allow efficient interaction of RepC with the origin or that they operate as accessories supporting normal DNA initiation by RepC. The mutational effects of *repA* and *repB* appear to be nearly the same, but no significant sequence similarity between the two genes was found.

pTi-2 and pTi-5 exerted incompatibility toward pAO359, pAL2821, pTiC58, pTiA6, and pTiB6S3 by either expelling the resident plasmid or forming a cointegrate with the resident plasmid. When pAO359 was introduced into cells harboring pTi-2 or pTi-5, the resident plasmids were ejected by the incoming plasmid, indicating that they are sensitive to incompatibility exerted by pTiB6S3. Therefore, it was concluded that both the incompatibility determinant(s) and its target site, which might overlap with each other, were present on the 1.8-kb *HindIII-SpII* region. In addition, we recently found that a small DNA fragment encompassing the spacer between *repB* and *repC* exerted incompatibility toward pTiB6S3, although it could no longer replicate by itself. A corresponding DNA fragment derived from pRiA4b also displayed incompatibility toward the parental pRiA4b (manuscript in preparation). Competition for the RepC initiation protein at the replication origin and/or the *repC* promoter might be the actual cause of the incompatibility of pTi and pRi.

pTi-2 was lost in bacterial cells carrying supervirulent pTiBo542 at a considerable frequency, as if it were a Rep⁻¹ mutant, suggesting that these two plasmids are partially incompatible. Since this result is inconsistent with an earlier report that parental pTiB6S3 is completely compatible with pTiBo542 (13), this phenomenon might be specific for pTi-2; it is also possible that the pTiBo542-carrying strain, AT1018, which has a chromosomal background nonisogenic with GV3101, is partly defective in pTiB6S3 replication.

Although pTiB6S3 and pRiA4b were completely compatible and their respective miniplasmids also were compatible, the gene size, organization, and possible functions of the replicator region in the two plasmids were nearly identical (17). However, the amino acid sequence of each Rep protein was not well conserved, and no complementation was detected between the Rep functions of these two plasmids. Thus, it appears that the three Rep proteins of pTi and pRi are functionally too specialized to support complementation, although the unlikely possibility that RepABC proteins are all *cis* acting cannot be completely ruled out at present. These results are in contrast to the fact that the structure and function of *vir* genes are well conserved and mostly complementary between the two plasmids (5, 8).

In plant genetic engineering technology, the binary vector strategy has commonly been used (e.g., see references 1 and 6). Generally, one plasmid derived from pTi is a supplier of Vir function, and the other plasmid carries T-DNA border repeats flanking cloning sites. In most cases, the latter is a broad-host-range plasmid vector derived from either RK2 or pSa. Such vectors, which are generally larger than 15 kb because of the relatively large size of their replicators, are lost from *Agrobacterium* cells at a higher frequency than are the mini-pTi or mini-pRi plasmids described here (unpublished data). Furthermore, these chimeric miniplasmids are about 7 kb long, easily allowing the cloning of foreign DNA fragments. Therefore, these chimeras are good candidates for *Agrobacterium* vectors. Two sets of vectors that are completely compatible with each other are now available.

Using these, we generally obtain better results in complementation analyses with mini-pTi chimeric vectors than with RK2-based vectors (manuscript in preparation).

ACKNOWLEDGMENTS

We are grateful to E. E. Hood, Y. Machida, M. van Montagu, and E. W. Nester for gifts of bacterial and plasmid strains.

This work was supported in part by grants-in-aid for scientific research from the Ministry of Education, Science and Culture of Japan (grant 62480463) and from Yamada Science Foundation.

LITERATURE CITED

- An, G., B. D. Watson, S. Stachel, M. P. Gordon, and E. W. Nester. 1985. New cloning vehicles for transformation of higher plants. *EMBO J.* 4:277-284.
- Constantino, P., P. J. J. Hooykaas, H. den Dulk-Ras, and R. A. Schilperoort. 1980. Tumor formation and rhizogenicity of *Agrobacterium rhizogenes* carrying Ti plasmids. *Gene* 11:79-87.
- Das, A., S. Stachel, P. Ebert, P. Allenza, A. Montoya, and E. Nester. 1986. Promoters of *Agrobacterium tumefaciens* Ti-plasmid virulence genes. *Nucleic Acids Res.* 14:1355-1364.
- Furth, M. E., and W. F. Dove. 1982. Specificity determinants for bacteriophage λ DNA replication. III. Activation of replication in λrI^c mutants by transcription outside of *ori*. *J. Mol. Biol.* 154:65-83.
- Hirayama, T., T. Muranaka, H. Ohkawa, and A. Oka. 1988. Organization and characterization of the *virCD* genes from *Agrobacterium rhizogenes*. *Mol. Gen. Genet.* 213:229-237.
- Hoekema, A., P. R. Hirsch, P. J. J. Hooykaas, and R. A. Schilperoort. 1983. A binary plant vector strategy based on separation of *vir*- and T-region of the *Agrobacterium tumefaciens* Ti-plasmid. *Nature (London)* 303:179-180.
- 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.
- Hooykaas, P. J. J., M. Hofker, H. den Dulk-Ras, and R. A. Schilperoort. 1984. A comparison of virulence determinants in an octopine Ti plasmid, a nopaline Ti plasmid and an Ri plasmid by complementation analysis of *Agrobacterium tumefaciens* mutants. *Plasmid* 11:195-205.
- Hooykaas, P. J. J., and R. A. Schilperoort. 1984. The molecular genetics of crown gall tumorigenesis. *Adv. Genet.* 22:209-283.
- Jouanin, L., F. Vilaine, C. d'Enfert, and F. Casse-Delbart. 1985. Localization and restriction maps of the replication origin regions of the plasmids of *Agrobacterium rhizogenes* strain A4. *Mol. Gen. Genet.* 201:370-374.
- Koekman, B. P., P. J. J. Hooykaas, and R. A. Schilperoort. 1980. Localization of the replication control region on the physical map of the octopine Ti plasmid. *Plasmid* 4:184-195.
- Koekman, B. P., P. J. J. Hooykaas, and R. A. Schilperoort. 1982. A functional map of the replicator region of the octopine Ti plasmid. *Plasmid* 7:119-132.
- Komari, T., W. Halperin, and E. W. Nester. 1986. Physical and functional map of supervirulent *Agrobacterium tumefaciens* tumor-inducing plasmid pTiBo542. *J. Bacteriol.* 166:88-94.
- Laemmli, U. K. 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* 227:680-685.
- Lark, K. G. 1972. Evidence for the direct involvement of RNA in the initiation of DNA replication in *Escherichia coli* 15T⁻. *J. Mol. Biol.* 64:47-60.
- Nester, E. W., M. P. Gordon, R. M. Amasino, and M. F. Yanofsky. 1984. Crown gall: a molecular and physical analysis. *Annu. Rev. Plant Physiol.* 35:387-413.
- Nishiguchi, R., M. Takanami, and A. Oka. 1987. Characterization and sequence determination of the replicator region in the hairy-root-inducing plasmid pRiA4b. *Mol. Gen. Genet.* 206:1-8.
- Sanger, F. 1981. Determination of nucleotide sequence in DNA. *Science* 214:1205-1210.
- Shine, J., and L. Dalgarno. 1974. The 3'-terminal sequence of

- Escherichia coli* 16S ribosomal RNA: complementarity to non-sense triplets and ribosome binding sites. Proc. Natl. Acad. Sci. USA 71:1342-1346.
20. White, F. F., and E. W. Nester. 1980. Relationship of plasmids responsible for hairy root and crown gall tumorigenicity. J. Bacteriol. 144:710-720.
 21. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.
 22. Zubay, G. 1973. In vitro synthesis of protein in a microbial system. Annu. Rev. Genet. 7:267-287.