## The Right End of the *vir* Region of an Octopine-Type Ti Plasmid Contains Four New Members of the *vir* Regulon That Are Not Essential for Pathogenesis

VIRGINIA S. KALOGERAKI, JUN ZHU, JOEL L. STRYKER, AND STEPHEN C. WINANS\* *Department of Microbiology, Cornell University, Ithaca, New York, 14853*

Received 15 September 1999/Accepted 15 December 1999

**We sequenced the** *virD-virE***,** *virE-virF***, and** *virF***–T-DNA intergenic regions of an octopine Ti plasmid. Four newly described genes were induced by the** *vir* **gene inducer acetosyringone, two of which are conserved in the nopaline-type Ti plasmid pTiC58. One gene resembles a family of phosphatase genes. Each of these genes is dispensible for tumorigenesis.**

Infection and colonization of plant or animal hosts often require a multifaceted attack against the host. For example, many pathogenic bacteria, including *Bordetella pertussis*, *Yersinia* spp., and *Pseudomonas syringae*, release multiple toxins and other virulence factors, any one of which may be dispensable for pathogenesis (5, 7, 8). It seems plausible that the plant pathogen *Agrobacterium tumefaciens* might similarly use more than one approach to invade and colonize host plants. This bacterium is well-known to transfer fragments of plasmid-encoded DNA (T-DNA) by a conjugation-like mechanism to the nuclei of host plants, where they are integrated into the host genome (23). This T-DNA directs the production of phytohormones (leading to formation of crown gall tumors) and of compounds called opines, which serve the colonizing bacteria as sources of nutrients. While it is not clear whether *A. tumefaciens* uses additional strategies to attack host plants, another species of *Agrobacterium* (*A. vitus*) releases a root macerating pectinase, as well as transferring T-DNA (22). This pectinase is not required for detectable tumorigenesis, while T-DNA transfer is not required for root maceration, indicating that this organism uses at least two independent strategies for pathogenesis.

One way to search for undescribed strategies for host infection and colonization is to identify genes that are induced during infection and not required for T-DNA transfer. Approximately 25 genes, called *vir* genes, are required for T-DNA transfer (23, 25). All *vir* genes are coordinately induced during infection by plant-released chemical signals, including phenolic compounds such as acetosyringone (11). This induction requires the sensory histidine protein kinase VirA, the response regulator VirG, and the periplasmic sugar-binding protein ChvE. Importantly, several of the genes in this regulon, including *virH*, *virK*, *virL*, and *virM*, are not required for efficient tumorigenesis (15). While it is possible that some or all of these genes play ancillary, dispensable roles in T-DNA transfer, at least some of these genes may direct other processes. We have recently demonstrated that the VirH2 protein catalyzes the O demethylation of several phenolic compounds, converting them to forms that are inactive as *vir* gene inducers (16). The functions of *virK*, *virL*, and *virM* are not understood, although *virK* strongly resembles a gene found on a symbiotic megaplas-

\* Corresponding author. Mailing address: Department of Microbiology, Wing Hall, Cornell University, Ithaca, NY 14853. Phone: (607) 255-2413. Fax: (607) 255-3904. E-mail: scw2@cornell.edu.

mid of *Rhizobium* sp. strain NGR234 and is thus unlikely to have any direct role in T-DNA transfer.

As part of an ongoing effort to identify new genes that are part of the *vir* regulon and to study their roles in pathogenesis, we have sequenced approximately 15 kb of DNA within and beyond the right end of the known *vir* region (Fig. 1). In doing so, we have closed all the gaps in DNA sequence between known *vir* loci in this region and closed the gap between the right end of the *vir* region and the left end of the T-DNA.

We identified a total of 14 new open reading frames (ORFs) that might encode proteins (Fig. 1). Of these, ORFs 17 to 21 strongly resemble one or more genes found in various insertion sequences elements (Table 1). These ORFs were not characterized further. We used two methods to test for expression of the remaining nine ORFs. Many of them were first tested by nuclease S1 protection assays, using 5'-radiolabeled oligonucleotides that are complementary to each ORF (Table 2). RNA was purified from strain VIK10 cultured at pH 5.3 in the presence or absence of 100  $\mu$ M acetosyringone and from the *virG* mutant strain VIK11 cultured in the presence of acetosyringone. VIK10 contains a *virG-lacZ* fusion on the Ti plasmid as a result of a Campbell integration and retains a functional copy of *virG*, while VIK11 is isogenic except for a *virG* mutation (14, 15). Each oligonucleotide contained four noncomplementary nucleotides at its  $3'$  end that were predicted to be removed by S1 digestion. Removal of these nucleotides (causing slightly faster gel migration) ensured that the resistance of the remaining part of the oligonucleotide against S1 digestion was due to hybridization with mRNA. Oligonucleotides complementary to *virB1* and to *rpoD* were used as inducible and constitutive controls, respectively.

ORFs that were acetosyringone inducible by nuclease S1 protection assays (as well as two additional ORFs) were fused to *lacZ* by using a suicide plasmid that creates *lacZ* fusions and gene disruptions in a single Campbell-type insertion (14). Internal fragments of *virD5*, *virE3*, ORF12, ORF16, *virP*, and *virR* were created by PCR amplification, using the oligonucleotides indicated in Table 2, and cloned into suicide vector pVIK107 or pVIK111 (14), creating in-frame translational fusions with *lacZ*. The fact that these plasmids create translational fusions demonstrates that any induced gene must encode a protein. The resulting plasmids were transferred by conjugation from S17-1/ $\lambda$ *pir* into *A. tumefaciens* strain R10 and selected using kanamycin. To confirm that integration of these suicide plasmids occurred by Campbell-type homologous recombination, we digested genomic DNA of these *A. tumefa-*



FIG. 1. Genetic map of the right end of the *vir* region. *virD1* to -*4*, *virE1* and -*2*, and *virF* have been previously described. Newly described genes that are induced more than fivefold by *vir* gene-inducing stimuli are designated *vir* genes; these include *virD5*, *virE3*, *virP*, and *virR*. All other open reading frames are designated ORFs. Shaded bars designate newly sequenced regions.

*ciens* strains with *Eco*RI, circularized the resulting fragments by using T4 DNA ligase, introduced them into *E. coli* strain S17-1/*\pir* by electroporation, and analyzed the rescued plasmids by restriction endonuclease digestion. In each case, the restriction map of the recovered plasmid indicated that a single homologous recombination event had occurred at the predicted site (data not shown).

We identified an ORF directly downstream of *virD4* that is conserved in the nopaline-type Ti plasmid pTiC58 and in the Ri plasmid pRi4Ab of *A. rhizogenes* (Table 1). S1 protection experiments showed that this ORF was strongly induced by acetosyringone in a strain that expresses VirG but not in a *virG* mutant (Fig. 2, upper right panel). When this ORF was fused to *lacZ*, the resulting strain was strongly induced by acetosyringone (Table 3). An earlier study suggested that this region of the Ti plasmid did not contain any inducible genes (24). This conclusion was based on two insertions of Tn*3*HoHo1 that were not induced during cocultivation with *Nicotiana* cultured mesophyll cells. Similarly, the orthologous gene of pTiC58 was described as being expressed constitutively and not regulated by VirA and VirG (17). The reasons for these apparent discrepancies are not clear. Since this ORF was found to be a member of the *vir* regulon by two different criteria, we designate it a *vir* gene. We designate it *virD5* to suggest that it is transcribed as part of the *virD* operon, although this point remains to be proven, especially since the stop codon of *virD4* and the start codon of *virD5* are separated by 91 nucleotides.

The *virD5-lacZ* fusion was constructed so as to disrupt this gene. The fusion strain was therefore tested for tumor-forming ability on *Kalanchoe¨* leaves, and seemed to form tumors at efficiencies similar to the wild type (data not shown), in agreement with earlier studies (24). However, such assays are qualitative in nature, and a moderate change in tumorigenesis efficiency might therefore not have been detected. The predicted VirD5 protein has 751 amino acid residues, a molecular mass of 83.5 kDa, and is largely hydrophilic. Interestingly, the amino-terminal half of VirD5 is composed of a six repeated sequences, each approximately 50 amino acids in length. Figure 3 shows a dot matrix alignment of VirD5 with itself and an alignment of these repeats.

A second ORF was found directly downstream from *virE2*. This ORF is 673 codons in length and would encode a hydrophilic protein having a molecular mass of approximately 76 kDa. Nuclease S1 protection assays indicate that this gene is strongly induced by acetosyringone in a VirG-dependent manner (Fig. 2, second row, left panel), and a fusion between this gene and *lacZ* was strongly induced by acetosyringone (Table 3). We therefore designate this gene *virE3* to indicate that it is a member of the *vir* regulon and to suggest that it is part of the

*virE* operon, although this has not yet been demonstrated. The stop codon of *virE2* is separated from the start codon of *virE3* by 65 nucleotides.

*virE3* strongly resembles a partially sequenced ORF downstream of the *virE2* gene of the nopaline-type Ti plasmid pTiC58 (12). To compare these genes more completely and to look for additional conserved genes, we extended the sequence of the *virE* operon of pTiC58. The two VirE3 proteins have similar molecular masses and hydrophilicity profiles. Their protein sequences are quite similar from amino acids 1 to 350 (74% identical, 85% similar) but are less similar in their carboxy-terminal halves (37% identical, 51% similar). We sequenced an additional 2.3 kb of DNA to the right of *virE3*, up to a region that was sequenced by Farrand and colleagues, but did not identify additional genes that are conserved between octopine-type and nopaline-type Ti plasmids. The *virE3-lacZ* fusion plasmid disrupted the *virE3* gene but did not cause any qualitative tumorigenesis deficiencies on *Kalanchoe¨* leaves. This conclusion is supported by earlier studies showing that insertions in *virE2* can be fully complemented by a cloned DNA fragment that lacks *virE3*, indicating that the insertions were not polar on any genes that are essential for virulence (24).

Seven additional new ORFs between *virE3* and ORF17 were

TABLE 1. Sequence similarities between ORF products described in this study and other proteins

Protein	Homolog (GenBank accession no.)		$\%$ Simi- $larity^a$
VirD5	pTiC58 ORF5 (AAA91607)	62	72
VirD <sub>5</sub>	pRi4Ab VirD5 (CAA31354)	61	73
VirE3	pTiC58 VirE3 (AAA91610)	54	66
ORF <sub>12</sub>	Integration host factor $\beta$ subunit, R. capsulatus (0.06607)	50	72
ORF <sub>13</sub>	Cold shock protein, S. meliloti (AE0000068)	46	70
ORF <sub>14</sub>	O antigen acetylase, S. typhimurium (AAC45706)	26	45
ORF <sub>15</sub>	TraA, 3' end, A. tumefaciens (AAC28116)	86	90
ORF <sub>16</sub>	TraF, A. tumefaciens (AAC28117)	77	82
VirP	Phosphoglycolate phosphatase, Synechocystis (BAA17565)	29	54
VirR	Hypothetical protein, <i>Thermotoga</i> (AAD35280)	44	61
ORF <sub>17</sub>	Invertase, Moraxella (P20665)	26	44
ORF <sub>18</sub>	IstB, IS1326 (AAA79726)	62	83
ORF <sub>19</sub>	IstA, IS1326 (AAA79725) (amino acids 202 to 501)	58	71
ORF <sub>20</sub>	Transposase, IS1111a (AAA23313)	41	59
ORF <sub>21</sub>	IstA, IS1326 (AAA79725) (amino acids 1 to 198)	54	70

*<sup>a</sup>* Sequence identities and similarities were determined using the BLAST program (4).



J. BACTERIOL.

described, each showing some degree of sequence similarity to one or more proteins deposited in the GenBank or SwissProt protein sequence databases (Table 1). Of particular note, ORF15 strongly resembles a portion of the *traA* gene of the Ti plasmid (3). However, ORF15 appears to be a pseudogene, since its 5' end is severely truncated and the remainder of the ORF contains one frameshift difference from *traA*. ORF16 strongly resembles the Ti plasmid *traF* gene (3) and does not contain any obvious deleterious mutations. The similarity between ORF16 and *traF* was previously described (9). *virP* strongly resembles a family of known or putative phosphatases. The significance of the similarity is unknown but suggests a role in hydrolysis of phosphoryl groups from an unknown substrate. *virR* resembles a family of uncharacterized genes in various organisms, including several archael species and the higher plant *Arabidopsis thaliana*. An example is shown in Table 1.

We used nuclease S1 protection assays to test for acetosyringone-inducible expression of each of the seven ORFs between *virE3* and ORF17. Of these, *virP* and *virR* were strongly induced, while most other ORFs were not detectably induced, although ORF12 may have been very weakly induced (Fig. 2). To further test for induction, *lacZ* fusions to ORF12, ORF16, *virP*, and *virR* were constructed. Of these, the *virP-lacZ* and *virR-lacZ* fusions were strongly induced by acetosyringone, while the ORF12-*lacZ* fusion was only weakly induced, and the ORF16-*lacZ* fusion was not detectably expressed (Table 3). The possible induction of ORF12 was very weak and therefore difficult to interpret, and we therefore do not designate ORF12 as a member of the *vir* regulon. The resulting disruptions of ORF12, ORF16, *virP*, and *virR* were tested for tumorigenesis on *Kalanchoë*, and no deficiencies were detected.

VirG-inducible promoters generally contain one or more VirG binding motifs (TNCAATTGAAAPy) directly upstream of their 235 sequences (11). Sequence inspection suggests that the 5 9 ends of *virP* and *virR* contain possible *virG* binding motifs. The sequences TGTAATTGAATT and TACTGTTG AAAC are found centered at 306 nucleotides and 229 nucleotides upstream of the putative *virP* translation start site, respectively, while the sequence GACAATTGAAAT is found centered 68 nucleotides upstream of the putative *virR* translation start site. No VirG binding motif was found upstream of *virD5* or *virE3*, providing further suggestive evidence that these two genes are expressed as part of their respective operons.

It is interesting that of the four new members of the *vir* regulon, none was essential for tumorigenesis. It is certainly possible that one or more of these genes plays ancillary, dispensable roles in T-DNA transfer, and our tumorigenesis assays might well not detect moderate quantitative defects. However, our sequence of the corresponding region of pTiC58, combined with additional sequences from the Farrand lab (GenBank accession no. AF065243), indicates that *virP* and *virR* are not conserved in pTiC58. This suggests that they are unlikely to play an important role in DNA processing or transfer. It is also possible that these genes could be redundant with chromosomal genes, as is the case with *virJ* (13), although Southern hybridization did not detect homologous genes (data not shown). Furthermore, if *virP* does encode a phosphatase, it

FIG. 2. Nuclease S1 protection assays of ORFs described in this study. Lanes 1, synthetic radiolabelled oligonucleotide in the absence of mRNA or nuclease S1. Lanes 2 to 4, nuclease S1-resistant oligonucleotides after hybridization with mRNA from a  $virG^+$  strain cultured without acetosyringone (lane 2), from a  $\frac{virG^+}{\text{strain}}$  cultured with 100  $\mu$ M acetosringone (lane 3), or from a *virG* mutant strain cultured with 100  $\mu$ M acetosyringone (lane 4).







FIG. 3. A repeated amino acid sequence motif found within the VirD5 protein. The dot matrix plot of VirD5 against itself (A) was obtained using GenePro (Riverside Scientific Enterprises), while the alignment of these repeated sequences (B) was obtained using MegAlign (DNASTAR, Inc).

TABLE 3. Induction of *vir-lacZ* fusions by acetosyringone*<sup>a</sup>*

Strain	Suicide plasmid	Fusion	<b>B-Galactosidase sp act</b>		
			Without acetosyringone	With acetosyringone	
VIK36	pVIK269	$virD5$ -lacZ	9	85	
VIK37	pVIK144	$virE3$ -lacZ	14	252	
VIK38	pVIK140	ORF12-lacZ	19	71	
<b>VIK39</b>	pVIK245	$ORF16$ -lac $Z$			
VIK40	pVIK268	$virP$ -lac $Z$	4	84	
VIK41	pVIK267	$virR$ -lacZ	2	632	

*<sup>a</sup>* Cells containing the indicated *lacZ* fusions were cultured at pH 5.5 with or without 100  $\mu$ M acetosyringone for 12 h and tested for  $\beta$ -galactosidase specific activity (19).

is difficult to imagine what role such an enzyme might have in these events. We hypothesize that one or both of these genes may direct a process unrelated to T-DNA transfer. It is tempting to speculate that a phosphatase might be useful in the dephosphorylation of isopentenyl-AMP, which is synthesized by the product of the *ipt* gene (located in the T-DNA). Although *ipt* is normally thought of as being expressed only after transfer to plant cells, it was recently shown to be expressed in *A. tumefaciens* as well (6). Isopentenyl-AMP might be expected to be membrane impermeable due to its negative charge, while a phosphatase might increase membrane permeability, thereby releasing isopentenyl-adenosine from the bacteria. Several other strains of *A. tumefaciens* are known to release cytokinins in response to *vir* gene-inducing stimuli (1, 2, 20, 21).

We thank John Helmann, Valley Stewart, and the members of our laboratory for helpful discussions.

This work was funded by a Public Health Service research grant from the National Institutes of Health (GM42893) and by the Cornell University College of Agriculture and Life Sciences.

## **REFERENCES**

- 1. **Akiyoshi, D. E., D. A. Regier, and M. P. Gordon.** 1987. Cytokinin production by *Agrobacterium* and *Pseudomonas* spp. J. Bacteriol. **169:**4242–4248.
- 2. **Akiyoshi, D. E., D. A. Regier, G. Jen, and M. P. Gordon.** 1985. Cloning and nucleotide sequence of the *tzs* gene from *Agrobacterium tumefaciens* strain T37. Nucleic Acids Res. **13:**2773–2788.
- 3. **Alt-Mo¨rbe, J., J. L. Stryker, C. Fuqua, S. K. Farrand, and S. C. Winans.** 1996. The conjugal transfer system of *Agrobacterium tumefaciens* octopinetype 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.
- 4. **Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman.** 1990. Basic local alignment search tool. J. Mol. Biol. **215:**403–410.
- 5. **Beier, D., T. M. Fuchs, H. Graeff-Wohlleben, and R. Gross.** 1996. Signal transduction and virulence regulation in *Bordetella pertussis*. Microbiologia **12:**185–196.
- 6. **Chou, A., J. Archdeacon, and C. I. Kado.** 1998. *Agrobacterium* transcriptional

regulator Ros is a prokaryotic zinc finger protein that regulates the plant oncogene *ipt*. Proc. Natl. Acad. Sci. USA **95:**5293–5298.

- 7. **Collmer, A.** 1998. Determinants of pathogenicity and avirulence in plant pathogenic bacteria. Curr. Opin. Plant Biol. **1:**329–335.
- 8. **Cornelis, G. R., and H. Wolf-Watz.** 1997. The Yersinia Yop virulon: a bacterial system for subverting eukaryotic cells. Mol. Microbiol. **23:**861–867.
- 9. **Farrand, S. K., I. Hwang, and D. M. Cook.** 1996. The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. J. Bacteriol. **178:**4233–4247.
- 10. **Greene, J. M., and K. Struhl.** 1993. Analysis of RNA structure and synthesis, p. 4.6.1. *In* Current Protocols in Molecular Biology. Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds). New York: John Wiley & Sons.
- 11. **Heath, J. D., T. C. Charles, and E. W. Nester.** 1995. Ti plasmid and chromosomally encoded two-component systems important in plant cell transformation by *Agrobacterium* species, p. 367–385. *In* J. A. Hoch and T. J. Silhavy (ed.), Two-component signal transduction. ASM Press, Washington, D.C.
- 12. **Hirooka, T., P. M. Rogowsky, and C. I. Kado.** 1987. Characterization of the *virE* locus of *Agrobacterium tumefaciens* plasmid pTiC58. J. Bacteriol. **169:** 1529–1536.
- 13. **Kalogeraki, V. S., and S. C. Winans.** 1995. The octopine-type Ti plasmid pTiA6 of *Agrobacterium tumefaciens* contains a gene homologous to the chromosomal virulence gene *acvB*. J. Bacteriol. **177:**892–897.
- 14. **Kalogeraki, V. S., and S. C. Winans.** 1997. Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. Gene **188:**69–75.
- 15. **Kalogeraki, V. S., and S. C. Winans.** 1998. Wound-released chemical signals may elicit multiple responses from an *Agrobacterium tumefaciens* strain containing an octopine-type Ti plasmid. J. Bacteriol. **180:**5660–5667.
- 16. **Kalogeraki, V. S., J. Zhu, A. Eberhard, E. L. Madsen, and S. C. Winans.** 1999. The phenolic *vir* pene inducer ferulic acid is O-demethylated by the VirH2 protein of an *A. tumefaciens* Ti plasmid. Mol. Microbiol. **34:**512–522.
- 17. **Lin, T. S., and C. I. Kado.** 1993. The *virD4* gene is required for virulence while *virD3* and orf5 are not required for virulence of *Agrobacterium tumefaciens*. Mol. Microbiol. **9:**803–812.
- 18. **Mantis, N. J., and S. C. Winans.** 1993. The chromosomal response regulatory gene *chvI* of *Agrobacterium tumefaciens* complements an *Escherichia coli phoB* mutation and is required for virulence. J. Bacteriol. **175:**6626–6636.
- 19. **Miller, J. H.** 1972. Experiments in molecular genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- 20. **Powell, G. K., N. G. Hommes, J. Kuo, L. A. Castle, and R. O. Morris.** 1988. Inducible expression of cytokinin biosynthesis in *Agrobacterium tumefaciens* by plant phenolics. Mol. Plant Microbe Interact. **1:**235–242.
- 21. **Regier, D. A., D. E. Akiyoshi, and M. P. Gordon.** 1989. Nucleotide sequence of the *tzs* gene from *Agrobacterium rhizogenes* strain A4. Nucleic Acids Res. **17:**8885.
- 22. **Rodriguez-Palenzuela, P., T. J. Burr, and A. Collmer.** 1991. Polygalacturonase is a virulence factor in *Agrobacterium tumefaciens* biovar 3. J. Bacteriol. **173:**6547–6552.
- 23. **Sheng, J., and V. Citovsky.** 1996. *Agrobacterium*-plant cell DNA transport: have virulence proteins, will travel. Plant Cell **8:**1699–1710.
- 24. **Stachel, S. E., and E. W. Nester.** 1986. The genetic and transcriptional organization of the *vir* region of the A6 Ti plasmid of *Agrobacterium tumefaciens*. EMBO J. **5:**1445–1454.
- 25. **Winans, S. C.** 1992. Two-way chemical signalling in *Agrobacterium*-plant interactions. Microbiol. Rev. **56:**12–31.
- 26. **Zhu, J., and S. C. Winans.** 1998. Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated dominant defective TraR-like protein. Mol. Microbiol. **27:**289–297.