DNA RESEARCH 5, 87–93 (1998)

Structural Characterization of the *virB* Operon on the Hairy-root-inducing Plasmid A4

Yajie LIANG, Takashi AOYAMA, and Atsuhiro OKA*

Laboratory of Molecular Biology, Institute for Chemical Research, Kyoto University, Uji, Kyoto 611-0011, Japan

(Received 20 March 1998)

Abstract

The hairy-root-inducing plasmid A4 (pRiA4) is capable of conferring tumorigenic symptoms on plants upon infection by its host bacterium, *Agrobacterium rhizogenes*. The *virB* operon on pRiA4 has been sequenced and found to be composed of 11 genes, *virB1* to *virB11*, whose products mostly appear to be associated with the cell membrane. A novel structural characteristic is frequent overlappings between the translation termination and initiation codons of adjacent genes. This is indicative of fine tuning of relative translation frequencies for each VirB protein. These results support the view that VirB multisubunit complexes provide facilities for T-DNA transfer at the bacterial cell membrane. The structural organization of the pRiA4 *virB* operon was essentially identical to that of the previously reported *virB* operons of tumor-inducing plasmids, pTiC58 and pTiA6, and the corresponding VirB proteins of the three plasmids were extremely homologous to one another. On the basis of the structural similarity of each VirB protein, the phylogenetic relationship among pRiA4, pTiC58, and pTiA6 is discussed.

Key words: *Agrobacterium rhizogenes*; Ri plasmid; *vir* gene; T-DNA transfer; pilus

1. Introduction

Agrobacterium rhizogenes is a plant pathogen that induces hairy roots on a wide variety of dicotyledonous plants. This symptom is caused by the hairy-root-inducing plasmid (pRi) that is similar to tumor-inducing plasmids (pTi) carried by *A. tumefaciens*, though pRi and pTi belong to different incompatibility groups. A portion of either plasmid DNA, defined as T-DNA, is transferred to plant cells and subsequently integrated into the plant nuclear genomes. Since T-DNA includes several genes that direct synthesis of phytohormones and/or their release from the glycosylated derivatives, tumorigenesis by pRi and pTi appears to be the results of disturbing the phytohormone balance in plant cells. Although the remaining T-DNA genes vary with plasmid species, T-DNA is always flanked by 24-bp imperfectly repeated sequences with the same polarity. These border repeats work as a *cis*-acting signal for T-DNA transfer (for a review, see ref. 1).

Although tumorigenesis by pRi and pTi needs the T-DNA genes directing the synthesis and release of phytohormones, T-DNA transfer itself occurs without any genes on T-DNA if the border repeats are present. *Trans*-factors required for T-DNA transfer are supplied from the plasmid virulence genes (*vir*), located outside of T-DNA, together with the chromosomal virulence genes (*chv*) scattered on the bacterial genome. The plasmid *vir* genes are usually comprised of six transcriptional units: *virA*, *virB*, *virC*, *virD*, *virE*, and *virG*. In the case of pRiA4, however, *virE* is missing except for its promoter region. The gene organization and transcriptional directions of these *vir* operons are well conserved among pRi and pTi though the spacers between the operons vary with the plasmid.[1,2] The expression of these genes is tightly regulated. Except for *virA* and *virG*, no expression occurs during vegetative growth of bacteria, and every *vir* gene including *virA* and *virG* is inducible by plant phenolic compounds such as acetosyringone, through the *virA* and *virG* gene products (two-component regulatory system).[3,4]

The entire *vir* region of pTiC58 has already been sequenced.[5] We have also deduced the sequence of all *vir* loci except *virB* of pRiA4. To complete structural analyses of pRiA4 *vir* and to shed light on the structural characteristics and functions of VirB proteins, we now report the nucleotide sequences of the *virB* operon, together with its flanking DNA regions.

Communicated by Mituru Takanami

* To whom correspondence should be addressed. Tel. +81-774-38-3262, Fax. +81-774-38-3259, E-mail: oka@molbio.kuicr.kyoto-u.ac.jp

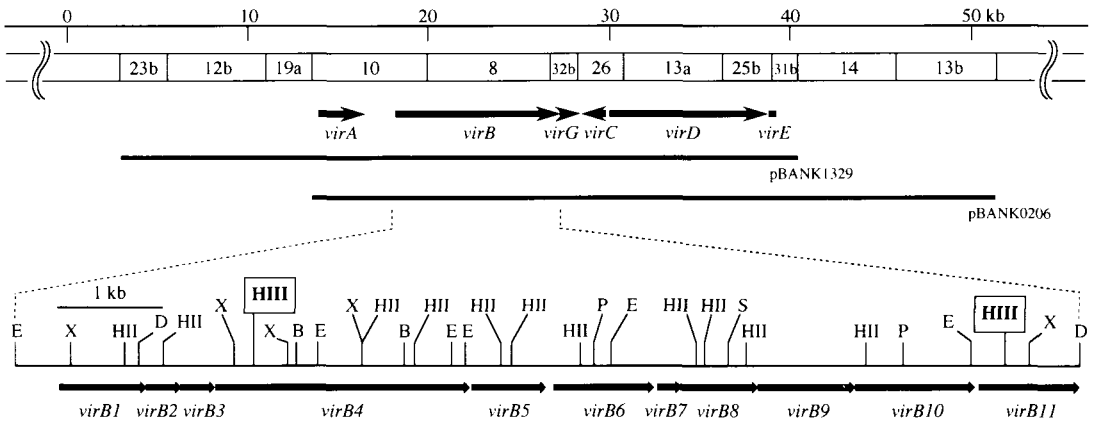


Figure 1. Organization of the pRiA4 *vir* genes. The bar at the top is a portion of pRiA4 DNA with the *Hin*dIII cleavage map, and the numerals on the bar show the *Hin*dIII restriction fragments. Each of the *virA* to *virG* operons[6] is indicated by a heavy line with an arrowhead in the 5′-to-3′ direction. A tiny square marked with *virE* is a promoter region of *virE*. The DNA regions carried by the two cosmid clones are shown by a thick line.[10] The lower half illustrates representative restriction sites in the *virB* region (B, *Bam*HI; D, *Dra*I; E, *Eco*RI; HII, *Hin*cII; HIII, *Hin*dIII; P, *Pst*I; S, *Sma*I; X, *Xho*I) together with the coding regions of 11 *virB* genes.

2. Materials and Methods

2.1. Reagents and general procedures

Restriction endonucleases were purchased from Takara Shuzo Co., Ltd. and Toyobo Inc., and used according to the instructions provided by the suppliers. DNA polymerase I Klenow fragment, T4 DNA ligase, and Taq polymerase were obtained from Takara Shuzo Co., Ltd. The procedures used for standard recombinant DNA experiments were as described previously.[6,7]

2.2. Bacterial strains, plasmids, and cosmids

The bacterial strains used were *A. rhizogenes* A4[6] and *Escherichia coli* K-12 DH5α [*supE44* Δ*lacU169* (Φ80 *lacZΔM15*) *hsdR17 recA1 endA1 gyrA96 thi-1 relA1*].[7] The plasmid vectors used were pUC18/19[8] and pHSG398/399.[9] For the DNA source of pRiA4 *virB* region, we used the two cosmid clones pBANK1329 and pBANK0206, which carry the *Hin*dIII-23b, -12b, -19a, -10, -8, -32b, -26, -13a, -25b, and -31b fragments of pRiA4 and the *Hin*dIII-10, -8, -32b, -26, -13a, -25b, -31b, -14, and -13b fragments of pRiA4, respectively (see Fig. 1).[10]

2.3. Nucleotide sequence analyses

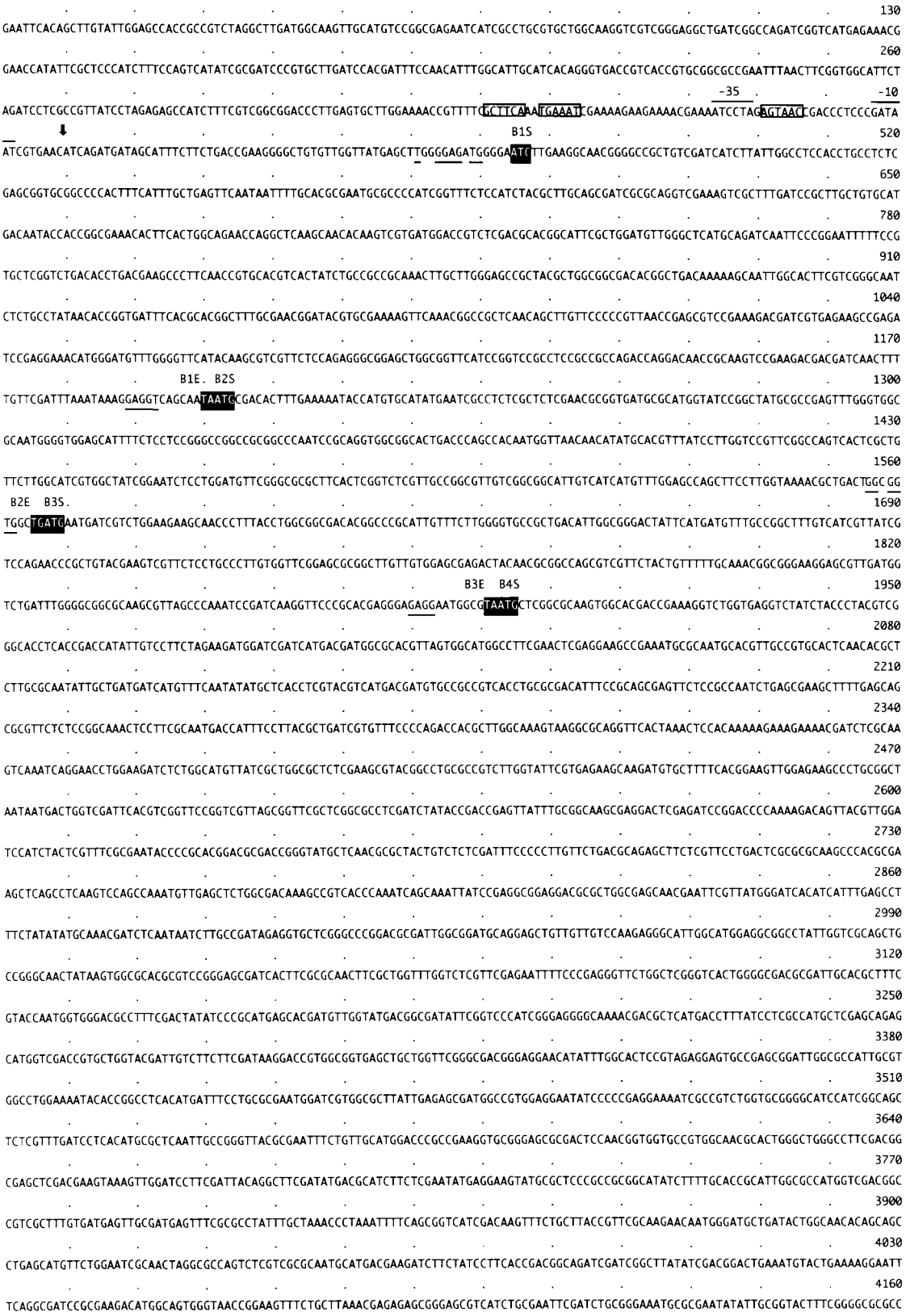
Various pRiA4 DNA fragments inserted in pUC18/19 or pHSG398/399 were sequenced with an ABI Prism 377 DNA Sequencer and a Dye Terminator Cycle Sequencing Kit with AmpliTaq DNA polymerase, FS (PE Applied Biosystems; Part no. 402079). Oligonucleotide primers used for sequencing by polymerase chain reactions were two universal forward and reverse primers (Takara Shuzo Co., Ltd.), corresponding to the DNA regions close to the poly-cloning sites on the plasmid vectors, and various custom primers synthesized by Biologica Co. The reaction products were separated by electrophoresis on a 5% polyacrylamide gel made of Super Reading DNA Sequence Solution (Toyobo Inc.). Nucleotide sequences thus obtained were arranged and analyzed using the software DNASIS-Mac (ver. 3.5, Hitachi Software Engineering Co., Ltd.) and various DNA-analyzing programs supplied by the GenomeNet WWW Server (http://www.genome.ad.jp/). The SOSUI,[11] PSORT,[12] MOTIF[13] computer programs were used for predicting transmembrane domains of proteins, protein localization sites in cells/sorting signals, and various amino acid sequence motifs, respectively.

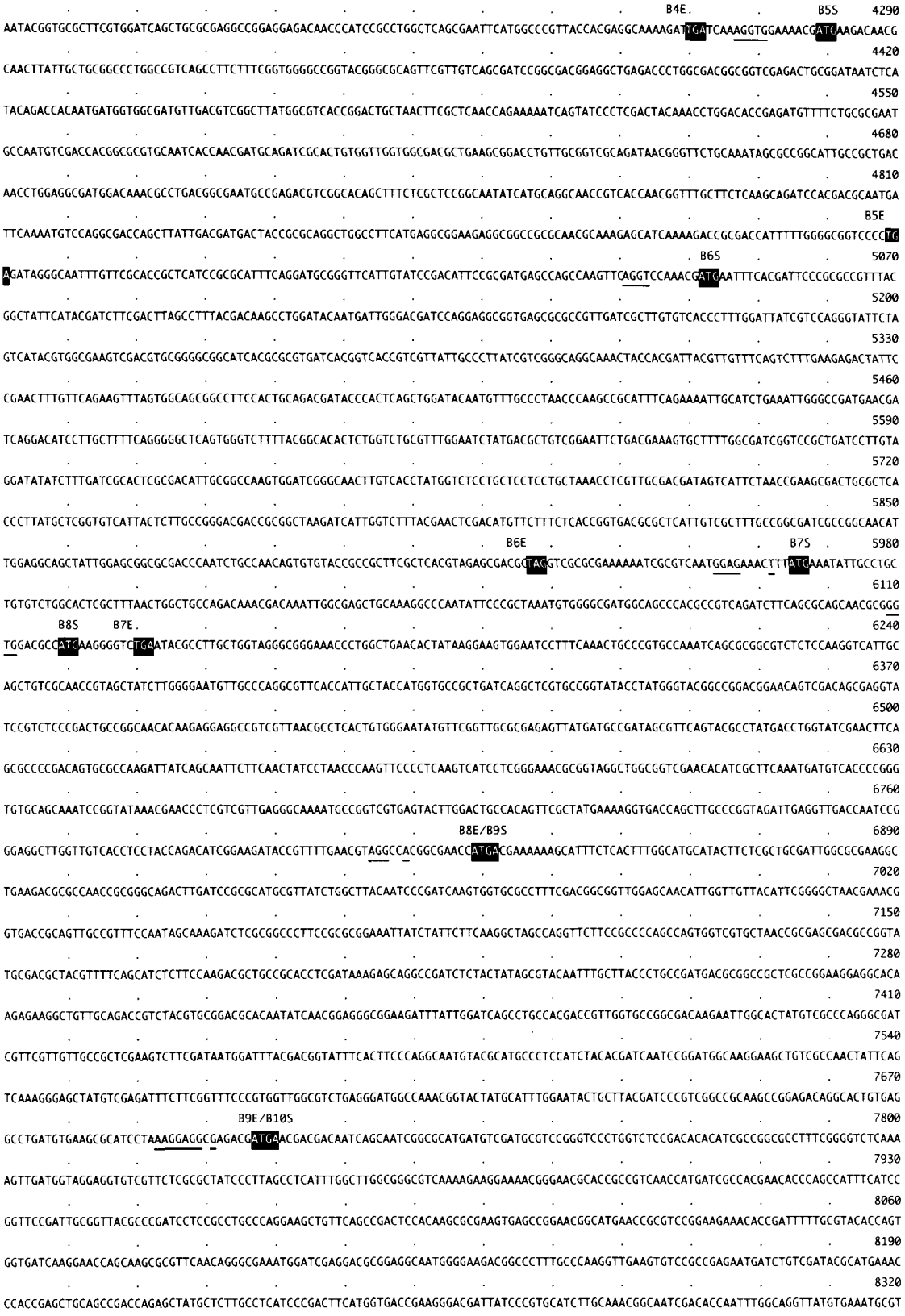
3. Results and Discussion

3.1. Sequencing strategy

It has been shown that pRiA4 *virB* is located in a region corresponding to the *Hin*dIII-10, -8 and -32b fragments from transposon-insertion mutagenesis and similarity tests to pTiA6.[6] These three fragments were purified from a cosmid clone (pBANK1329) (Fig. 1), and

Figure 2. Nucleotide sequence of the pRiA4 *virB* operon and its flanking regions in the 5′-to-3′ direction. The first and last nucleotide positions correspond to nucleotide position 4236 in Fig. 2 of ref. 15 and to nucleotide position 6 in Fig. 1 of ref. 16. The start sites of the *virB1* to *virB11* coding regions are indicated by B1S to B11S, respectively, and the corresponding ending sites by B1E to B11E, respectively, with outline letters indicating their translation initiation and termination codons. Potential ribosome-binding sequences are underlined. The transcription start site is marked by a vertical arrow, and the promoter −35/−10 regions are shown by upperlines. Helically phased hexamer sequences with the upstream inverted one for VirG binding are boxed.





then separately inserted into the plasmid vectors. With the resulting recombinant plasmids, progressive deletion derivatives were constructed by either ejecting appropriate restriction fragments or recloning smaller fragments. Each of the inserted DNA segments was sequenced from either side of both strands. In addition, adjacent restriction fragments were confirmed to be contiguous by sequencing overlapping fragments or pBANK0206 cosmid DNA with appropriate custom primers. By arranging overlapping sequences of both strands, the nucleotide sequence from the *Eco*RI site downstream of *virA* to the *Dra*I site upstream of *virG* (9,926 bp) was deduced, as shown in Fig. 2. The sequence data were deposited in the DDBJ/EMBL/GenBank data banks with accession numbers AB011800 and AB011801.

3.2. Characterization of the virB *nucleotide sequence*

The occurrence of initiation and termination codons for protein synthesis in all possible reading frames suggested that the *virB* locus contains information to code for 11 polypeptides of 26.4 kDa, 12.3 kDa, 11.7 kDa, 87.7 kDa, 23.1 kDa, 31.8 kDa, 5.9 kDa, 26.2 kDa, 32.2 kDa, 40.3 kDa, and 38.2 kDa in the same polarity (nucleotide positions 465 to 9905 in Fig. 2). Each putative coding frame was named *virB1* to *virB11*, respectively, all of which were accompanied by a potential ribosome-binding sequence,[14] as marked by underlines in Fig. 2. The *virB1* gene starts 1747 bp downstream of the *virA* termination codon,[15] and the *virB11* gene ends 255 bp upstream of the *virG* initiation codon.[16] The putative initiation codon for each *virB* gene except *virB1* was the conventional AUG, whereas *virB1* appeared to use AUC, a rare translation initiation codon. This is not the result of a base substitution occurred during cloning experiments because direct sequencing of the corresponding region of pRiA4 DNA gave an identical result. The transcription start site previously found in a *virB* region was the only one that corresponds to 65 bp upstream of the *virB1* translation initiation site.[3] In this DNA region, there were the −10 region of the promoter (−10 in Fig. 2) and the helically phased hexamer sequences for VirG binding (marked by boxes in Fig. 2).[3] Therefore, all of *virB1* to *virB11* are likely to be actually structural genes and to constitute an operon. A novel structural characteristic of the *virB* operon is frequent overlappings of the translation termination codon of a given gene with the initiation codon of the following gene. Those from *virB1* to *virB4* overlap for 1 bp such that two adjacent coding regions are separated by 2 bp. Those from *virB8* to *virB10* overlap for 2 bp where the adjacent coding regions themselves share 1 bp. Furthermore, in the cases of *virB7* and *virB8*, 11 bp of the coding regions overlap with each other. As a consequence, their potential ribosome-binding sequences should invade the preceding genes. This kind of overlapping and compactness have never been found in the other *vir* operons. Conversely, short spacer sequences are present between the two genes from *virB4* to *virB7*.

3.3. Characterization of VirB proteins

To obtain information on characteristics of the *virB* gene products, the amino acid sequence of each VirB protein was analyzed. Possible N-terminal signal peptides were found in VirB1, VirB5, VirB7, and VirB9, and

Table 1. Similarity of the VirB proteins among the three plasmids, pRiA4, pTiC58, and pTiA6.[a]

Comparison	VirB1	VirB2	VirB3	VirB4	VirB5	VirB6	VirB7	VirB8	VirB9	VirB10	VirB11
pRiA4/pTiC58	93	95	94	93	90	94	94	91	93	92	93
pRiA4/pTiA6	70	86	88	90	85	89	80	88	85	83	92
pTiC58/pTiA6	71	89	87	89	87	89	82	86	85	83	89

[a]Similarity of each VirB protein between two plasmids shown in the leftmost column is expressed as the percent identity of amino acid residues.

potential transmembrane domains were found in VirB2, VirB3, VirB5, VirB6, VirB8, VirB9, and VirB10. The smallest VirB7 protein contained a putative target sequence for signal peptidase II, an enzyme involved in lipoprotein recognition and processing,[17] and thus it appears to associate with lipid. In addition, an ATP-binding motif was found in VirB4 and VirB11.

3.4. VirB multisubunit complexes

As mentioned above, the pRiA4 *virB* operon is highly compacted. These traits are identical with those of pTiC58 and close to those of pTiA6.[18–20] Conservation of these traits seems to reflect the control of translation initiation frequency of a given gene relative to that of the preceding gene, providing the exact molar ratios of synthesized VirB proteins. Another possibility is that translation of a gene depends on the preceding gene product. Either view is compatible with the idea that VirB proteins form multisubunit complexes, probably coupled with their translation processes. In fact, several protein complexes have been reported with pTi (e.g., VirB7-VirB9 heterodimer and VirB10 homodimer),[21–24] and the presence of some VirB proteins appears to stabilize other VirB proteins.[25,26] ATP-binding proteins, VirB4 and VirB11, might act as energy suppliers. Furthermore, VirB2, VirB3, and VirB4 are significantly homologous to TraA, TraL, and TraC, respectively, which are essential for the assembly of the pilus of F plasmid.[27] These findings are all consistent with the view that the VirB multisubunit complexes provide facilities at the bacterial cell membrane for the passage of T-DNA from a bacterium to a plant cell, as does the F pilus for conjugal transfer of F plasmid.

3.5. Evolution of pRiA4, pTiC58, and pTiA6

The percent identity of amino acid residues between VirB of pRiA4 and pTiC58 and between VirB of pRiA4 and pTiA6 were in the ranges 90%-95% and 70%-92%, respectively (Table 1). These similarity scores are roughly comparable to those for other *vir* operons. However, for any *vir* operons and their spacer regions, pRiA4 is always much closer to pTiC58 than to pTiA6. In addition, these similarity scores between pRiA4 and pTiC58 (compatible combination) are significantly higher than those between pTiA6 and pTiC58 (incompatible combination), whereas the similarity scores between pTiA6 and either pRiA4 or pTiC58 are comparable. These structural features suggest that the *vir* regions of the three plasmids have evolved from a common ancestor of the *vir* gene set from which pTiA6 *vir* was initially separated. Since plasmids belonging to the same incompatibility group are usually more cognate to each other than those involved in different incompatibility groups, functional domains of pRi and pTi, at least those determining virulence on plants (*vir* and T-DNA) and autonomous replication and incompatibility (*ori*/*inc*), have presumably been shuffled during the evolution of these plasmids.

Acknowledgments: We thank the Supercomputer Laboratory of our Institute for providing computational time. L. Y. was supported by a scholarship for foreign graduates from the Ministry of Education, Science, Sports, and Culture of Japan.

References

1. Hooykaas, P. J. and Schilperoort, R. A. 1992, *Agrobacterium* and plant genetic engineering, *Plant Mol. Biol.*, **19**, 15–38.
2. Oka A. 1993, Transcriptional control of the *Agrobacterium* virulence genes by two proteins VirA and VirG, *Bull. Inst. Chem. Res., Kyoto Univ.*, **71**, 355–368.
3. Aoyama, T., Takanami, M., and Oka, A. 1989, Signal structure for transcriptional activation in the upstream regions of virulence genes on the hairy-root-inducing plasmid A4, *Nucleic Acids Res.*, **17**, 8711–8725.
4. Stachel, S. E. and Zambryski, P. C. 1986, VirA and VirG control the plant-induced activation of the T-DNA transfer process of *A. tumefaciens*, *Cell*, **46**, 325–333.
5. Rogowsky, P. M., Powell, B. S., Shirasu, K., Lin, T.-S., Morel, P., Zyprian, E. M., Steck, T. R., and Kado, C. I. 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.
6. Hirayama, T., Muranaka, T., Ohkawa, H., and Oka, A. 1988, Organization and characterization of the *virCD* genes from *Agrobacterium rhizogenes*, *Mol. Gen. Genet.*, **213**, 229–237.
7. Sambrook, J., Fritsch, E. F., and Maniatis, T. 1989, Molecular cloning: a laboratory manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
8. Yanisch-Perron, C., Vieira, J., and Messing, J. 1985, Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors, *Gene*, **33**, 103–119.

9. Takeshita, S., Sato, M., Toba, M., Masahashi, W., and Hashimoto-Gotoh, T. 1987, High-copy-number and low-copy-number plasmid vectors for *lacZα*-complememtation and chloramphenicol- or kanamycin-resistance selection, *Gene*, **61**, 63–74.
10. Nishiguchi, R. and Oka, A. 1986, Structure of the hairy-root-inducing plasmid and identification of its replicator region, *Bull. Inst. Chem. Res., Kyoto Univ.*, **64**, 79–87.
11. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. 1998, SOSUI: Classification and secondary structure prediction system for membrane proteins, *Bioinformatics*, in press.
12. Nakai, K. and Kanehisa, M. 1991, Expert system for predicting protein localization sites in gram-negative bacteria, *Proteins: Structure, Function, and Genetics*, **11**, 95–110.
13. Bairoch, A., Bucher, P., and Hofmann, K. 1997, The PROSITE database, its status in 1997, *Nucleic Acids Res.*, **25**, 217–221.
14. Shine, J. and Dalgarno, L. 1974, The 3′-terminal sequence of *Escherichia coli* 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites, *Proc. Natl. Acad. Sci. USA*, **71**, 1342–1346.
15. Endoh, H., Hirayama, T., Aoyama, T., and Oka, A. 1990, Characterization of the *virA* gene of the agropine-type plasmid pRiA4 of *Agrobacterium rhizogenes*, *FEBS Lett.*, **271**, 28–32.
16. Aoyama, T., Hirayama, T., Tamamoto, S., and Oka, A. 1989, Putative start codon TTG for the regulatory protein VirG of the hairy-root-inducing plasmid pRiA4, *Gene*, **78**, 173–178.
17. Yu, F., Inouye, S., and Inouye, M. 1986, Lipoprotein-28, a cytoplasmic membrane lipoprotein from *Escherichia coli*. Cloning, DNA sequence, and expression of its gene, *J. Biol. Chem.*, **261**, 2284–2288.
18. Kuldau, G. A., De Vos, G., Owen, J., McCaffrey, G., and Zambryski, P. 1990, The *virB* operon of *Agrobacterium tumefaciens* pTiC58 encodes 11 open reading frames, *Mol. Gen. Genet.*, **221**, 256–266.
19. Shirasu, K., Morel, P., and Kado, C. I. 1990, Characterization of the *virB* operon of an *Agrobacterium tumefaciens* Ti plasmid: nucleotide sequence and protein analysis, *Mol. Microbiol.*, **4**, 1153–1163.
20. Ward, J. E., Akiyoshi, D. E., Regier, D., Datta, A., Gordon, M. P., and Nester, E. W. 1988, Characterization of the *virB* operon from an *Agrobacterium tumefaciens* Ti plasmid, *J. Biol. Chem.*, **263**, 5804–5814.
21. Anderson, L. B., Hertzel, A. V., and Das, A. 1996, *Agrobacterium tumefaciens* VirB7 and VirB9 form a disulfide-linked protein complex, *Proc. Natl. Acad. Sci. USA*, **93**, 8889–8894.
22. Baron, C., Thorstenson, Y. R., and Zambryski, P. C. 1997, The lipoprotein VirB7 interacts with VirB9 in the membranes of *Agrobacterium tumefaciens*, *J. Bacteriol.*, **179**, 1211–1218.
23. Das, A., Anderson, L. B., and Xie, Y. H. 1997, Delineation of the interaction domains of *Agrobacterium tumefaciens* VirB7 and VirB9 by use of the yeast two-hybrid assay, *J. Bacteriol.*, **179**, 3404–3409.
24. Beaupre, C. E., Bohne, J., Dale, E. M., and Binns, A. N. 1997, Interactions between VirB9 and VirB10 membrane proteins involved in movement of DNA from *Agrobacterium tumefaciens* into plant cells, *J. Bacteriol.*, **179**, 78–89.
25. Fernandez, D., Spudich, G. M., Zhou, X. R., and Christie, P. J. 1996, The *Agrobacterium tumefaciens* VirB7 lipoprotein is required for stabilization of VirB proteins during assembly of the T-complex transport apparatus, *J. Bacteriol.*, **178**, 3168–3176.
26. Jones, A. L., Shirasu, K., and Kado, C. I. 1994, The product of the *virB4* gene of *Agrobacterium tumefaciens* promotes accumulation of VirB3 protein, *J. Bacteriol.*, **176**, 5255–5261.
27. Kado, C. I. 1994, Promiscuous DNA transfer system of *Agrobacterium tumefaciens*: role of the *virB* operon in sex pilus assembly and synthesis, *Mol. Microbiol.*, **12**, 17–22.