IS292: a novel insertion element from *Agrobacterium*

Cécile Ponsonnet,¹ Philippe Normand,¹ Gilles Pilate² and Xavier Nesme^{1,3}

Author for correspondence: Cécile Ponsonnet. Fax: +33 72 43 12 23.
e-mail: nesmc@ecosol.univ-lyon1.fr

¹ Laboratoire d'Ecologie Microbienne du Sol, Centre National de la Recherche Scientifique, URA 1977, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, F-69622 Villeurbanne, France

² Station d'Amélioration des Arbres Forestiers, Institut National de la Recherche Agronomique, F-45160 Ardon, France

³ Institut National de la Recherche Agronomique

A new insertion sequence, IS292, located in the 6b gene of a nopaline-type *Agrobacterium* strain (X88-292) isolated from poplar was identified and sequenced. IS292 is 2494 bp long, has 21 bp inverted terminal repeats with two mismatches, and generates 10 bp direct repeats upon integration. No sequence similarity was found between IS292 and other insertion elements associated with *Agrobacterium*, but it shows strong similarity with ISR11 from *Rhizobium leguminosarum* bv. *viciae*. The occurrence of IS292-like sequences in various *Agrobacterium* isolates, especially different *Agrobacterium* strains isolated from the same biotope, was demonstrated by DNA hybridization.

Keywords: *Agrobacterium*, insertion sequence

INTRODUCTION

Insertion sequences (ISs) are transposable genetic elements which are commonly found in the genomes of bacteria. They can coexist in the cell in multiple copies, located on the chromosome, plasmids or phages (see reviews by Galas & Chindler, 1989; Grindley & Reed, 1985). Many reports have described the presence of insertion sequences in the genome of *Rhizobiaceae* (see reviews by Flores *et al.*, 1987; Martinez *et al.*, 1990; Otten *et al.*, 1992). Several insertion elements have also been found in various *Agrobacterium* Ti plasmids (Bonnard *et al.*, 1989; De Meirsmann *et al.*, 1989, 1990; Machida *et al.*, 1984; Vanderleyden *et al.*, 1986). Systematic surveys of the distribution of insertion elements in various *Agrobacterium* strains isolated from grapevine (Bonnard *et al.*, 1989; Paulus *et al.*, 1989a, b, 1991a, b) have permitted a better understanding of evolutionary links between the different octopine/cucumopine-type strains of *Agrobacterium vitis*. In this case, some of the observed Ti plasmid variability was due to the activity of insertion elements leading to rearrangements involving transpositions and deletions (Otten *et al.*, 1992; Paulus *et al.*, 1991b).

Abbreviations: PCR-RFLP, polymerase chain reaction-restriction fragment length polymorphism; T-DNA, transferred DNA.

The GenBank accession number for the nucleotide sequence reported in this paper is L29283.

In a previous study of *Agrobacterium* populations isolated from poplar tumours, we observed the occurrence of a particular group of isolates identified as chromosome type 2177 by the PCR-RFLP method (Ponsonnet & Nesme, 1994). We later found that these isolates could harbour one of the two types of Ti plasmid, pTi2177 or pTi292, differing in the size of the T-DNA region between the *tmr* and *nos* genes. This observation led us to hypothesize the presence of an insertion element in the T-DNA of pTi292. Here, we report on the identification and characterization of a new insertion element, IS292, from the nopaline-type *Agrobacterium* isolate X88-292. IS292 is distinct from any of the previously described *Agrobacterium*-associated insertion sequences, but shows similarity to ISR11 associated with *Rhizobium leguminosarum* bv. *viciae*. The implications of the occurrence of IS292 in a natural *Agrobacterium* population are discussed.

METHODS

Bacterial strains. The characteristics of the bacterial strains used in this study are summarized in Table 1. The *Escherichia coli* recipient used for plasmid transformation and as a host for M13 phage cloning was strain DH5αF' (Bethesda Research Laboratories). M13mp18 (Yanisch-Perron *et al.*, 1985) was obtained from BRL and pBluescript-SK from Stratagene. *Agrobacterium* strains were grown on YPGA medium (yeast extract, 5 g l⁻¹; Bacto-Peptone, 5 g l⁻¹; glucose, 10 g l⁻¹; Bacto-Agar, 15 g l⁻¹; all from Difco Laboratories) for 48 h at 28 °C. *E. coli* was grown on LB medium (yeast extract, 5 g l⁻¹; Bacto-Tryptone, 10 g l⁻¹;

Table 1. Bacterial strains used in this study

Bacterium*	Opine type	Host plant	Origin	Source†
<i>Agrobacterium tumefaciens</i>				
C58	Nopaline	<i>Prunus cerasus</i>	New York	CFBP
T37	Nopaline	Walnut	California	CFBP
1904	Nopaline	<i>Vitis vinifera</i>	Greece	CFBP
2411	Nopaline	<i>Salix purpurea</i>	New Zealand	CFBP
CG1028	Nopaline	<i>Prunus</i> rootstock cv. 'Damil'	USA	T. Burr
Ach5	Octopine	<i>Prunus</i> sp.	USA	CFBP
B6	Octopine/agropine	<i>Lycopersicon esculentum</i>	USA	CFBP
Bo542	Agropine/mannopine	<i>Dahlia</i> sp.	Germany	CFBP
<i>Agrobacterium rhizogenes</i>				
2178 (82-139)	Nopaline	<i>Prunus</i> rootstock cv. 'F12-I'	Orléans	CFBP
1905	Nopaline	<i>Vitis vinifera</i>	Greece	CFBP
1936	Nopaline	<i>Rosa</i> sp.	Tahiti	CFBP
1961	Nopaline	<i>Populus bolleana</i>	France	CFBP
1962	Nopaline	<i>Prunus</i> rootstock 'SL64'	France	CFBP
2417	Nopaline	<i>Prunus</i> rootstock 'Colt'	France	CFBP
CG1029	Nopaline	<i>Prunus</i> rootstock 'Damil'	USA	T. Burr
<i>Agrobacterium vitis</i>				
2512	Unknown	<i>Vitis vinifera</i>	Spain	CFBP
2620	Nopaline	<i>Vitis vinifera</i>	Australia	CFBP
<i>Agrobacterium tumefaciens</i>				
X88-292	Nopaline	<i>Populus tremula</i> × <i>P. alba</i>	Orléans and Peyrat-le-Château	This study
2177				
2517				
2516				
2518				
85-3				
85-46				
A68-18				
B100-11				
B89-11				
<i>Agrobacterium rhizogenes</i>				
85-30	Nopaline	<i>Populus tremula</i> × <i>P. alba</i>	Orléans and Peyrat-le-Château	This study
2520				
85-110				
85-178				
M62				
C104-12				
A134-34				
<i>Agrobacterium tumefaciens</i>				
C91-1	Non-tumorigenic	Soil	Peyrat-le-Château	This study
C91-2				
C91-3				
C91-4				
C91-5				
C91-12				
C91-13				
C91-15				
C91-16				
C91-54				
Bacterium	Characteristics			Reference
<i>Escherichia coli</i> DH5αF'	F80d <i>lacZ</i> ΔM15 Δ(<i>lacZ</i> Y _A – <i>argF</i>)U169 <i>deoR recA1 endA1</i> <i>hsdR17</i> (rK [–] mK [–]) <i>supE44</i> , l [–] <i>thi-1 gyrA relA1</i>			Miller (1972)

* The names *A. tumefaciens* and *A. rhizogenes* are used for *Agrobacterium* biovars 1 and 2, respectively; see Sawada *et al.* (1993) and Bouzar (1994).

† CFBP, Collection Française de Bactéries Phytopathogènes, INRA, Angers, France; T. Burr, Barton Laboratory, Geneva, NY, USA.

Fig 1. Map of the T-DNA region containing IS292. (a) Mapped region between the *tmr* and *nos* genes from strain X88-292. The thick line represents the 2.5 kb insertion. Amplification primer positions are shown below the map. (b) Extended map of IS292 and the DNA sequencing strategy. Amplification primers flanking the insertion element are indicated below the map. Arrows indicate the orientation and extent of each sequenced fragment. (1), Fragments cloned into *Sal*I (*Hpa*II), *Bam*HI (*Mbo*I), *Hind*III and *Xho*I sites of M13mp18; (2), fragments cloned into *Cla*I, *Hind*III, *Ksp*I, *Pst*I, *Sac*I and *Xho*I sites of pBluescript; (3), fragments obtained with internal primers. Primers used for sequencing are indicated below the map and designated as is1, is2', is3, is6, is7', is10', is11' and is12.

NaCl, 10 g l⁻¹) overnight at 37 °C and recombinants were selected with ampicillin (50 µg ml⁻¹).

DNA isolation. Total DNA isolation from *Agrobacterium* strains was performed as described by Dhaese *et al.* (1979). Plasmid DNA from *E. coli* was prepared by the alkaline-lysis procedure (Maniatis *et al.*, 1982). Single-stranded DNA was obtained from phage preparations according to the manual accompanying the M13 cloning kit (BRL).

Amplification reactions. Amplification of DNA from *Agrobacterium* strains 2177 and X88-292 was performed using the primer sets FGP*tmr*530 and FGP*nos*975 (Ponsonnet & Nesme, 1994), FGP6b489' (5'-CGTCGACTTCTCCTTCCTCA-3') and FGP*nos*975, or FGPis292 (5'-TAAGCGGCAAGCTGAC-3') alone since it annealed to insertion terminal inverted repeats. Amplification conditions were those previously described (Ponsonnet & Nesme, 1994), except in the case of FGPis292 where a 52 °C hybridization temperature was used.

Restriction endonuclease mapping and cloning strategy. Restriction mapping of the T-DNA region between the *tmr* and *nos* genes was performed for strains 2177 and X88-292 with *Hind*III, *Sal*I, *Cla*I and *Rsa*I. DNA fragments from strains 2177 and X88-292, amplified with primers FGP6b489' and FGP*nos*975, were restricted by *Bcl*II, *Cla*I, *Eco*RI, *Hinc*II, *Hind*III, *Pst*I, *Sac*I, *Ksp*I, *Sal*I and *Xho*I (Boehringer Mannheim), singly and in combination. The restriction map was constructed according to restriction patterns supported by the T-DNA map

of the T37 strain obtained by Bevan *et al.* (1983). The 2.5 kb fragment (corresponding to the entire insertion) obtained by amplification of strain X88-292 DNA with primer FGPis292) was partially restricted with *Hpa*II and *Mbo*I, or totally digested with *Hind*III and *Xho*I, and cloned into M13mp18. The following supplementary fragments were cloned in pBluescript-SK: *Sal*I-*Cla*I and *Sal*I-*Ksp*I from strains X88-292 and 2177; *Pst*I-*Xho*I, *Hind*III-*Sac*I, *Sac*I and *Hinc*II from strain X88-292.

DNA sequence analysis. Sequence reactions were conducted according to Sanger *et al.* (1977) using the T7 sequencing kit (Pharmacia). In addition, synthetic 15 nt primers were used to determine the sequence of some areas of the complementary strand. The DNA sequence data were assembled and analysed using the Assembly Lign and Mac Vector programs (IBI-Kodak). GenBank release 80 was scanned for similar sequences using the algorithm of Lipman & Pearson (1985) as implemented by Higgins & Gouy (1987).

DNA labelling and dot-blot hybridization. The probe consisted of the amplified fragment, obtained with primer FGPis292, from pSC292.5 recombinant plasmid DNA containing the entire insertion. The DNA probe was labelled during the amplification reaction with digoxigenin using DIG-UTP (Boehringer Mannheim). DNA solutions (approx. 50 ng per dot) from *Agrobacterium* strains were denatured and loaded on GeneScreen Plus nylon filters (Dupont) using a slot-blot apparatus. High stringency conditions were used with filter

Fig 2. Sequence of IS292 and the duplicated target sequence. (a) Complete nucleotide sequence of IS292. Base positions +1 and 2494 are the first and last nucleotides. The 10 bp target site is indicated in italics at both extremities of the insertion. The 21 bp terminal repeats are boxed. ORFs 1, 2, 3 and 4 are indicated above the sequence; ORFs 5, 6, 7 and 8

washes at 63 °C with 0.01 × SSC/0.1 % (w/v) SDS. Detection of hybrid DNAs was achieved by a chemiluminescent reaction using Lumigen PPD as substrate according to the protocol accompanying the non-radioactive DNA labelling and detection kit (Boehringer Mannheim). Filters were placed in plastic pouches, and photographed using Kodak X-omat film exposed for 30 min.

RESULTS

Nucleotide sequence and structural features of IS292

When analysing the Ti plasmids of *Agrobacterium* strains X88-292 and 2177 we showed that the amplified fragments targeted by primers FGP*tmr*530 and FGP*nos*975 (corresponding to the T-DNA region between the *tmr* and *nos* genes) differed in size. In pTi292, harboured by strain X88-292, this fragment was approximately 6 kb long, that is 2.5 kb more than the amplification products obtained from pTi2177 (data not shown). Restriction maps of the amplified DNA region between the *tmr* and *nos* genes from strains 2177 and X88-292 indicated that a supplementary sequence was inserted in the *6b* gene of strain X88-292 (data not shown).

The terminal borders of the insertion were first sequenced in order to determine their structure and insertion site. The exact site of integration (nt 226 in the *6b* gene) was determined by comparing the nucleotide sequences flanking the insertion for strain X88-292 and the corresponding region in strain 2177. At the target site a 10 bp sequence, 5'-CGACGCAAGG-3', is duplicated at both recombination junctions in strain X88-292. The inserted fragment has 21 bp terminal inverted repeats with two mismatches. Fig. 1 shows the cloning strategy used to determine the entire DNA sequence of the insertion on both strands. The 2494 bp nucleotide sequence of the insertion, flanked by the duplicated insertion site, is given in Fig. 2. The *6b* gene sequence surrounding the insertion in strain X88-292 was identical to that of the corresponding region in strains 2177 (Fig. 2b) and C58 (Brasileiro, 1992).

Using ATG, GTG and TTG as start codons, TAA, TGA and TAG as stop signals, and a 100-amino-acid-length threshold, eight ORFs were identified in the inserted sequence. As shown in Figs 2 and 3(b), ORFs 1 (501 bp), 2 (369 bp), 3 (423 bp) and 4 (573 bp) were identified on the upper strand; ORFs 5 (510 bp), 6 (687 bp), 7 (516 bp) and 8 (345 bp) on the opposite strand. The upstream regions of these ORFs were screened for the occurrence of prokaryotic regulatory consensus sequences. Using the algorithm of Staden (1984) putative transcriptional elements -35 and -10 were identified upstream of each ORF except ORF8. Moreover, ORFs 4, 5 and 8 were preceded by a sequence similar to the ribosome binding site (Shine & Dalgarno, 1974) 6 bp from the initiation

Fig 3. Comparison between IS292 and ISR1/a. (a) Percentage nucleotide sequence similarity between the two insertion elements. (b) Location of ORFs 1–8 identified in IS292. (c) Location of the nine major ORFs identified in ISR1/a. Arrows indicate the orientation and extent of each ORF. Conserved amino acid regions are represented by thick lines. Position of the switch between ORF4 and ORF5 from ISR1/a allowing alignment with ORF5 from IS292 is indicated by an asterisk.

codon (Fig. 2a). Using the algorithm of Fickett (1982) for calculating the probability of coding based on mono-nucleotide frequencies, both strands of regions 19–390 and 403–2494 had a high probability of coding. In addition, DNA sequence analysis showed the presence of two Chi-like sequences at positions 1700 and 2424 (Fig. 2a).

Comparison of IS292 with known sequences

Since no identity between the insertion element detected in pTi292 and other prokaryotic insertion elements from the GenBank database was found, it was considered to be a new insertion sequence and designated IS292. However, it appears strongly related to ISR1/a from *R. leguminosarum* bv. *viciae* (GenBank accession number L19650). These two elements are almost identical in size and show an average of 80 % nucleotide sequence similarity (with domains ranging from 20 % to 90 % similarity along the insertion (Fig. 3a). Moreover, ORFs 2, 5 and 8 (Fig. 3b) show amino acid sequence similarity with part of three ORFs of ISR1/a (Figs 3c and 4). The highest score was between ORF5 and the corresponding region of ISR1/a, which showed 69.5 % similarity when a switch in the ISR1/a coding frame at amino acid coordinate 127 in ORF5 was taken into account (Fig. 4).

Comparison of IS292 with insertion elements described for other *Agrobacterium* strains did not reveal notable similarities. However, when only the inverted repeats were considered, nucleotide sequence similarity was found between IS292 and IS66, IS866, IS1131 from

on the opposite strand are shown below. Start codons (>) and stop signals (Δ) are underlined, putative -35 and -10 consensus sequences are indicated by double lines above (plus strand) or below (minus strand) the sequence. The putative Shine–Dalgarno sequences (at positions 546, 1911 and 2420 from the beginning of the insertion +1) are indicated by stars and the two Chi sequences are located by circles (at positions 1700 and 2424). Restriction sites are shown above the sequence. (b) Nucleotide sequence of the *6b* gene at the insertion site of IS292 in strain 2177. The 10 bp target site is underlined.

ORF5						
	1	11	21	31	41	51
ISRLa	MVAYRFEDSR	ATECVARHLN	GYRGILQVDG	YAAYNKLVK	DGGNNSAILA	GCWSHSRRKF
IS292	-----	-----	-----	-----	-----	-----
ISRLb	-----	-----	-VESCRRMDM	PLITSSSGKM	EAITAPSWPA	VGRTAGESST
	61	71	81	91	101	111
ISRLa	YELHVAKSSK	VATDTVERMA	RLWEIEERVR	AQSPEARVAA	RQEISAAIVR	DLFTLWQATL
	II IIIII	III I IIII	IIIII II	IIIIIIIII	III IIIIII	IIIIII II
IS292	--LHIKSSK	VATETLERMA	TLWEIEDTVR	GQSPEARVAA	RQERSAAIVR	CLFTLWQTTL
	I					
ISRLb	SCMSQARAR-K	SLPTPSSVWR	GYGRSRVCA	PKALKLVSLH	ARRSPQRSSV	TSSLSGRRPC
	121	131	141	151	161	171
ISRLa	PRVSGKSKPR	RGSALCHLAS	RHL-RALPDR	RPHRT--RLQ	HSRASNQTPS	DHE---KK
	IIIIIII		I I	I		
IS292	PRVSGKSNLA	EALRYSIKRR	DVFERFLTDG	RVELDSNIVE	RAIRPQAITR	KNSLFGAGSDG
	III IIIII	I II I	IIIIIIII	I IIIIIIII	IIIIIIIII	IIIIIIIII
ISRLb	RGSLENPNLA	EALRYAISRR	DIFERFLTDG	RIELDSNIVE	RAIRPQAITR	KNSLFGAGSDG
	181	191	201	211	221	231
ISRLa						
IS292	GGRWATIAS	LLQTAKMSV	DPQAWLTQTL	ERLANGWPSS	EIHTLMPWSY	KA
	III IIIII	IIIIIIII	I IIIIIIIII	I IIIII I	II IIII I	I
ISRLb	GGRWTATAT	LLQTAKMNAV	DPQAWLTQTL	EALANGWPNS	EIDALMPWNY	AA
ORF2						
	1	11	21	31	41	51
ISRL	-----VCRSR	--RLDDTILV	PKGILRPAGD	DDTELCHDI	QPLGDVFANQ	DLGLTCMFWK
			I I		I I I	I III I
IS292	MPASMICVGA	AAWITPSSFL	-KAYLGRRVT	MTRNFAATTS	KRSGYVLSQ	DFGLTRMLRQ
	61	71	81	91	101	111
ISRL	LFRFDHDLDA	LKMRRKALAW	PRSAFAVRAT	AAFRDLSLDR	RNAGLDLDFED	KGMLLVIDRW
	II II II	I IIIIIII	II II	I I I II II	II III I	I IIIII
IS292	IFRLDHHLDT	LQMRRKALAW	PRLTFAIFVD	AT-RNLGPDR	RNTGLDLLKD	KSMMLVIN--
	121	131	141	151	161	171
ISRL	MMQAFPSVGQ	IGRDRKPSGS	CVSRSMRSSA	SALRPLDRQP	VSPKPRRGLL	LHSGKNHGFE
			I I	I		
IS292	-----	-----TGR	CK-----	----FL----	-----	-----
	181	191	201	211		
ISRL	SLYIIRELNL	GPRHGYDQSI	FCRAPMLSG	T		
IS292	-----	-----	-----	-		
ORF8						
	1	11	21	31	41	51
ISRL	MSAPSSGTRN	FHMIEAVPER	LEGAPRQFRR	RWSDDFKAQA	VAEAMEPGAS	VSAIAHRIGI
	IIII IIIII	IIIIIIIIII	IIIIIIII	IIIIII I	IIIIIIII I	IIIII I
IS292	MSAPISGTRN	FHMIEAVPER	FEGAPRQFRR	RWSDDFKERA	VAEAMEPGPS	VSAIARRTRH
	61	71	81	91	101	111
ISRL	HPSQLFGWRR	DARDGQRSFS	QDRAGQTGTR	TIGTRAMIEL	VIGDVIIRAD	ADIGEALQOR
			II I I		I II I I	
IS292	YIHHNYS---	-----AGVV	MPRAEQGGSS	G-ACVCRGGF	GAG---LRAG	GDRGGYRRRN
	121					
ISRL	VIRAVRSA					
IS292	YPGGGGCR					

Fig 4. Comparison of IS292 and ISRI1 amino acid sequences. The three ORFs in ISRI1 that extend from positions 655 to 26 and 722 to 207 together, 1054 to 1665, and 2412 to 2029 are aligned with ORFs 5, 2 and 8 in IS292. The switch between ISRI1 ORF4 (ISRI1a) and ORF5 (ISRI1b) is located at coordinate 127 in ISRI1a.

5' TAAGCGGCAAGCTGACGCGT	----- IS292 (2494)	----- ATGGGGTCAGCTTGCGCTTA	3'
IIIIII IIIII		I IIIII	
GTAAAGCGACAAGC	----- ISR11 (2495)	----- GCTTGTCGCTTAC	
IIIIII I I III		I III I IIIII	
GTAAAGCGCAAGCTGAGGCG	----- IS66 (2458)	----- GGCCTTCACCGAGGCTTAC	
II III I I I II I		II II III II	
GTATGCGCGCGCTCCATCCATTGATT	----- IS866 (2716)	----- AATCAATGGGACGGAGAGCGCCATAC	
I IIIII I I		I IIIII I	
GTAAAGCGTCCGG	----- IS1131 (2773)	----- CGGAGCGCTAC	
IIIIII I I II		I I I IIIII	
GTAAAGCGCTCATTTCCATGCGC	----- ISRm2 (2500)	----- GCGCATGGAAATTAGCGCTTAC	
TaaGCG		cGCTtA	

Fig 5. Comparison of the terminal inverted repeats of IS292, ISR11, IS66, IS866, IS1131 and ISRm2. Left (5') and right (3') insertion borders are indicated. Conserved nucleotides are connected by vertical lines. The 6 nt consensus sequence is given on the last line (upper case letters indicate nucleotides identical in all the elements, lower case letters indicate nucleotides conserved in most of the elements). The size of each element is indicated in bp in parentheses.

Fig 6. Hybridization of an IS292 probe to total DNA from *Agrobacterium*. (a) Reference strains. Dots: 1, strain X88-292; 2–5 and 17–18, *A. rhizogenes* strains 2417, 1961, 1962, 1905, 2178, CG1029 and 1936; 6 and 7, *A. vitis* strains 2620 and 2512; 8–15, *A. tumefaciens* strains C58, T37, B6, Ach5, Bo542, CG1028, 1904 and 2411. (b) *Agrobacterium* isolated from poplar tumours. Dots: 1, strain 2177; 2–7, *A. rhizogenes* isolates 2520, 85-110, 85-178, 85-30, C104-12 and M62; 8 and 9, strain X88-292 (50 ng and 5 ng); 10–18, *A. tumefaciens* isolates 2517, 2516, 2518, 85-3, 85-46, A68-18, B100-11, B89-11 and A134-34. (c) *Agrobacterium* isolated from soil. Dots: 1, strain X88-292; 2–11, *A. tumefaciens* isolates C91-1, C91-2, C91-3, C91-4, C91-5, C91-12, C91-13, C91-15, C91-16 and C91-54.

Agrobacterium and ISRm2 from *Rhizobium meliloti* (Fig. 5). These elements have sizes ranging from 2494 to 2773 bp, with terminal repeats differing in size and an 8 bp insertion site, except for IS292 in which it is 10 bp long. Moreover, IS292 can be distinguished by the absence of terminal 5'-G and C-3' residues (Fig. 5). We also searched for DNA similarity to pTi sequences and detected 50% similarity between the first 440 nucleotides of IS292 and the region between T_R- and T_L-DNA of pTi15955 at position

14558–14988 (GenBank accession number X00493). Moreover, the amino acid sequence of ORF1 shows similarity to that of the hypothetical protein 14 (NBRF accession number X00493) corresponding to this same part of pTi15955 and for which no activity is known.

Occurrence of IS292 in various *Agrobacterium* strains

Several *Agrobacterium* strains were screened for the presence of DNA sequences similar to IS292. The detection procedure was based on hybridization of total DNA with an amplified fragment corresponding to the entire IS292 as a probe. Among the 17 *Agrobacterium* reference strains, belonging to *A. tumefaciens*, *A. rhizogenes* or *A. vitis* isolated from various host plants and geographical origins (Table 1), only the DNA from the agropine/mannopine-type *A. tumefaciens* strain Bo542 gave a hybridization signal (Fig. 6a). However, the signal was less intense than with strain X88-292.

In a second step, we investigated whether any sequences similar to IS292 were present in other *Agrobacterium* poplar isolates (Table 1) from the same crown gall outbreak as strains 2177 and X88-292 which developed in Orléans and Peyrat-le-Château nurseries (Nesme *et al.*, 1992). All these isolates had been previously differentiated at the chromosomal and Ti plasmid levels by the PCR-RFLP method (data not shown). Sequences similar to IS292 were detected in two *A. tumefaciens* strains, 2177 and 2517, from poplar tumours (Fig. 6b) which have different Ti plasmid and chromosomal backgrounds. In these cases, the sequences similar to IS292 were not located in the *6b* gene as demonstrated by the size of the T-DNA region between the *tmr* and *nos* genes (data not shown). No hybridization signal was observed for *A. rhizogenes* isolates [*Agrobacterium* biovar 2; see Sawada *et al.* (1993)]. A hybridization test was also carried out on non-tumorigenic *A. tumefaciens* isolates from soil of the poplar nursery sampled in an area where no galled tree had been observed (Table 1). Under high stringency conditions, a low hybridization signal was detected for two (C91-15 and C91-54) out of ten isolates tested.

DISCUSSION

We determined the entire nucleotide sequence of IS292 present in the *6b* gene of the nopaline-type *Agrobacterium* strain X88-292. It is 2494 bp long with imperfectly matched 21 bp terminal inverted repeats, flanked by a 10 bp target sequence duplication. These structural features are characteristic of a bacterial insertion element. Since no identity with other insertion sequences listed in databases has been revealed, this element is considered to be a novel one. Sequence analysis of IS292 indicated eight ORFs longer than 100 amino acids. Several ORFs are preceded by -35 and -10 consensus sequences and could therefore putatively be transcribed in bacteria. However, a ribosome binding site motif was only identified for ORFs 5 and 8. The presence of such motifs increases the likelihood that these two ORFs encode a polypeptide. Similarity was found between ORF5 and the putative protein 14 from Tc-DNA of pTi15955 whose

function is unknown. However, no known sequence was found related to ORF4.

IS292 is readily distinguishable from IS426 and IS427 isolated from the *Agrobacterium* nopaline-type strain T37 (Vanderleyden *et al.*, 1986; De Meirsmen *et al.*, 1989, 1990), from IS1131 isolated from the nopaline-type strain PO22 (Wabiko, 1992), from IS866 and IS867 isolated from the octopine/vitopine-type strain Tm4 (Bonnard *et al.*, 1989; Paulus *et al.*, 1989b, 1991a) and from IS66 isolated from the octopine-type strain A66 (Machida *et al.*, 1984). However, we found similarity between terminal inverted repeats of IS292, IS66, IS866 and IS1131 from *Agrobacterium*, ISR11 from *R. leguminosarum* and ISRM2 from *R. meliloti*, but limited to a six nucleotide consensus sequence (Fig. 5). In fact, IS292 differs from these elements by the lack of 5'-G and C-3' residues in the border repeats, and by the size of its insertion site and terminal repeats. These observations suggest that the enzymes involved in the transposition process of IS292 are different from those of these other insertion elements since the activity of insertion elements depends on the structure of their extremities. No similarity was found between IS292 and a known transposase sequence in agreement with the organization of ORFs and occurrence of transcriptional signals.

Although IS292 differed from other *Agrobacterium*-associated insertion elements, sequence comparison revealed a strong similarity with ISR11 from *R. leguminosarum* bv. *viciae*. These two insertion sequences are identical in size, show an average of 80% similarity in nucleotide sequence, and ORFs 5, 2 and 8 of IS292 showed amino acid sequence similarity with the corresponding regions of ISR11. This is the first report of such high similarity between two insertion elements from the genera *Agrobacterium* and *Rhizobium* and implies the possibility of a common ancestor. However, data are too scarce to fully understand the evolutionary divergence between IS292 and ISR11, and the chronology of their establishment in their respective hosts.

The presence of insertion elements has been described in the T-DNA of various Ti plasmids (Otten *et al.*, 1992), and they have been proposed to be an important means of evolution of Ti plasmids. In this study, sequence analysis has revealed similarity between the left border of IS292 and the Tc-DNA of pTi15955 (Barker *et al.*, 1983). Similarity between an insertion region and parts of Tc-DNA has also been reported for IS426, IS66 and IS1131 from *Agrobacterium* (De Meirsmen *et al.*, 1987; Machida *et al.*, 1984; Wabiko, 1992), ISR11 from *R. leguminosarum*, IS51 from *Pseudomonas syringae* pv. *savastanoi* (Yamada *et al.*, 1986) and for a plasmid sequence of *R. leguminosarum* 300 (Yun *et al.*, 1987). From these data, Machida *et al.* (1984) proposed that the octopine T-DNA had been split into two parts (T_L and T_R) as a result of insertion events in the domain then called Tc-DNA. The impact of genetic events caused by insertion elements in the evolution of the T-DNA structure has clearly been demonstrated by Paulus *et al.* (1991b) who were able to replace octopine/cucumopine-type *A. vitis* strains in a common evolutionary scheme. In our study, the two nopaline-type

strains X88-292 and 2177 appear closely related since they share a common chromosomal type as determined by the highly resolutive PCR-RFLP analysis of the ribosomal region (unpublished results). Ti plasmid analysis indicated that their main difference is the presence of IS292 in the *6b* gene of pTi292. Moreover, an IS292-like element has been detected in strain 2177, implying that pTi292 could have evolved from pTi2177 as a consequence of IS292 transposition in the *6b* gene. As established by Hooykaas *et al.* (1988), the *6b* gene is an oncogene interfering in the auxin/cytokinin ratio. The *6b* gene alteration might then have an influence on tumour morphology. Preliminary analysis has showed that a strain harbouring pTi292 induces a shooty phenotype on poplar (data not shown).

Although almost all strains from various crown gall origins had no IS292-like sequence, sequences similar to IS292 were detected in several tumorigenic *Agrobacterium* strains involved in a single crown gall outbreak on poplar. IS292 may, therefore, be used as a population marker suitable for epidemiological studies. There was, however, a low hybridization signal with non-tumorigenic soil isolates of *Agrobacterium* from the same biotope. This could be due to fewer copies of the sequence or to a low level of sequence conservation with IS292. This point could be resolved by the analysis of IS292 copy number and distribution in these isolates. The presence of IS292-like sequences in *Agrobacterium* isolates with different chromosomal backgrounds could be the result of horizontal transfers, perhaps through Ti plasmid exchange. This hypothesis is highly probable since a previous study (Nesme *et al.*, 1992) showed the occurrence of natural plasmid transfers among *Agrobacterium* strains involved in this crown gall outbreak on poplar. In addition, the occurrence of insertion elements on the conjugative Ti plasmid could increase opportunities for genetic transfer among strains. The propagation of such elements might lead to an increase in genomic diversity, raising the question as to the role of these mobile DNA sequences in the modification of the genetic structure of the *Agrobacterium* population involved in a crown gall outbreak and the selective effect of a mutation in the T-DNA.

ACKNOWLEDGEMENTS

Thanks are expressed to M. O'Connell and J. K. O'Brien for providing information about ISR11 and to C. Chapelon for technical assistance. C. Ponsonnet received a doctoral grant from the French 'Ministère de la Recherche et de la Technologie'.

REFERENCES

- Barker, R. F., Idler, K. B., Thompson, D. V. & Kemp, J. D. (1983). Nucleotide sequence of the T-DNA region from the *Agrobacterium tumefaciens* octopine Ti plasmid pTi15955. *Plant Mol Biol* **2**, 335–350.
- Bevan, M., Barnes, W. M. & Chilton, M. D. (1983). Structure and transcription of the nopaline synthase gene region of T-DNA. *Nucleic Acids Res* **11**, 369–385.
- Bonnard, G., Vincent, F. & Otten, L. (1989). Sequence and distribution of IS866, a novel T-region associated insertion sequence from *Agrobacterium tumefaciens*. *Plasmid* **22**, 70–81.

- Bouzar, H. (1994). Request for a judicial opinion concerning the type species of *Agrobacterium*. *Int J Syst Bacteriol* **44**, 373–374.
- Brasileiro, A. C. M. (1992). *Caractérisation de la souche d'Agrobacterium tumefaciens 82-139 et son utilisation pour la transformation génétique du peuplier (Populus × canescens)*. Thèse de l'Université Paris-Sud, Centre d'Orsay.
- De Meirsmann, C., Desair, J., Vanderleyden, J., Van Gool, A. P. & Jen, G. C. (1987). Similarities between nucleotide sequences of insertion elements of *Agrobacterium tumefaciens* and *Pseudomonas savastanoi* in relation to *Agrobacterium tumefaciens* Tc-DNA. *Nucleic Acids Res* **15**, 10591.
- De Meirsmann, C., Croes, C., Desair, J., Verreth, C., Van Gool, A. & Vanderleyden, J. (1989). Identification of insertion element IS427 in pTiT37 plasmid DNA of an *Agrobacterium tumefaciens* T37 isolate. *Plasmid* **21**, 129–137.
- De Meirsmann, C., Van Soom, C., Verreth, C., Van Gool, A. & Vanderleyden, J. (1990). Nucleotide sequence analysis of IS427 and its target sites in *Agrobacterium tumefaciens* T37. *Plasmid* **24**, 227–234.
- Dhaese, P., De Greve, H., Decraemer, H., Schell, J. & Van Montagu, M. (1979). Rapid mapping of transposon insertion and deletion mutations in the large Ti plasmids of *Agrobacterium tumefaciens*. *Nucleic Acids Res* **7**, 1837–1849.
- Fickett, J. W. (1982). Recognition of protein coding regions in DNA sequences. *Nucleic Acids Res* **10**, 5303.
- Flores, M., Gonzales, V., Brom, S., Martinez, E., Pinero, D., Romero, D., Davila, G. & Palacios, R. (1987). Reiterated sequences in *Rhizobium* and *Agrobacterium* spp. *J Bacteriol* **169**, 5782–5788.
- Galas, D. J. & Chandler, M. (1989). Bacterial insertion sequences. In *Mobile DNA*, pp. 109–162. Edited by D. E. Berg & M. M. Howe. Washington, DC: American Society for Microbiology.
- Grindley, N. D. F. & Reed, R. R. (1985). Transpositional recombination in prokaryotes. *Ann Rev Biochem* **54**, 863–896.
- Higgins, D. & Gouy, M. (1987). Interfacing similarity search software with the sequence retrieval system ACNUC. *Comput Appl Biosci* **3**, 329–341.
- Hooykaas, P. J. J., Den Dulk-Ras, H. & Schilperoort, R. A. (1988). The *Agrobacterium tumefaciens* T-DNA 6b gene is an *onc* gene. *Plant Mol Biol* **11**, 791–794.
- Lipman, D. J. & Pearson, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227**, 1435–1441.
- Machida, Y., Sakurai, M., Kiyokawa, S., Ubasawa, A., Suzuki, Y. & Ikeda, J. (1984). Nucleotide sequence of the insertion sequence found in the T-DNA region of mutant Ti plasmid pTiA66 and distribution of its homologues in octopine Ti plasmid. *Proc Natl Acad Sci USA* **81**, 7495–7499.
- Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Martinez, E., Romero, D. & Palacios, R. (1990). The *Rhizobium* genome. *Crit Rev Plant Sci* **9**, 59–93.
- Miller, J. H. (1972). *Experiments in Molecular Genetics*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- Nesme, X., Ponsonnet, C., Picard, C. & Normand, P. (1992). Chromosomal and pTi genotypes of *Agrobacterium* strains isolated from *Populus* tumors in two nurseries. *FEMS Microbiol* **101**, 189–196.
- Otten, L., Canaday, J., Gerard, J. C., Fournier, P., Crouzet, P. & Paulus, F. (1992). Evolution of agrobacteria and their Ti plasmids – a review. *Mol Plant–Microbe Interact* **5**, 279–287.
- Paulus, F., Huss, B., Bonnard, G., Ridé, M., Szegedi, E., Tempé, J., Petit, A. & Otten, L. (1989a). Molecular systematics of bt3 Ti plasmids of *Agrobacterium tumefaciens*. *Mol Plant–Microbe Interact* **2**, 64–74.
- Paulus, F., Ridé, M. & Otten, L. (1989b). Distribution of two *Agrobacterium tumefaciens* insertion elements in natural isolates: evidence of stable association between Ti plasmids and their bacterial hosts. *Mol & Gen Genet* **219**, 145–152.
- Paulus, F., Canaday, J. & Otten, L. (1991a). Limited host-range Ti plasmids: recent origin from wide host-range Ti plasmids and involvement of a novel IS element IS868. *Mol Plant–Microbe Interact* **4**, 190–197.
- Paulus, F., Canaday, J., Vincent, F., Bonnard, G., Klares, C. & Otten, L. (1991b). Sequence of *iaa* and *ipt* region of different *Agrobacterium tumefaciens* bt3 octopine strains: reconstruction of octopine Ti plasmid evolution. *Plant Mol Biol* **16**, 601–614.
- Ponsonnet, C. & Nesme, X. (1994). Identification of *Agrobacterium* strains by PCR-RFLP analysis of pTi and chromosomal regions. *Arch Microbiol* **161**, 300–309.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* **74**, 5463–5467.
- Sawada, H., Ieki, H., Oyaizu, H. & Matsumoto, S. (1993). Proposal for the rejection of *Agrobacterium tumefaciens* and revised descriptions for the genus *Agrobacterium* and for *Agrobacterium radiobacter* and *Agrobacterium rhizogenes*. *Int J Syst Bacteriol* **43**, 694–702.
- Shine, J. & 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.
- Staden, R. (1984). Computer methods to locate signals in nucleic acids sequences. *Nucleic Acids Res* **12**, 505–519.
- Vanderleyden, J., Desair, J., De Meirsmann, C., Michiels, K., Van Gool, A. P., Chilton, M. & Jen, G. C. (1986). Nucleotide sequence of an insertion sequence (IS) element identified in the T-DNA region of a spontaneous variant of the Ti plasmid pTiT37. *Nucleic Acids Res* **14**, 6699–6709.
- Wabiko, H. (1992). Sequence analysis of an insertion element, IS1131, isolated from the nopaline-type Ti plasmid of *Agrobacterium tumefaciens*. *Gene* **114**, 229–233.
- Yamada, T., Lee, P. D. & Kosuge, T. (1986). Insertion sequence elements of *Pseudomonas savastanoi*: nucleotide sequence and homology with *Agrobacterium tumefaciens* transfer DNA. *Proc Natl Acad Sci USA* **83**, 8263–8267.
- Yanisch-Perron, C., Vieira, J. & Messing, J. (1985). Improved M13 phage cloning vectors and host strains: nucleotide sequence of the M13mp18 and pUC19 vectors. *Gene* **33**, 103–119.
- Yun, A. C., Hadley, R. G. & Szalay, A. A. (1987). A plasmid sequence from *Rhizobium leguminosarum* 300 contains homology to sequences near the octopine T_L-DNA right border. *Mol & Gen Genet* **209**, 580–584.

Received 11 July 1994; revised 14 November 1994; accepted 24 November 1994.