Mol Gen Genet (1993) 240:49–57

MGG

© Springer-Verlag 1993

Rapid divergence of *Agrobacterium vitis* octopine-cucumopine Ti plasmids from a recent common ancestor

Marc van Nuenen, Patrice de Ruffray, Léon Otten

C.N.R.S. Institute of Plant Molecular Biology, Rue du Général Zimmer 12, F-67084 Strasbourg – Cedex, France

Received: 22 January 1993 / Accepted: 2 March 1993

Abstract. The octopine/cucumopine (o/c) Ti plasmids of the grapevine-associated *Agrobacterium vitis* strains constitute a family of related DNA molecules. Restriction maps were established of two limited-host-range o/c Ti plasmids, pTiAg57 and pTiAB3, and of the wide-host-range o/c Ti plasmid pTiHm1. Together with the previously obtained map of the wide-host-range o/c Ti plasmid pTiTm4, about 1000 kb were mapped with a resolution of 0.2 kb, allowing a detailed comparison of the various structures. One region of the o/c Ti plasmids is highly conserved and differs mainly by the presence or absence of relatively small DNA fragments (0.9–2.7 kb); the other region has been modified more extensively and carries large sequences specific for each Ti plasmid type. The sequence similarity within large conserved regions shows that these plasmids have diverged recently and that their evolution was driven by large-scale genetic events rather than single nucleotide changes. These results have important implications for studies on bacterial evolution.

Key words: *Agrobacterium* – Microbial evolution – RFLP analysis – Ti plasmids

Introduction

Agrobacterium tumefaciens induces plant tumours by transferring part of its DNA into the cells of infected plants. The transferred DNA or T-DNA carries growth-inducing genes and genes coding for the synthesis of small molecules (collectively called opines), which can be metabolized by the bacterium. The T-DNAs are located on a large plasmid, the tumour-inducing or Ti plasmid, which also carries genes necessary for T-DNA transfer (virulence genes), opine utilization, plasmid replication and conjugational transfer (reviewed in Kado 1991).

Communicated by A. Kondorosi

Correspondence to: L. Otten

The species *A. vitis* (previously called *A. tumefaciens* biotype III; Ophel and Kerr 1990) is specifically associated with grapevine *(Vitis vinifera)* and can be distinguished from the well-known biotype I and II strains on the basis of chromosomally encoded properties (Kerr and Panagopoulos 1977). Opine utilization characteristics and hybridization studies with different T-DNA probes define three *A. vitis* subgroups: octopine/cucumopine (o/c), nopaline and vitopine strains, with o/c, nopaline and vitopine Ti plasmids, respectively (Szegedi 1985; Paulus et al. 1989a). The o/c strains occur worldwide; they have been divided into wide-host-range (WHR) strains, which induce tumours on tobacco and tomato, and limited-host-range (LHR) strains which do not (Knauf et al. 1982). WHR plasmids and LHR plasmids differ in host range and *Sma*I fingerprints but share extensive DNA homology (Knauf et al. 1983; Thomashow et al. 1981). The host range of LHR o/c Ti plasmids can be extended by the introduction of a cytokinin synthesis *(ipt)* gene from the T-DNA of WHR biotype I plasmids (Yanofsky et al. 1985a; Buchholz and Thomashow 1984; Hoekema et al. 1984). The *virA* and *virC* genes from biotype I strains further enlarge the host range of LHR strains (Yanofsky et al. 1985b). These host range studies were carried out with the LHR strains Ag162 (Yanofsky et al. 1985b), Ag63 (Buchholz and Thomashow 1984) and Ag57 (Hoekema et al. 1984). The restriction map of pTiAg162 has been published (Knauf et al. 1984).

The LHR and WHR o/c Ti plasmids carry two T regions, TA and TB (Knauf et al. 1984; Buchholz and Thomashow 1984; Yanofsky et al. 1985a). The TA region is highly homologous to the TL region of the octopine Ti plasmids of the biotype I strains and has the same overall organization; the TB region is unique for o/c Ti plasmids and carries a set of auxin genes (Knauf et al. 1984; Buchholz and Thomashow 1984; Yanofsky et al. 1985a) and the cucumopine synthase gene (Paulus et al. 1989a; Fournier et al. 1993). The structures of these regions differ in different strain types, but the sequences in the common regions are more than 99.7% homolo-

Table 1. Bacterial strains and plasmids

Bacteria	Description	Reference or origin
Escherichia coli:		
NM522	Host for pUC clones	Gough and Murray (1983)
Agrobacterium:		
AB3, Zw2	*A. vitis,* o/c, small TA	Szegedi (1985)
2607, 2612, 2613, 2614, 2617, 2644, 2651, 2653, 2656, 2676	*A. vitis,* o/c, small TA	Paulus et al. (1989a)
NW90, NW102	*A. vitis,* o/c, small TA	Paulus et al. (1989b)
Tm4, Hm1	*A. vitis,* o/c, large TA	Szegedi (1985)
Ag82	*A. vitis,* o/c, large TA	C. Panagopoulos
2686	*A. vitis,* o/c, large TA	Paulus et al. (1989a)
NCPPB3554	*A. vitis,* reference strain	Ophel and Kerr (1990)
550–22	*A. vitis,* o/c, large TA	M. Lopez
48, 49	*A. vitis,* o/c, large TA	M. Ridé
Ag57	*A. vitis,* o/c, small TA	Panagopoulos and Psallidas (1973)
LBA649	C58-C9(pTiAg57), exconjugant	Hoekema et al. (1984)
Plasmids:		
pUC18	Cloning vector	Yanisch-Perron et al. (1985)
pPM1016	Kanamycin-resistant pUC18 derivative	Huss et al. (1989)
H176	Partial *Hin*dIII clone pTiTm4 (coordinates 51.2–69.0)	Otten et al. (1992b)

gous. A detailed comparison of the TA regions of five different o/c Ti plasmids (LHR: pTiAg57, pTiAB3 and pTiNW233; WHR: pTiHm1 and pTiTm4) revealed that they can be derived from a common large ancestral TA region by insertion of and recombination between bacterial insertion elements (reviewed in Otten et al. 1992a). pTiAg57, pTiAB3 and pTiNW233 have a small TA region, whereas pTiHm1 and pTiTm4 have conserved a large TA region. The IS elements that contributed to the divergence of the TA region (IS*866*, IS*868*, IS*869*, ISX and ISY) are not present in other lineages and were probably introduced by horizontal gene transfer (Otten et al. 1992a). Up to now, we have found the TA region structure of pTiAB3 in 24 isolates, that of pTiTm4 in 31 isolates and that of pTiNW233 in two isolates. The TA regions of pTiAg57 and pTiHm1 are unique. The TA and TB regions of pTiAB3 (cloned by our group) are very similar and possibly identical to those of pTiAg162 (Knauf et al. 1984). Here we present the complete restriction maps of pTiAg57, pTiAB3 and pTiHm1 and a detailed comparison of the four main o/c Ti plasmid types.

Materials and methods

Bacterial strains and plasmids. Bacterial strains and plasmids used in this study are listed in Table 1.

Construction of pTi maps. Ti plasmid DNA of strain LBA649 [C58-C9(pTiAg57)], AB3 and Hm1 was isolated according to Currier and Nester (1976) and partially digested with various restriction enzymes. Fragments were separated on a sucrose gradient and 15–25 kb fragments were ligated to pPM1016 (a kanamycin-resistant pUC18 derivative) or pUC18. Restriction analysis of about 200 overlapping clones per Ti plasmid yielded partial restriction maps for *Pst*I, *Eco*RI and *Hin*dIII. Remaining gaps were closed by using end fragments of unlinked groups as probes against clone banks or by hybridizing probes from related Ti plasmids to Southern blots of Ti plasmid DNA digested with different enzymes (for general methods see Sambrook et al. 1989).

Results

Cloning and restriction mapping of pTiAg57, pTiAB3 and pTiHm1

In previous studies it was established that pTiAg57, pTiAB3, pTiHm1 and pTiTm4 represent four major types of o/c Ti plasmids, whereas pTiNW233 is closely related to pTiAg57 (summarized in Otten et al. 1992a). The restriction map of pTiTm4 (256.5 kb) has been published recently (Otten et al. 1992b), those of pTiAg57 (224.5 kb), pTiAB3 (234.5 kb) and pTiHm1 (258.0 kb) were established by the same methods. Different regions were cloned with different frequencies: one region conserved in all three plasmids (between coordinates 0 and

Fig. 1a, b. Restriction maps of pTiAg57, pTiAB3, pTiHm1 and pTiTm4. Abbreviations: H; *Hin*dIII; P, *Pst*I; E, *Eco*RI. Sizes are in kb. **a** Region I. The right TB border of the o/c Ti plasmids has been localized to within 2 kb inside the 7.3 kb *Hin*dIII fragment (Yanofsky et al. 1985a). VIR, Virulence region (Knauf et al. 1984); OCC, octopine catabolism region (Knauf et al. 1984). The *numbers* indicate the different features of the Ti plasmids, detailed in Table 3. **b** Region II. B, *Bam*HI sites of clone AB1; INC, incompatibility region of pTiAg162 (Knauf et al. 1984); ORI, region of pTiTm4 containing origin of replication (Otten et al. 1992b)

a

pTiAg57

VIR

OCC

TA

TB

pTiAB3

pTiHm1

pTiTm4

b

pTiAg57

pTiAB3

INC

pTiHm1

pTiTm4

P663 (ORI)

H176

7, see below) was particularly difficult to clone and found only once: Ap199 from Ag57. The reason for this was not further investigated. Hybridization of Ap199 to *Hin*dIII-, *Pst*I- and *Eco*RI-digested DNA from pTiAB3 and pTiHm1 yielded the same restriction patterns as for pTiAg57. A 0.7 kb gap in pTiAB3 was closed by first establishing detailed restriction maps for the clones flanking this gap and then hybridizing Southern blots of restricted pTiAB3 DNA with each of the two flanking regions. To eliminate possible cloning artefacts due to ligation of non-contiguous fragments, representative clones covering the entire Ti plasmid were hybridized against plasmid DNA digested with *Eco*RI, *Pst*I or *Hin*dIII. The hybridization patterns (not shown) confirmed the maps derived from comparison of the clones. In general, very few clones with aberrant restriction patterns were found. The maps of the three Ti plasmids are shown in Fig. 1a and b together with the map of pTiTm4 (Otten et al. 1992b). The *Eco*RI map of pTiAB3 showed several differences from the *Eco*RI map of pTiAg162 (Knauf et al. 1984). Since we did not have Ag162 or pTiAg162 clones available, these differences could not be investigated further.

The 0 coordinate of pTiTm4 has been defined as the right *Hin*dIII site of the 0.94 kb *Hin*dIII fragment located to the right of the TB region (Otten et al. 1992b). This site is also present in pTiAg57, pTiAB3 and pTiHm1 and can therefore be used as the general 0 coordinate for all four o/c Ti plasmids. The 0 coordinate chosen earlier for pTiAg162 (Knauf et al. 1984) is an *Xho*I site, which is only found in plasmids of the pTiAB3/pTiAg162 type and corresponds to coordinate 166.0 of pTiAB3. The alignment of the four o/c Ti plasmid maps showed that about one-half of the o/c Ti plasmids is well-conserved (see below). This region was called region I and contains the virulence region, the TA region, the octopine utilization region and the TB region (Fig. 1a). The other half is much less conserved and was called region II. In pTiAg162, this region contains the incompatibility functions (Knauf et al. 1984) and in pTiTm4 was shown (Otten et al. 1992b) to contain the origin of replication (Fig. 1b). In the following comparisons, different Ti fragments are indicated with a letter (A, B, H and T for pTiAg57, pTiAB3, pTiHm1 and pTiTm4 respectively), followed by the corresponding coordinates in kb within brackets.

Region I

The restriction maps of regions I of the four Ti plasmids, defined as A (108–224.5), B(115–234.5), H(129–258) and T(131.5–256.5), show a number of differences which are indicated by numbers 1 to 20 in Fig. 1a (see also Table 3). The left end of region I is highly conserved: T(131.5–174) and H(129.5–171.5) are identical (all 48 restriction sites, covering 42.5 kb, are conserved, indicating a DNA homology of more than 99%). A(108–150.5) and B(115–158) differ from pTiTm4 and pTiHm1: they lack a *Pst*I site (number 2, Fig. 1a); pTiAB3 carries in addition an extra DNA fragment of 1.3 kb (number 1,

Table 2. Selected pTi clones

Clone	Restriction fragments (in kb)
pTiAg57 clones:	
Ap199	9.2
Ap21	1.47–6.5
Ap292	3.6–0.35–4.0–0.2–4.3–2.6–5.3–(1.7–0.5)
Ab1	8.0
Ap23	3.7–6.05–1.7–1.5–6.1–2.9
Ah136	10.7–0.9–10.0–0.95–1.1–2.2–(1.4–2.1)
Ah88	1.1–2.2–(1.4–2.1)–0.4–8.9
Ah11	0.4–8.9–3.9–13.5–4.4
Ap93	3.0–6.0–2.6–5.4
Ap2	2.6–5.4–5.35–2.4–0.87
Ah201*	2.8–0.7–2.9–0.8–10.2
Ah148	0.8–10.2–0.2–4.0–1.3
Ah10	1.3–8.1–1.7–3.5–7.5–1.1–5.1
Ah203*	5.1–2.8–0.35–0.5–0.35–6.8–4.2
pTiAB3 clones:	
[Ap199]	9.2
Bp31	7.5–3.6–0.35–4.0
Bh1*	2.6–8.4–2.85–0.82
Bp200	5.1
Bh2*	12.2–1.65–1.7–(1.3–2.3)
Bh3*	1.9–3.4–0.35–1.45–2.1–10.7
Bp40	1.8–(0.5–0.3)–3.7–4.8–0.2–5.0–0.35
Bp1	5.0–0.35–1.45–2.9–3.5–1.3–3.8
Bp201	4.2
Bh4*	2.6–13.5
Bp39	1.5–3.2–1.3–3.0–7.05
Bp66	2.6–5.4–5.35–2.4–0.8–4.75–2.4–0.2
Bh5*	1.8–0.65–4.1–2.2–0.7–2.2–0.3–3.7–0.35
Bp42	8.25–3.0
Bp41	3.0–2.1–4.8
Bh6*	1.3–8.1–1.7–3.5–7.5
Bh7*	3.5–7.5–1.1–6.5–0.35–0.52–0.35
Bh8*	0.35–0.6–0.35–7.3
Bh9*	7.3–4.2–0.94
pTiHm1 clones:	
[Ap199]	9.2
[Ap21]	1.47–6.5
He121	7.9–12.25–2.3–2.5–0.4–3.2–0.95
He134	0.4–3.2–0.95–0.4–0.7–4.2–5.6
He183	0.95–0.4–0.7–4.2–5.6–2.2–1.6–3.9–1.55–2.3
He156	2.3–10.4–5.2
He152	5.2–3.2–4.9–1.4–0.5–0.2–0.9–7.2–0.75–0.7
He208	7.2–0.75–0.7–2.0–0.9–0.9–2.1–7.8–5.2
He225	0.9–2.1–7.8–5.2–1.75–1.6–6.2–3.5
He93	3.5–1.9–5.0–0.4–5.07–5.4
Hp151	3.0–6.0–2.6–5.4–5.35–2.4
He51	2.0–6.05–3.85–1.55–2.0–0.2–3.9–1.5
He182	1.5–0.2–4.3–4.8–4.8–4.75–2.5
He65	2.5–0.2–10.4–3.2
He35	0.2–10.4–3.2–6.2–2.0–3.5–0.5
He145	3.5–0.5–0.3–0.9–1.6–2.3–4.1–1.15–1.47–5.5
He138	1.47–5.5–16.5–6.1

A, B, H: pTiAg57, pTiAB3 and pTiHm1 clones. b, e, h and p: *Bam*HI, *Eco*RI, *Hin*dIII and *Pst*I partials
Fragments in *parentheses* represent non-ordered fragments; clones shown in *square brackets* represent heterologous clones with an identical structure; an *asterisk* indicates pPM1016 derivatives and the other clones are pUC18 derivatives

Table 3. Specific features of Ti plasmids (numbering is as in Fig. 1a)

Position	pTi coordinates	Nature of difference	Reference
1	B	+1.3 kb fragment	This work
2	T, H	Extra *Pst*I site	This work
3	T	Lack of 3.2 kb fragment	This work
4	H	+1.3 kb fragment	This work
5	A, B	IS*868*-induced deletion (8 kb)	Paulus et al. (1991b)
6	B	+1.3 kb (IS*868*)	Paulus et al. (1991b)
7	B	+0.9 kb (IS*869*)	Paulus et al. (1991c)
8	H	+1.3 kb (ISX)	Paulus et al. (1992)
9	T	+2.7 kb (IS*866*)	Bonnard et al. (1989)
10	A	IS*867*-induced deletion (0.6 kb)	Paulus et al. (1991a)
11	A	0.1 kb deletion	Paulus et al. (1991a)
12	T	Additional *Pst*I site	This work
13	T	Lack of *Pst*I site	This work
14	T	Additional 0.15 kb fragment	This work
15	T	Lack of 0.5 kb fragment	This work
16	T	Additional 1.0 kb	This work
17	T, H	Additional 1.3 kb	This work
18	A	+1.1 kb (IS*870*)	Fournier et al. (1993)
19	B	+0.9 kb (IS*869*-like)	This work
20	H	+1.3 kb (ISX-like)	This work

A, B, H and T, pTiAg57, pTiAB3, pTiHm1 and pTiTm4

Fig. 1a). Mutation analysis of pTiAg162 (Knauf et al. 1984) and hybridization with the biotype I Ti plasmid pTiB$_6$806 (Thomashow et al. 1981) showed the pTiAg162 *vir* genes to be located as indicated in Fig. 1a. This region is conserved between pTiAg162 and the four Ti plasmids. To the right of it, the regions corresponding to A(150.5–157.5) and ending at the left TA border (Paulus et al. 1991a) differ in size: 7.1 kb for pTiAg57 and pTiAB3, 8.4 and 3.9 kb for pTiHm1 and pTiTm4 respectively. The exact nature of these differences (3 and 4 in Fig. 1a) requires further study.

The TA regions of the four plasmids have been described earlier and the differences within these regions (5–10, Fig. 1a) have been explained as insertions and deletions due to bacterial insertion sequences (Otten et al. 1992a; Table 2). Between the TA and TB region, pTiAg162 carries the genes for octopine utilization (Knauf et al. 1984) and hybridizes to the octopine catabolism region of pTiB$_6$806 (Thomashow et al. 1981) as does the corresponding region of pTiTm4 (Otten et al. 1992b). The 27.5 kb region from the common *Hin*dIII site at A(164.5) up to the left border of the TB region (identified in pTiAB3 by sequence analysis, Fournier et al. 1993) is identical in pTiHm1, pTiAB3 and pTiAg57 (all 28 sites conserved, indicating more than 99% sequence identity). pTiTm4 differs from the three other Ti plasmids at four positions (numbers 12–15 in Fig. 1a, Table 3).

The TB regions of pTiAg57, pTiAB3, pTiHm1 and pTiTm4 have recently been compared in detail (Fournier et al. 1993); the differences are due to insertions and/or deletions within a common ancestral structure (numbers 16–20 in Fig. 1a, Table 3).

Considered overall, the nucleotide sequences of the ancestral parts of region I are extremely well conserved: only three restriction site differences (numbers 2, 12 and 13 in pTiTm4) may result from single nucleotide changes. For a total of about 100 common sites per Ti plasmid this represents a nucleotide sequence conservation of 99.5–99.9%.

Region II

Whereas the restriction maps of region I can be aligned without difficulty, those of region II cannot. On the basis of their restriction patterns the following subregions can be distinguished (Figs. 1b and Fig. 2):

Subregion a. A 38 kb fragment to the right of the TB region (indicated by the letter a in Fig. 2) is identical for pTiHm1 and pTiAg57: H(252.5–32.5) and A(219–32.5). B(229–32.5) is very similar; it is slightly shorter at its right end and carries an additional 1.0 kb fragment indicated by the number 22 in Fig. 1b. Only 14 kb of pTiTm4 [T(251.0–7.0)] (indicated by a′ in Fig. 2) can be aligned with subregion a, with an additional *Pst*I site at T(0.7) (Fig. 1b, number 21).

Subregion b. To the right of subregion a, pTiAg57 and pTiAB3 differ considerably for 6 and 11.5 kb respectively. To the right of this region, A(37.5–83.5) and B(44.5–90.5) are again identical for 46 kb (53 out of 53 sites conserved, subregion b). The restriction pattern of subregion b is not found in pTiHm1 or pTiTm4.

Subregion c. To the right of subregion b, pTiAg57 and pTiAB3 differ for over 13 kb. Subsequently, they are again identical for 13.5 kb [subregion c, A(96.5–110.0) and B(103.5–117.0)]. Homology with pTiHm1 and pTiTm4 only exists for the right part of this region (c′).

Subregion d. To the right of subregion a of pTiHm1, H(34.5–38.0) is identical to T(66.0–69.5) in subregion d.

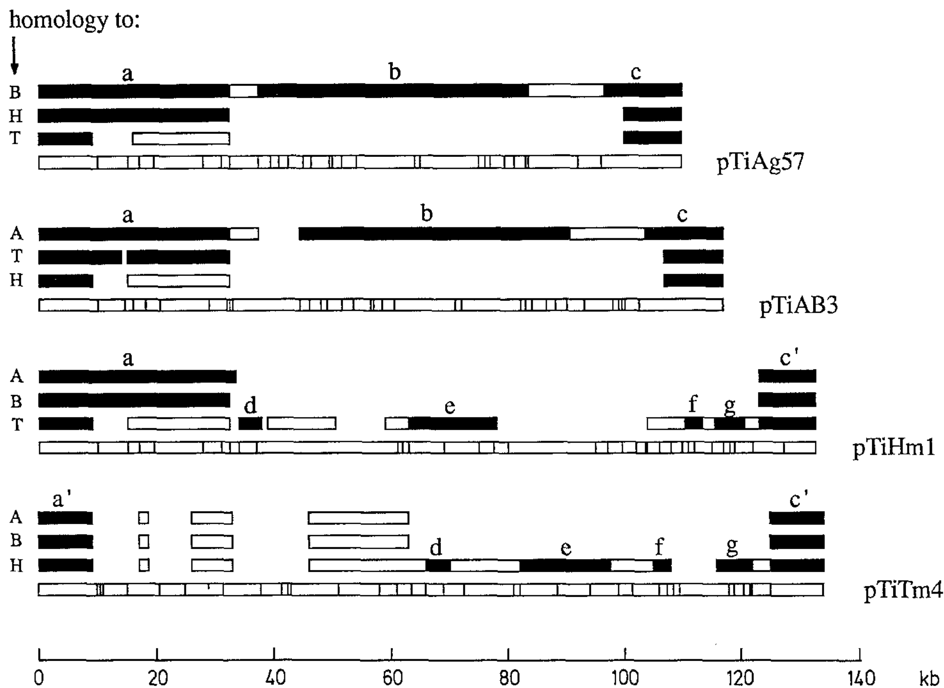

Fig. 2. Schematic comparison of region II of pTiAg57, pTiAB3, pTiHm1 and pTiTm4. For each Ti plasmid (identified on the right) the *Hin*dIII map is shown and above this map the regions of homology with the three other Ti plasmids (identified on the left). *Black boxes,* regions with common restriction sites; *white boxes,* regions without common restriction sites but with DNA homology as determined by Southern analysis. The *lower case letters* refer to subregions discussed in the text. A, B, H and T: pTiAg57, pTiAB3, pTiHm1 and pTiTm4 respectively

T(66.0–69.5) is situated to the right of a 57.5 kb pTiTm4 region which has a unique restriction pattern.

Subregion e. The 15 kb regions H(63.0–78.0) and T(82.0–97.0), although not completely identical, have several restriction sites in common.

Subregions f and g. Two additional small regions of pTiHm1 and pTiTm4 have common restriction sites: subregion f, H(110.0–113.0) and T(105.0–108.0), and subregion g, H(114.5–119.5) and T(117.0–121.5).

DNA homology between the o/c Ti plasmids

Regions that lack common restriction sites may still be related; this was investigated by hybridizing clones of pTiHm1 with pTiTm4, pTiAB3 and pTiAg57. The DNA homology between pTiHm1 and pTiTm4 covers H(0.0–9.0), H(15.0–51.0), except the 0.95 kb *Eco*RI fragment at 38.0–39.0, H(59.5–78.0) and H(104.0–132.0) whereas the homology between pTiHm1 and pTiAB3/pTiAG57 covers H(0.0–31.5) and H(122.5–132.0). In an earlier study, different pTiTm4 clones were hybridized with pTiAB3 and pTiAg57 (Otten et al. 1992b). T(0.0–10.0), T(17.0–18.5), T(26.0–33.0), T(46.5–63.5) and T(125.0–134.5) hybridized. Clone H176 of pTiTm4 (Table 1) was hybridized with pTiAg57 and pTiAB3 to establish the end of the homologous region. H176 hybridizes to the end of subregion a; at this position all plasmids have a 2.85 kb *Hin*dIII fragment in common. In pTiAg162 this region has been shown to carry incompatibility genes (Knauf et al. 1984). The results of the hybridization studies of region II are summarized in Fig. 2. Compared with pTiAg57, pTiAB3 and pTiHm1, pTiTm4 seems to have a modified subregion a with about 30 kb of additional DNA. To the right of subregion a, pTiHm1 shows partial homology with pTiTm4 interrupted by specific regions (between subregions d and e and between e and f). Compared to pTiHm1, pTiTm4 has an additional fragment between subregions f and g. Region II seems to have undergone more large-scale changes than region I. The precise nature of these changes requires further investigation.

Divergence pattern of the o/c Ti plasmids

The genetic distance between related DNAs is often measured by comparing the restriction fragment patterns of homologous regions (restriction fragment length polymorphism or RFLP analysis, Nei and Li 1979). This analysis is based on the assumption that divergence mainly results from single nucleotide changes (point mutations). However, the o/c Ti plasmids obviously do not evolve in this way and classical RFLP analysis therefore leads to an erroneous estimate of the divergence rates. In the case of pTiAg57 and pTiHm1 for example, the proportion of shared fragments (also called similarity coefficient, F) is 0.50 which would indicate 4% nucleotide divergence. However, only one out of 101 ancestral sites has been changed by nucleotide substitution (marked by number 2 in Fig. 1a). This yields a nucleotide substitution value per site of only 0.05%.

Due to the very low sequence divergence, a phylogenetic tree for the o/c Ti plasmids can only be constructed by comparing distinguishing features shared by the various regions I (shared feature analysis). As an example, consider the 1.3 kb fragment in the TB region indicated by number 17 (Fig. 1a). This fragment is present in pTiHm1 and pTiTm4, but not in pTiAg57 and pTiAB3. Therefore, pTiTm4 and pTiHm1 share a com-

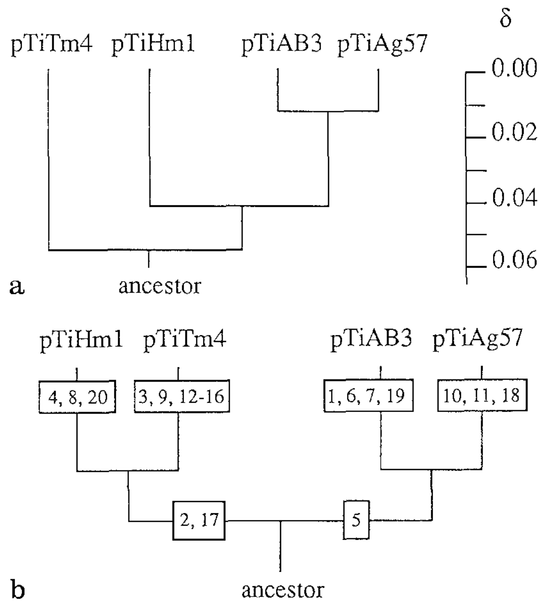

Fig. 3a, b. Phylogenetic trees of the four main o/c Ti plasmid types. **a** Reconstruction on the basis of shared restriction fragments (Nei and Li 1979). Proportion of shared fragments (F) between two Ti plasmids was calculated according to the formula $F = 2n_{XY}/(n_X + n_Y)$ where n_X and n_Y are the numbers of fragments in plasmids X and Y respectively, whereas n_{XY} is the number of fragments shared by the two plasmids. F values were (A, Ag57; B, AB3; H, Hm1; T, Tm4): A/B, 0.79; A/H, 0.50; A/T, 0.38; B/H, 0.47; B/T, 0.39; H/T, 0.47. Divergence values (δ) calculated from F values according to Nei and Li (1979) were 1.31, 4.03, 5.49, 4.38, 5.6 and 4.4 respectively. **b** Reconstruction on the basis of shared features within region I. The *numbers* refer to the positions indicated in Fig. 1a

mon ancestor (X) which contained this fragment, whereas pTiAg57 and pTiAB3 share a common ancestor (Y) which lacked it. Until more is known about the fragment and its surrounding region, whether or not the common ancestor of X and Y contained this fragment cannot be decided. In the case of features 5–11 of the TA region (Otten et al. 1992) and 18–20 of the TB region (Fournier et al. 1993) sequence analysis showed that all features are due to secondary events. The analysis of the distribution of the 20 discriminating features of region I (Fig. 1a, Table 3) yields a coherent branching pattern. Since the various features do not result from the same types of event, the relative sizes of the branches remain unknown. It should be noted that the branching pattern thus obtained differs from the pattern derived from RFLP analysis: whereas classical RFLP analysis suggests that pTiHm1 should be grouped with pTiAB3 and pTiAg57, shared feature analysis groups pTiHm1 with pTiTm4 (Fig. 3a and b).

Other o/c Ti plasmids

The TA region structures of pTiAB3 and pTiTm4 were also found in several other o/c isolates. These strains were therefore called "AB3-like" and "Tm4-like" strains. To establish whether the maps of pTiAB3 and pTiTm4 were representative, total DNAs of a number of AB3- and Tm4-like strains were hybridized with cloned Ti fragments. A Southern blot with *Pst*I-digested DNA of seven Tm4-like strains (all containing IS*866* within their TA region: Paulus et al. 1989b; L. Otten, unpublished data) was hybridized with pTiAB3 clone Bp42 which covers the pTiTm4-specific 1.9+1.0 *Pst*I fragments (Fig. 1a, number 12). Whereas 550–22, NCPPB3554, 2686 and Tm4 show the pTiTm4 pattern, 48, 49 and Ag82 show the 3.0 kb *Pst*I fragment of pTiAg57, pTiAB3 and pTiHm1 (results not shown). This strongly suggests that the 3.0 kb fragment belongs to the ancestral o/c Ti plasmid and underwent a mutation after the insertion of the pTiTm4-characteristic IS*866* element in the TA region had occurred.

In a second experiment, a Southern blot was prepared with *Pst*I-digested total DNA of 14 AB3-like strains (AB3, Zw2, NW102, NW90, 2607, 2612, 2613, 2614, 2617, 2644, 2651, 2653, 2656 and 2676), which are characterized by the presence of IS*868* within their TA region: Paulus et al. 1989b; L. Otten, unpublished data). The blot was hybridized to Bp39 which covers the pTiAB3-specific 7.05 kb *Pst*I fragment (Fig. 1a, number 1): all strains had the characteristic pTiAB3 pattern. Finally, hybridization of the same set of DNAs to another pTiAB3 clone, Bh5, yielded the expected *Pst*I patterns, except for 2612 which showed a 13 kb *Pst*I fragment instead of a 8.25 kb fragment (results not shown). This indicates that one of the immediate ancestors of pTi2612 underwent a mutation within the region covered by Bh5.

Discussion

In this study we have compared four related Ti plasmids of the o/c group: pTiAg57, pTiAB3, pTiHm1 and pTiTm4. In certain areas these plasmids are more than 99.9% homologous as shown by sequence analysis (TA regions: Paulus et al. 1991a and b; TB regions: Fournier et al. 1993). The same conclusion emerges from comparisons of the restriction maps of common areas. Nevertheless many differences exist between the Ti plasmids, although these differences are not distributed evenly. Region I with the *vir* genes and the two T-regions (about 125 kb) is well conserved: almost all of the differences are due to relatively small (0.9–2.7 kb) insertions and/or deletions. If it is assumed that each insertion/deletion represents a single molecular event, region I has diverged in only a few steps. The origin of the various region II structures is more complicated; this region contains large fragments specific for each Ti plasmid type. However many o/c isolates are available (Paulus et al. 1989a, b) and further analysis may allow the reconstruction of the evolutionary history of region II.

Classical phylogenetic analysis relies on nucleotide sequence divergence between related organisms. Divergence values can be obtained by the alignment of well-defined and generally relatively short DNA regions (like 16S RNA genes), or by counting common restriction fragments, mostly by using randomly chosen DNA fragments as probes (Upholt 1977; Nei and Li 1979; McFadden et al. 1987; Clark-Curtiss and Walsh 1989; Gabriel

et al. 1988; Cook et al. 1989; Regnery et al. 1991). Neither of these methods can be used for the o/c Ti plasmids since homologous regions do not show sufficiently high levels of nucleotide sequence divergence. Restriction fragment analysis yields similarity values, but these cannot be used to calculate percentages of single nucleotide changes necessary for the construction of a phylogenetic tree. The possible occurrence of large-scale alterations in bacterial genomes and the resulting problems for phylogenetic studies have been noted before (see for example Arber 1991). The present study of four related plasmids with sizes between 225 and 258 kb constitutes one of the first demonstrations that such complications do occur within natural populations of a common bacterial species and that such cases cannot be studied by RFLP analysis.

The phylogeny of DNA molecules in which large-scale changes predominate can only be obtained by a painstaking comparison of related structures and detailed sequence analysis of crucial areas. Such studies may or may not reveal the nature and the temporal sequence of the events that were required to produce the observed differences. Rapid divergence of the type observed here might make if difficult to find surviving intermediates. Reconstruction (as was done for the evolution of the o/c TA regions; Otten et al. 1992a) may still be carried out if divergence has occurred recently, but the low amount of sequence divergence makes it very difficult to obtain a time-scale for the phylogenetic tree. If the phenomena reported here occur more generally the reconstruction of bacterial phylogeny becomes much more difficult than anticipated. The changes leading to the different o/c Ti plasmids have taken place recently. We have postulated (Paulus et al. 1991b) that the rapid divergence of an ancestral o/c Ti plasmid coincided with the development of large-scale grapevine culture. Further studies are required to define the functional consequences of these changes under both natural and laboratory conditions. It would also be worthwile to extend the analysis of o/c strains to a consideration of the evolution of their chromosomes.

Acknowledgements. We thank C. Panagopoulos, M. Lopez, M. Ridé and E. Bien for *Agrobacterium* strains. M.v N. was supported by an ERASMUS grant from the European Communities.

References

Arber W (1991) Elements in microbial evolution. J Mol Evol 33:4–12

Bonnard G, Vincent F, Otten L (1989) Sequence and distribution of IS*866*, a novel T-region associated insertion sequence from *Agrobacterium tumefaciens*. Plasmid 22:70–81

Buchholz WG, Thomashow MF (1984) Host range encoded by the *Agrobacterium tumefaciens* tumor-inducing plasmid pTiAg63 can be expanded by modification of its T-DNA oncogene complement. J Bacteriol 160:327–332

Clark-Curtiss JE, Walsh GP (1989) Conservation of genomic sequences among isolates of *Mycobacterium leprae*. J Bacteriol 171:4844–4851

Cook D, Barlow E, Sequeira L (1989) Genetic diversity of *Pseudomonas solanacearum:* detection of restriction fragment length polymorphisms with DNA probes that specify virulence and the hypersensitive response. Mol Plant-Microbe Interact 2:113–121

Currier TC, Nester EW (1976) Isolation of covalently closed circular DNA of high molecular weight from bacteria. Anal Biochem 76:431–441

Fournier P, Paulus F, Otten L (1993) IS*870* requires a 5′-CTAG-3′ target sequence to generate the stop codon for its large ORF1. J Bacteriol, in press

Gabriel DW, Hunter JE, Kingsley MT, Moller JW, Lazo GR (1988) Clonal population structure of *Xanthomonas campestris* and genetic diversity among citrus canker strains. Mol Plant-Microbe Interact 1:59–65

Gough J, Murray N (1983) Sequency diversity among related genes for recognition of specific targets in DNA molecules. J Mol Biol 166:1–19

Hoekema A, de Pater BS, Fellinger AJ, Hooykaas PJJ, Schilperoort RA (1984) The limited host range of an *Agrobacterium tumefaciens* strain extended by a cytokinin gene from a wide host range T-region. EMBO J 3:3043–3047

Huss B, Bonnard G, Otten L (1989) Isolation and functional analysis of a set of auxin genes with low root-inducing activity from an *Agrobacterium tumefaciens* biotype III strain. Plant Mol Biol 12:271–283

Kado CI (1991) Molecular mechanisms of crown gall tumorigenesis. Crit Rev Plant Sci 10:1–32

Kerr A, Panagopoulos CG (1977) Biotypes of *Agrobacterium radiobacter* var. *tumefaciens* and their biological control. Phytopath Z 90:172–179

Knauf VC, Panagopoulos CG, Nester EW (1982) Genetic factors controlling host range of *Agrobacterium tumefaciens*. Phytopathology 12:1545–1549

Knauf VC, Panagopoulos CG, Nester EW (1983) Comparison of Ti plasmids from three different biotypes of *Agrobacterium tumefaciens* isolated from grapevines. J Bacteriol 153:1535–1542

Knauf VC, Yanofsky M, Montoya A, Nester EW (1984) Physical and functional map of an *Agrobacterium tumefaciens* plasmid that confers a narrow host range. J Bacteriol 160:564–568

McFadden JJ, Butcher PD, Chiodini R, Hermon-Taylor J (1987) Crohn's disease-isolated mycobacteria are identical to *Mycobacterium paratuberculosis,* as determined by DNA probes that distinguish between mycobacterial species. J Clin Microbiol 25:796–801

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Ophel K, Kerr A (1990) *Agrobacterium vitis* sp. nov. for strains of *Agrobacterium* biovar 3 from grapevines. Int J Syst Bacteriol 40:236–241

Otten L, Canaday J, Gérard J-C, Fournier P, Crouzet P, Paulus F (1992a) Evolution of *Agrobacteria* and their Ti plasmids. A review. Mol Plant-Microbe Interact 5:279–287

Otten L, Gérard J-C, de Ruffray P (1992b) The Ti plasmid from the wide host range *Agrobacterium vitis* strain Tm4; map and homology with other Ti plasmids. Plasmid 29:154–159

Panagopoulos CG, Psallidas PG (1973) Characteristics of Greek isolates of *Agrobacterium tumefaciens* (EF Smith and Townsend) Conn. J Appl Bacteriol 36:233–240

Paulus F, Huss B, Bonnard G, Ridé M, Szegedi E, Tempé J, Petit A, Otten, L (1989a) Molecular systematics of biotype III 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 for stable association between Ti plasmids and their bacterial hosts. Mol Gen Genet 219:145–152

Paulus F, Huss B, Tinland B, Herrmann A, Canaday J, Otten L (1991a) Role of T-region borders in *Agrobacterium* host range. Mol Plant-Microbe Interact 4:163–172

Paulus F, Canaday J, Otten, l (1991b) Limited host range Ti plasmids: recent origin from wide host range Ti plasmids and in-

volvement of a novel IS element, IS*868*. Mol Plant-Microbe Interact 4:190–197

Paulus F, Canaday J, Vincent F, Bonnard G, Kares C, Otten L (1991c) Sequence of the *iaa* and *ipt* region of different *Agrobacterium tumefaciens* biotype III octopine strains: reconstruction of octopine Ti plasmid evolution. Plant Mol Biol 16:601–614

Regnery RL, Spruill CL, Plikaytis BD (1991) Genotypic diversification of *Rickettsiae* and estimation of intraspecies sequence divergence for portions of two rickettsial genes. J Bacteriol 173:1576–1589

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Szegedi E (1985) Host range and specific L(+)-tartrate utilization of biotype 3 *Agrobacterium tumefaciens*. Acta Phytopathol Acad Sci Hung 20:17–22

Thomashow MF, Knauf VC, Nester EW (1981) Relationship between the limited and wide host range octopine-type Ti plasmids of *Agrobacterium tumefaciens*. J Bacteriol 146:484–493

Upholt WB (1977) Estimation of DNA sequence divergence from comparison of restriction endonuclease digests. Nucleic Acids Res 4:1257–1265

Yanisch-Perron C, Vieira J, Messing J (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103–119

Yanofsky M, Montoya A, Knauf VC, Lowe B, Gordon M, Nester, EW (1985a) Limited-host-range plasmids of *Agrobacterium tumefaciens:* molecular and genetic analysis of transferred DNA. J Bacteriol 163:341–348

Yanofsky M, Lowe B, Montoya A, Rubin R, Krul W, Gordon MP, Nester EW (1985b) Molecular and genetic analysis of factors controlling host range in *Agrobacterium tumefaciens*. Mol Gen Genet 201:237–246