The *repABC* Plasmids with Quorum-Regulated Transfer Systems in Members of the Rhizobiales Divide into Two Structurally and Separately Evolving Groups

Margaret E. Wetzel[1], Gary J. Olsen[1,2], Vandana Chakravartty[1], and Stephen K. Farrand[1,*]

[1]Department of Microbiology, University of Illinois at Urbana-Champaign

[2]Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign

*Corresponding author: E-mail: stephenf@life.illinois.edu.

Accepted: November 13, 2015

Data deposition: All GenBank numbers for all plasmids included in this study are located in table 1.

Abstract

The large *repABC* plasmids of the order Rhizobiales with Class I quorum-regulated conjugative transfer systems often define the nature of the bacterium that harbors them. These otherwise diverse plasmids contain a core of highly conserved genes for replication and conjugation raising the question of their evolutionary relationships. In an analysis of 18 such plasmids these elements fall into two organizational classes, Group I and Group II, based on the sites at which cargo DNA is located. Cladograms constructed from proteins of the transfer and quorum-sensing components indicated that those of the Group I plasmids, while coevolving, have diverged from those coevolving proteins of the Group II plasmids. Moreover, within these groups the phylogenies of the proteins usually occupy similar, if not identical, tree topologies. Remarkably, such relationships were not seen among proteins of the replication system; although RepA and RepB coevolve, RepC does not. Nor do the replication proteins coevolve with the proteins of the transfer and quorum-sensing systems. Functional analysis was mostly consistent with phylogenies. TraR activated promoters from plasmids within its group, but not between groups and dimerized with TraR proteins from within but not between groups. However, *oriT* sequences, which are highly conserved, were processed by the transfer system of plasmids regardless of group. We conclude that these plasmids diverged into two classes based on the locations at which cargo DNA is inserted, that the quorum-sensing and transfer functions are coevolving within but not between the two groups, and that this divergent evolution extends to function.

Key words: plasmid evolution, RepABC, horizontal transfer, quorum-sensing, Alphaproteobacteria, conjugative transfer genes.

Introduction

Most of the large plasmids native to members of the Rhizobiales share in common a *repABC*-type replication system (Cevallos et al. 2008; Castillo-Ramirez et al. 2009; Pinto et al. 2012). This system is remarkable in that it can acquire, stably maintain, and vertically transfer large amounts of genetic information thereby conferring a large number and wide variety of functions to its host. The best described of these plasmids carry genes that are involved in plant–microbe interactions, the functions of which often confer the defining characteristic to the host bacterium. For example, the Ti (tumor inducing) and Ri (root inducing) plasmids of *Agrobacterium* spp. encode most of the virulence genes responsible for plant diseases correspondingly called crown gall and hairy root (reviewed in Tempé et al. 1984). In addition, Ti and Ri plasmids, as well as the Ao (*Agrobacterium* opine catabolic) and accessory plasmids of pathogenic and nonpathogenic isolates of *Agrobacterium* spp., may encode genes for ancillary traits including uptake and catabolism of opines, unique organic conjugates that are produced by crown gall tumors and hairy roots (reviewed in Dessaux et al. 1998). Other members of the Rhizobiales including species of *Rhizobium*, *Ensifer* and *Sinorhizobium* also harbor such *repABC* plasmids, the most recognizable of these being the Sym plasmids that confer nodulation and nitrogen fixation when the bacteria are in symbiosis with a suitable plant host (Galibert et al. 2001; Cevallos 2002). In some cases these large *repABC* plasmids have evolved or are evolving into second chromosomes, now called chromids (Harrison et al. 2010). For example, the 1.7-Mb *repABC* element in *Sinorhizobium*

meliloti and the 0.5-Mb *repABC* replicon in *Rhizobium etli* are chromids (Harrison et al. 2010; Landeta et al. 2011; diCenzo et al. 2013).

Many of these *repABC* family plasmids also encode a conjugative transfer system responsible for horizontal transfer of the plasmid among and between bacterial species. There are at least four classes of such transfer systems associated with the *repABC* plasmids (fig. 1*A*) (Giusti et al. 2012 and reviewed in Ding and Hynes 2009), two of which are well-characterized. The Class I system is composed of a chimeric IncQ- and IncP-like DNA metabolism (Dtr) and *oriT* region (Cook and Farrand 1992; Farrand et al. 1996a), the former of which is encoded by the *traAFBH* and *traCDG* operons. The mating pair formation (Mpf) system is composed of an IncP-like type IV secretion system (T4SS) (Li et al. 1998). The Class II system is composed of an IncQ-like *oriT* region and a Dtr system similar to that of the Class I system, but lacking the *traF*, *traB,* and *traH* genes (Chen et al. 2002). The Mpf system of Class II transfer systems, called *avhB,* is similar to the pathogenesis-associated VirB T4SSs of the Ti and Ri plasmids and of species of *Brucella* and *Bartonella* (Chen et al. 2002). The large genetic carrying capacity and relatively broad replication host range lend an overarching level of importance of these transfer systems to the evolution of traits carried by these plasmids, as well as the host bacteria that harbor them.

Classes I and II transfer systems have known regulatory mechanisms; the former being controlled by a quorum-sensing (QS) mechanism, whereas the latter is regulated by a *rctA*/*rctB*-like mechanism (reviewed in Ding and Hynes 2009). Although the nature of the inducing signal, if any, for Class II systems remains unknown, the genes and regulation of Class I transfer systems are well-studied.

Perhaps, the best-characterized Class I *repABC*-associated transfer systems are those of the Ti plasmids of *Agrobacterium tumefaciens*. These plasmids encode most of the *cis*- and *trans*-acting functions required by the bacterium to induce crown gall tumors on susceptible host plants. It has been known for more than 40 years that conjugative transfer of the Ti plasmids is highly regulated, and is strongly inducible by one or more of the opines produced by the crown gall tumors induced by the bacterium (reviewed in Dessaux et al. 1998; Farrand 1998). Thus, induction of the transfer system of these elements is intimately linked to the habitats resulting from the pathologies induced by the bacteria.

Although opines induce transfer of Ti, and some Ao and accessory plasmids, in all studied cases transcription of the genes of the Class I transfer systems of these plasmids is directly regulated by a LuxR-family QS system composed of the transcriptional activator TraR and an acyl-homoserine lactone (acyl-HSL) quormone (Piper et al. 1993; Zhang et al. 1993). The acyl-HSL is a population-dependent QS signal and is a product of TraI, the acyl-HSL synthase encoded by *traI*, the first gene of the plasmid *trb* operon (Hwang et al. 1994; Li et al. 1998). The acyl-HSL, in this case *N*-(3-oxooctanoyl)l-L-homoserine lactone (3-oxo-C8-HSL), is bound by TraR, where it promotes dimerization and stability of the activator (Zhu and Winans 1999, 2001; Qin et al. 2000). The dimerized form of TraR directly activates transcription of the *tra* and *trb* operons (Fuqua and Winans 1996; He et al. 2003).

One additional component, TraM, is common to these QS systems, and serves to inhibit premature activation of the *tra* regulon by basal levels of TraR when the appropriate signal is absent (Fuqua and Winans 1994; Fuqua et al. 1995; Hwang et al. 1995; Danino et al. 2003). TraM, an antiactivator, functions by binding to TraR, thereby inhibiting the transcription factor (Luo et al. 2000). In the Class I systems described to date, this effect is overcome and transfer is induced by an increase in the transcription of *traR* in response to some specific external signal. Such signals include opines produced by the plant neoplasias induced by pathogenic *Agrobacterium* spp. (Hwang et al. 1995; Piper and Farrand 2000; Oger and Farrand 2002) and in the case of *Rhizobium leguminosaurum*, an orphan LuxR homolog that responds to an acyl-HSL produced by an appropriate recipient (Danino et al. 2003).

Although the *tra*, *trb*, *rep,* and QS genes are conserved among the Class I group of these large plasmids in the Rhizobiales, as first noted by Moriguchi et al. (2001), the organization of these genes and operons can be divided into two categories. Here, we denote these categories as Group I and Group II (fig. 1*A*). In both groups, the *traI*/*trb* operon is invariably adjacent and divergently oriented to the canonical *repABC* operon. In Group I plasmids, a locus encoding *traR*, *traM*, the two divergently oriented *tra* operons, and the *cis*-acting *oriT* is separated, often by more than 60 kb, from the *traI*/*trb* genes (fig. 1*B*, left side and supplementary fig. S1, left side, Supplementary Material online). Moreover, *traR* generally is located in an operon the expression of which can be regulated by a specific external signal (fig. 1*B*, left side and supplementary fig. S1, left side, Supplementary Material online). In the Group II plasmids the *traI*/*trb* operon again is divergently linked to *repABC*, but the *tra* locus is contiguous to the *trb* region with *traM* and *traR* located between these two components (fig. 1*B*, right side and supplementary fig. S1, right side, Supplementary Material online). Moreover, unlike the Group I plasmids, *traR* appears to be monocistronic in the Group II plasmids.

That these groups of otherwise conserved genes are organized in two different patterns raises the question of whether the component gene systems represent divergent evolutionary lineages, and if so, whether within a lineage, the genes are coevolving in a given plasmid or group of plasmids. In this study, we analyzed the evolutionary and functional relationships of select genes of these core systems

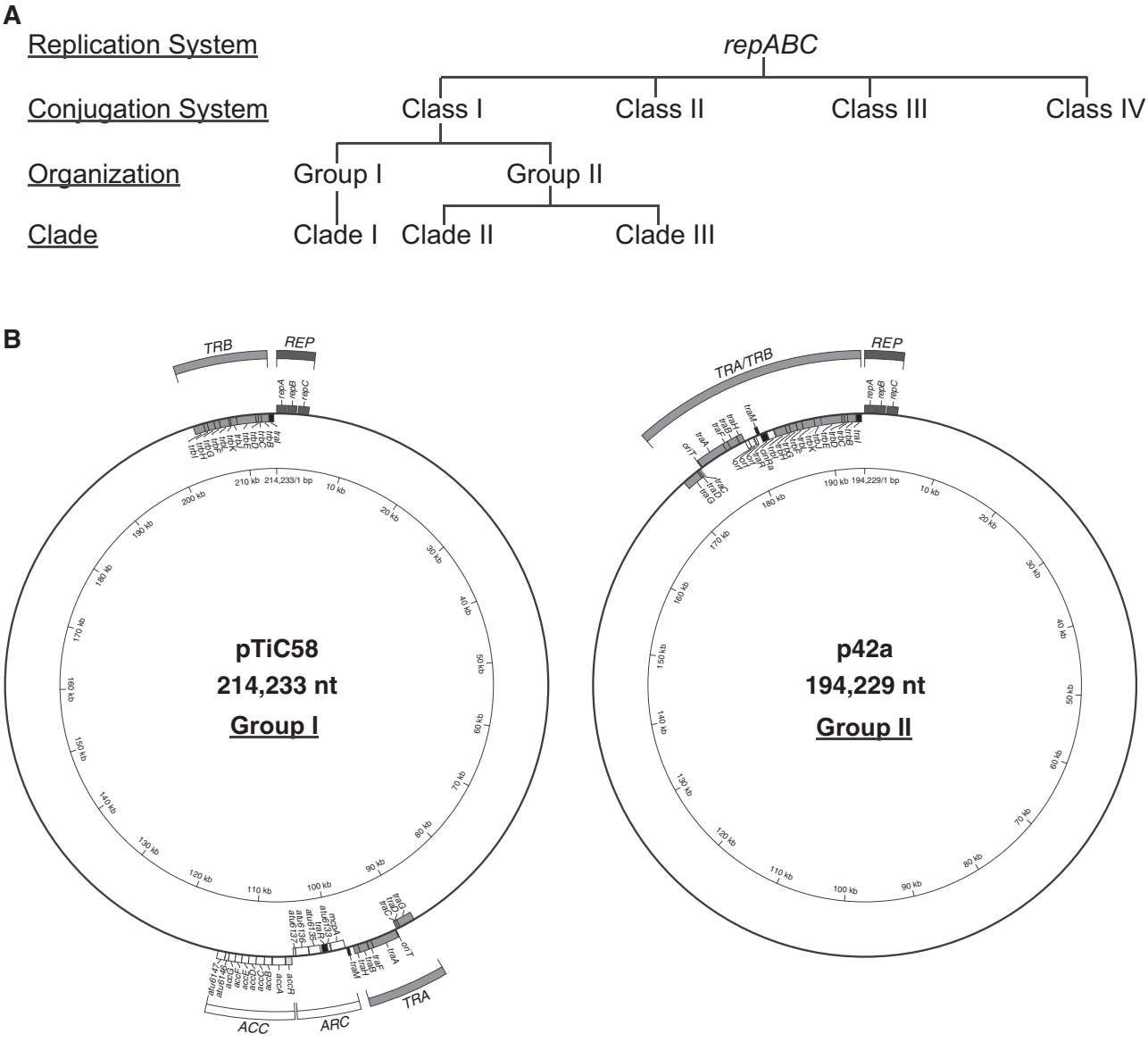

Fig. 1.—The *repABC* plasmids with Class I conjugative transfer systems in *Agrobacteria* and *Rhizobium* divide into two organizational groups. (*A*) Flow chart categorizing *repABC* plasmids with Class I conjugative transfer systems. Plasmids with Class I transfer and regulatory genes have two distinct organizations of the genes involved in transfer and QS regulation, which can be further divided into three evolutionary clades. (*B*) Examples of the two organizational groups. Left: pTiC58 from *Agrobacterium tumefaciens* strain C58, a representative of plasmids with Group I organization. Right: p42a from *Rhizobium etli* strain CFN 42, a representative of plasmids with Group II organization. The *traI*, *traR* and *traM* genes are in black, the *repABC* genes are in dark gray, and the *tra* and *trb* genes are in medium gray. On pTiC58, genes involved in regulation of transfer of pTiC58 include *accR* in light gray, and the genes for catabolism of agrocinopines A+B, the conjugative opine, in white.

encoded by 18 plasmids from members of the families Rhizobiaceae and Bradyrhizobiaceae. Here we report that, based on amino acid sequence comparisons, the QS and transfer proteins belonging to plasmids within Group I cluster together but separately from the orthologous proteins encoded by Group II plasmids. Additionally, in any given plasmid the QS and transfer proteins appear to evolve together, but separately from their adjacent Rep proteins. The *cis*-acting *oriT* sequences are highly conserved among all of the analyzed plasmids, and unlike the proteins, they do not neatly separate into two major clades. Functionally, TraR activates *tra* box-containing promoters within, but not between the two major clades. On the other hand, recombinant plasmids containing different *oriT* regions can be mobilized by plasmids from either group, but the efficiency of transfer is dependent upon how related the *oriT* is to the *oriT* that is cognate to the *trans*-acting *tra* system.

Materials and Methods

Strains Media and Growth Conditions

Bacterial strains and plasmids used in this study are listed in supplementary table S1, Supplementary Material online. Cultures of *A. tumefaciens* strain NTL4 (Luo et al. 2001) and its derivatives were grown with shaking at 30 °C in liquid MG/L (Cangelosi et al. 1991) or in AB minimal medium (Cangelosi et al. 1991) supplemented to 0.2% with mannitol as the sole source of carbon (ABM) and with 0.005% yeast extract or on solid 2% agar medium using either Nutrient both (Difco) or ABM media at 28 °C. Strains of *Escherichia coli* were grown in L broth (Fischer Scientific) at 30 or 37 °C. When required for selection, antibiotics were added at the following concentrations (μg/ml): Ampicillin, 100; carbenecillin, 50 or 100; gentamicin, 25; kanamycin, 25 or 50; rifampicin, 50; spectinomycin, 50 or 100; streptomycin, 50 or 100; and tetracycline, 5 or 10. X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was used at a concentration of 40 μg/ml, whereas isopropyl-β-D-thiogalactopyranoside (IPTG) was used at 1 mM unless otherwise stated.

Alignments and Phylogenies

The GenBank accession numbers and information concerning the parental strains for the 18 plasmids investigated in this study are presented in table 1. Plasmid pTiBo542 contains two complete copies of *repABC* in tandem direct repeat. We used the protein sequences of *repA*, *repB* and *repC* of the first repeat, which is directly linked to the *tral/trb* operon. Plasmid pAtK84b encodes two opine-inducible copies of *traR*. The first is inducible by agrocinopines A+B ($traR_{acc}$) and is located in close proximity to *traM* and the two *tra* operons. The second copy of *traR* is located near the nopaline catabolic operon and is inducible by nopaline ($traR_{noc}$) (Oger and Farrand 2002). We used the sequence encoded by *traR* that is most closely linked with the Dtr system ($traR_{acc}$). All protein and nucleotide sequences used in this study (supplementary table S2, Supplementary Material online) were annotated by hand. The TrbK protein sequence of pRi1724 appears to have a premature stop codon with a conserved downstream sequence of DNA. For this study, this stop codon was annotated as unknown amino acid, X, and the downstream sequence was translated and included in the protein sequence. pNGR234a is annotated as having two consecutive reading frames containing portions of TrbE. Further analysis of the nucleotide sequence suggested to us that this was due to the addition of an extra nucleotide which changed the reading frame of this protein. For the purposes of this study, we removed the extra nucleotide and reannotated $TrbE_{pNGR234a}$. Protein sequences were aligned using three programs, MAFFT (Katoh and Toh 2008), ClustalW (Thompson et al. 1994), and Muscle (Edgar 2004a, 2004b). As MAFFT gave alignments with the best likelihood scores, all sequence alignments were conducted using this program. We tested five phylogeny programs in MEGA 6.06 (Tamura et al. 2013) to construct trees for each MAFFT alignment: The Neighbor-Joining (NJ) method (Saitou and Nei 1987) using the bootstrap test (Felsenstein 1985) with 1,000 replicates, the Minimum Evolution method (Rzhetsky and Nei 1992) using 1,000 replicates of the bootstrap test (Felsenstein 1985), and the UPGMA method (Sneath and Sokal 1973). These three programs used the Poisson (Zuckerkandl and Pauling 1965) and γ correction implemented in MEGA 6.06. The maximum likelihood method using a Poisson correction (Zuckerkandl and Pauling 1965) and the maximum parsimony method, which uses the Subtree-Pruning–Regrafting algorithm (Nei and Kumar 2000), also were assessed using MEGA 6.06. The majority of the trees was either identical or highly similar and so only the trees constructed using the NJ method are shown.

Nucleic acid sequences were aligned using MAFFT and were visualized using the TEXshade (Beitz 2000) program in the SDSC biology workbench (http://workbench.sdsc.edu, last accessed December 1, 2015) alignment suite (Subramaniam 1998). The *tra* box and *oriT* trees were constructed using the NJ method with 1,000 bootstrap replicates implemented in MEGA 6.06 (Felsenstein 1985; Saitou and Nei 1987; Tamura et al. 2013).

Cloning

All polymerase chain reaction (PCR) reactions performed for cloning purposes used either Pfu DNA polymerase (Promega) or Phusion DNA polymerase (NEB). $traR_{pTiC58}$ and $traR_{pAoF64/95}$ were cloned into the pBBR1MCS derivative pSRKGm as previously described (Khan et al. 2008; Wetzel et al. 2014). $traR_{pRi1724}$ was amplified by PCR using the following primers: traRpRi1724-F (5′-GCCGAATTCATATGGACGGTGACTTTCGTTCT-3′) and traRpRi1724-R (5′-CGCAAGCTTTCAAACCAAGCCGTGATCTTTAGCG-3′). We directionally cloned the PCR product into pBBR1MCS-derived vector pZLQ (Luo and Farrand 1999) using the *Nde*I and *Hin*dIII sites underlined in the primer sequences. To construct the *traA::lacZ* fusions and the *oriT* mobilization vectors, the *traA-C* intergenic region containing *tra* box I and the *oriT* sequence from each plasmid tested was amplified by PCR. The *traA-C* intergenic region of pAoF64/95 was amplified using primers traCpAoF64XmaI (5′-CAGATAACCCGGGATCGTCTCCTGGGTGAGAAAG-3′) and traApAoF64BamHI (5′-CGAGTCCGGATCCGGTTGCGAACAATATCAAAGGG-3′), whereas the *traA-C* intergenic region from pRi1724 was amplified using the primers traCpRi1724XmaI (5′-CGCCCCGGGTCCGTCTGTCTCCTTGGGTG-3′) and traApRi1724BamHI (5′-CGCGGATCCGGTTGCAAACGAAATCAATG-3′). These products were subsequently directionally cloned into pRG970b (Van den Eede et al. 1992) using *Xma*I and *Bam*HI (underlined in the primer sequence), such that the *traA* promoter was transcriptionally fused to *lacZ*. pZLb251 which contains the *traA-C* intergenic

Table 1
Descriptions and GenBank Accession Numbers for Plasmids or Contigs Used in This Study[a]

Genus/Species	Strain(s)	Plasmid	Type	Group[b]	GenBank Number
Agrobacterium radiobacter	F64/95	pAoF64/95[c]	OC	II	JX683454.1
Agrobacterium radiobacter	K84	pAtK84b[c]	OC	I	CP000630.1
Agrobacterium rhizogenes	A4	pRiA4b (replication)	VR	II	X04833.1
Agrobacterium rhizogenes	A4	pRiA4b (transfer)	VR	II	AB050904.1
Agrobacterium rhizogenes	K599	pRi2659	VR	II	EU186381.1
Agrobacterium rhizogenes	MAFF03-01724	pRi1724	VR	II	AP002086.1
Agrobacterium tumefaciens	Bo542	pTiBo542[c]	VT	I	DQ058764.1
Agrobacterium tumefaciens	Composite	pTiOctopine[c]	VT	I	AF242881.1[d]
Agrobacterium tumefaciens	C58	pTiC58[c]	VT	I	AE007871.2
Agrobacterium tumefaciens	MAFF 301001	pTi-SAKURA	VT	I	AB016260.1
Agrobacterium vitis	S4	pTiS4[c]	VT	I	CP000637.1
Agrobacterium vitis	S4	pAtS4c	TC	I	CP000636.1
Ensifer adhaerens	OV14	pOV14c	CR	I	CP007238.1
Nitrobacter hamburgensis	X14	pB11	CR	II	CP000322.1
Oligotropha carboxidovorans	OM5	pHCG3	CC	II	CP002827.1
Rhizobium etli	CFN 42	p42a[c]	CR	II	CP000134.1
Rhizobium leguminosaurum bv. *trifolii*	CB782	pCB782	SM	I	CP007070.1
Sinorhizobium fredii	GR64	pSfr64a[c]	CR	II	CP002245.1
Sinorhizobium fredii	NGR234	pNGR234a	SM	II	U00090.2

NOTE.—CC, carbon monoxide utilization; CR, cryptic; OC, opine catabolism; SM, symbiosis; TC, tartrate utilization; VR, virulence-rhizogenic; VT, virulence-tumorigenic.
[a]The plasmids were originally identified in the species and strain indicated.
[b]Based on the organization of the *rep*, *tra,* and *trb* genes as described in the text.
[c]Plasmids experimentally known to be self-conjugative.
[d]The sequence of the octopine-type Ti plasmid is an assembly of sequences from several virtually identical Ti plasmids including pTiR10, pTi15955, pTiA6NC, pTiAch5, and pTiB6S3 (Zhu et al. 2000).

region from pTiC58 cloned into pRG970b is described elsewhere (Luo and Farrand 1999).

Construction of Mutant Strains

All PCR reactions used to construct mutant strains were carried out with either Pfu DNA polymerase (Promega) or Taq DNA polymerase (NEB). In-frame deletion mutants of *traM* and *mrtR* on pAoF64/95 were constructed using the method of Datsenko and Wanner (2000). Briefly, the kanamycin-resistance cassette of pKD4 was amplified using the following primers which contained 5′-overhang sequences for *traM*: Forward primer, 5′-CTTGAGCGTGGGGTTTTCGAAAA AAGGGAGGAGAATGGTGTGTAGGCTGGAGCTGCTTCG-3′, reverse primer, 5′-CCTCGTCGCGATCGCCAAGGACCAC GGCCTGCTGTAGCGCATATGAATATCCTCCTTAGT-3′. The PCR product was transformed into *E. coli* (pKD46) carrying a cosmid clone of the appropriate region of pAoF64/95 (Wetzel et al. 2014). λ red-mediated replacement of *traM* with the kanamycin resistance cassette was confirmed in the cosmid by PCR analysis using the traMcheckdown (5′-CTATG ATGTTGACGTTTGCATCTT-3′) and traMcheckup (5′-GATCG CCATGACCTCTTTGA-3′) primers. The mutant allele of *traM* was marker exchanged into pAoF64/95 in strain NTL4 as described previously (Wetzel et al. 2014). The same method was used to construct the indel mutation in *mrtR*. The kanamycin cassette of pKD4 was amplified with the following primers: Forward, 5′-TTGGACACCGAGCCAATGT ACATCCAACTGCAGCAAGATGTGTAGGCTGGAGCTGCTT CG-3′ and reverse, 5′-CTAGCCACCTTTGTGTGGTGTCCTAC CGCCCATCATCATCATATGAATATCCTCCTTAGT-3′. The λ red-mediated *mrtR* mutation in the cosmid was confirmed by PCR using the mrtRcheckup primer (5′-GCCCTCCGCTCC CAGTTAAA-3′) and the mrtRcheckdwn primer (5′-AGCGGC TACAATCTTCCTTG-3′), and the mutant allele of *mrtR* was marker exchanged into pAoF64/95 as described previously (Wetzel et al. 2014).

β-Galctosidase Assays

The $traA_{pTiC58}$::*lacZ*, $traA_{pAoF64/95}$::*lacZ* and $traA_{pRi1724}$::*lacZ* reporter vectors described above were electroporated into *A. tumefaciens* strain NTL4. Strain NTL4 harboring the empty vector pRG970b (Van den Eede et al. 1992) also was constructed as a control. We subsequently electroporated pSRKGm, pSRKGm::$traR_{pTiC58}$, pSRKGm::$traR_{pAoF64/95}$, or pZLQ::$traR_{pRi1724}$ individually into each of the four *traA*::*lacZ* reporter strains. The resulting 16 strains were assessed for β-galactosidase activity on solid medium. Briefly, single colonies of each strain were individually suspended in 250 µl of 0.9% NaCl and 5 µl volumes of each suspension were spotted onto ABM media containing X-gal, spectinomycin, and IPTG to induce expression of TraR. Strains were tested on medium both with and without 50 nM 3-oxo-C8-HSL, *Agrobacterium* autoinducer (AAI, Sigma-Aldrich). β-galactosidase activity was assessed visually after 48 h of incubation at 28 °C.

Mobilization Experiments

Empty vector pRG970b (Van den Eede et al. 1992), or pRG970b containing the *traA-C/oriT* region from pAoF64/95, pTiC58 or pRi1724 was electroporated into strains NTL4(pTiC58Δ*accR*) (Ellis, Murphy, et al. 1982) and NTL4(pAoF64/95Δ*mrtR*) (Wetzel ME and Farrand SK, unpublished data, see above). The *repABC* plasmids in these strains are constitutive for conjugative transfer (tra^C). The resulting tra^C donor strains carrying the *oriT* vectors and the recipient strain, C58C1RS, were grown in liquid MG/L medium overnight. Filter matings were conducted as described previously (Farrand et al. 2002). Briefly, donor and recipient strains were mixed together in a 10:1 v/v ratio and cells in 50 µl volumes of each mixture were collected by vacuum filtration onto a sterile 0.22-µm filter disc. The filter was then placed bacterial side up onto solid media and incubated at 28 °C for 22–24 h. The cells were collected by vortexing each disk in a 1 ml volume of 0.9% NaCl. Volumes of 100 µl of a decade dilution series of each mating then were plated onto solid medium supplemented with rifampicin, streptomycin, carbenecillin, and spectinomycin. Rifampicin and streptomycin select for the recipient strain, whereas carbenecillin and spectinomycin select for the mobilized *oriT* plasmid. Colonies of transconjugates were enumerated after 5–6 days of incubation at 28 °C. Frequencies of transfer are expressed as transconjugates per input donor (Farrand et al. 2002).

TrlR–TraR Cross-Dimerization

To assess the ability of TraR to cross-dimerize with orthologous proteins, we utilized pPOKKTrlRA (Oger et al. 1998), a vector expressing a cloned copy of *trlR* from pTi15955, or pKK38 as an empty vector control. These plasmids were electroporated into *A. tumefaciens* strain NTL4 harboring the tra^C *traM* deletion derivatives of pTiC58 (Hwang et al. 1995), pTiR10 (Fuqua et al. 1995) and pAoF64/95, all of which confer resistance to kanamycin. If TrlR cross-dimerizes with the TraR cognate to the system, then transfer frequencies of these tra^C plasmids should decrease. The resulting strains were assessed for conjugative transfer frequency using the drop-plate mating technique described previously (Farrand et al. 2002; Wetzel et al. 2014). Transconjugates were selected on medium containing rifampicin, streptomycin, and kanamycin. Transfer frequencies are expressed as the number of transconjugates obtained per input donor.

Analysis of Protein Coevolution

Analysis of trees constructed from the RepA, RepB, RepC, TrbE, TrbK, TraI, TraR, TraM, TraG, and TraA protein alignments compared with the MAFFT alignments of each protein was accomplished by using the likelihood scores. We used the formula:

$$(\mathrm{LnLikelihood(Data|GivenTree)} - \mathrm{LnLikelihood(Data|RandomTree)}) / (\mathrm{LnLikelihood(Data|OwnTree)} - \mathrm{LnLikelihood(Data|RandomTree)}).$$

This formula essentially compares the fit of the sequence data in an alignment to the trees constructed from different protein alignments, with each adjusted for the component of the score attributable to random similarity of trees. Proteins that are coevolving have a score close to 1.00, whereas proteins that are not coevolving have scores closer to 0.00 (negative values are possible because a tree can be worse than random for the given data). The data were entered into a program written to display output values on a gray scale heat map, with values closer to 1.00 being white and values closer to 0.00 being black.

We noticed that using relatedness values from proteins from plasmids that are highly similar skewed the entire data set toward coevolution as proteins from those plasmids will always group closely together. We therefore excluded results from analysis of pTi-SAKURA, which is highly similar to pTiC58 (Suzuki et al. 2000; Goodner et al. 2001), and pRi1724, which is highly similar to pRi2659 (Moriguchi et al. 2001).

Results

The Two Distinct Organizations of the Genes Involved in Conjugative Transfer and Its Regulation Correspond to Plasmid Type

Structurally, all 18 of the Class I-type *repABC* plasmids studied divide into two distinct genetic organizations, Group I and Group II (see fig. 1*B*, table 1, and supplementary fig. S1, Supplementary Material online). Common to all of the plasmids examined, the *trai/trb* operon is invariably adjacent to the divergently oriented *repABC* operon. However, in the Group I plasmids such as pTiC58, pTiBo542, pTiOctopine, and pCB782 (fig. 1*B*, left side, supplementary fig. S1, left side, Supplementary Material online, and table 1) as well as in at least one opine catabolic plasmid (pAtK84, table 1) the two *tra* operons, along with *traR* and *traM,* are located near the region encoding catabolism of the conjugative opine or some unknown compound and are separated by a large but variable distance from the *trai/trb* region. In the Group II plasmids, such as the auxiliary plasmid p42a of *R. etli* (fig. 1*B*, right side and table 1), pAoF64/95, pRi1724, pNGR234a (supplementary fig. S1, right side, Supplementary Material online, and table 1), two sinorhizobial plasmids, pSfr64a and pNGR234a, and the two bradyrhizobial plasmids, pHCG3 and pB11 (table 1), the *tra* and *trb* regions are contiguous, and *traR* and *traM* are located between the *trai/trb* operon and the last gene of the *traAFBH* operon. Moreover, unlike in the Group I plasmids where *traR* generally is a member of an

operon, in the Group II plasmids examined to date *traR* is monocistronic (fig. 1*B*, right side and supplementary fig. S1, right side, Supplementary Material online). Remarkably, all of the Ti plasmids characterized to date fall into Group I, whereas all of the characterized Ri plasmids fall into Group II (table 1). Two of the Class I-type plasmids from species of *Rhizobium*, *Ensifer* and *Sinorhizobium* (pOV14c and pCB782) fall into Group I, whereas three (p42a, pSfr64a, and pNGR234a) fall into Group II (table 1). The two bradyrhizobial plasmids (pHCG3 and pB11) fall into Group II (table 1).

Proteins of the QS Systems of Plasmids from Members of the Rhizobiales Form Two Major Clades that Correspond to Plasmid Organization

The observation that *traR* is organized either as a member of an operon, as seen in the Group I plasmids, or monocistronic, as in Group II plasmids, along with the bimodal organization of the *tra–trb* regulon (fig. 1*B* and supplementary fig. S1, Supplementary Material online), led us to hypothesize that TraR proteins that regulate conjugative transfer could be divergently evolving between the two groups of plasmids. In addition, considering that TraM interacts with TraR, and that the acyl-HSL produced by TraI is the ligand of TraR, we postulated that these three QS proteins would evolve together. To test these two hypotheses, we assessed amino acid sequence relatedness of the TraR, TraM and TraI proteins derived from the Class I plasmids described in table 1.

Figure 2*A–C* shows the phylogenetic trees of the three QS proteins for the 18 plasmids examined, all determined as described in Materials and Methods. The trees built from these three proteins overall are topologically consistent. Sixteen of the protein sets divide into two major clades whereas two, those from the plasmids of *Oligotropha carboxidovorans* and *Nitrobacter hamburgensis*, divide as distant relatives to both of the two major clades. Whether either of these plasmids is conjugative remains to be determined. The members of the two major clades divide precisely by plasmid organization. The three proteins from Group I plasmids, in which *traR* is polycistronic and where the *tra* and *trb* regions are physically separate on the plasmid, form one clade whereas the three proteins from Group II plasmids, in which *traR* is monocistronic and the *tra* and *trb* regions are adjacent to one another, group together as the second clade.

The *Tra* Box Sequence, while Having a Small Core Set of Nucleotides in Common, Divides into Two Distinct Conserved Sequences: Group I- and Group II-Like

TraR binds to an 18-bp inverted repeat sequence, called the *tra* box, from which it activates transcription from the promoters of the associated operons (Luo and Farrand 1999; Zhu and Winans 1999). The Class I rhizobial and Ti plasmids typically contain between two and four such boxes (Fuqua et al. 1995; Fuqua and Winans 1996; Danino et al. 2003; He et al. 2003; White and Winans 2007). *tra* box I, located in the intergenic region between and controlling transcription of the divergently oriented *traAFBH* and *traCDG* operons, is the most highly conserved of these sequences (reviewed in White and Winans 2007). *tra* box II is located upstream of and controls transcription of the *traI/trb* operon, whereas *tra* box III, when present, is located just upstream of *tra* box II and contributes to the control of transcription of the divergently oriented *repABC* operon (Li and Farrand 2000; Pappas and Winans 2003). A fourth and less conserved *tra* box, *tra* box IV, was described in the Ti plasmids and is located in the promoter region of *traM* (Fuqua et al. 1995). We aligned the nucleotide sequences of the putative *tra* box I from each of the 18 plasmids examined in this study.

Like TraR, TraM and TraI, the putative *tra* box I sequences divide into two distinct sets of sequences corresponding to the Group I and Group II clades (fig. 2*D* and *E*). The *tra* box I sequences from six of the Group II plasmids are virtually identical with one sequence, that from p42a, differing by a single nucleotide and that from pSfr64a differing at eight nucleotides (fig. 2*E*). On the other hand, the sequences of the Group I clade subdivide into groups that represent changes to five separately located bases (fig. 2*D* and *E*). The Group I *tra* box I consensus sequence and the Group II *tra* box I consensus sequence share in common only four fully conserved bases, whereas three additional bases are conserved in the *tra* box I sequences of most of the Rhizobial plasmids examined (fig. 2*F*). Additionally, the Group I sequences form a considerably more perfect inverted repeat in comparison to the Group II sequences (fig. 2*E* and *F*). Consistent with the divergence of their TraR proteins, the putative *tra* box I sequences of pB11 from *N. hamburgensis* and pHCG3 from *O. carboxidovorans* are more distantly related to the boxes of Group I and Group II plasmids, and most closely related to each other (fig. 2*D*).

TraR Activates Transcription from a *tra* Box-Dependent Promoter among Members of the Same Clade, but Not between Members of the Other Clade

Given the division of TraR and its DNA binding site into two major groups, we tested the ability of activators of each of the two major clades to induce transcription from cognate and noncognate *tra* boxes. This was accomplished by assessing activation of cognate and noncognate promoters from the *traAFBH* operon by alleles of *traR* from different sources. We cloned *traR* from pTiC58 ($traR_{pTiC58}$), pAoF64/95 ($traR_{pAoF64/95}$), and pRi1724 ($traR_{pRi1724}$) into pBBR1MCS-derived vectors (Luo and Farrand 1999; Khan et al. 2008) and constructed *ptraA*::*lacZ* transcriptional fusions that contain the entire *traA–traC* intergenic regions, including the *tra* box I sequences, from pTiC58 ($ptraA_{pTiC58}$), pAoF64/95 ($ptraA_{pAoF64/95}$), and pRi1724 ($ptraA_{pRi1724}$), all as described in Materials and Methods. The three TraR orthologs were tested for their ability to activate transcription of the three *traAFBH* promoters as

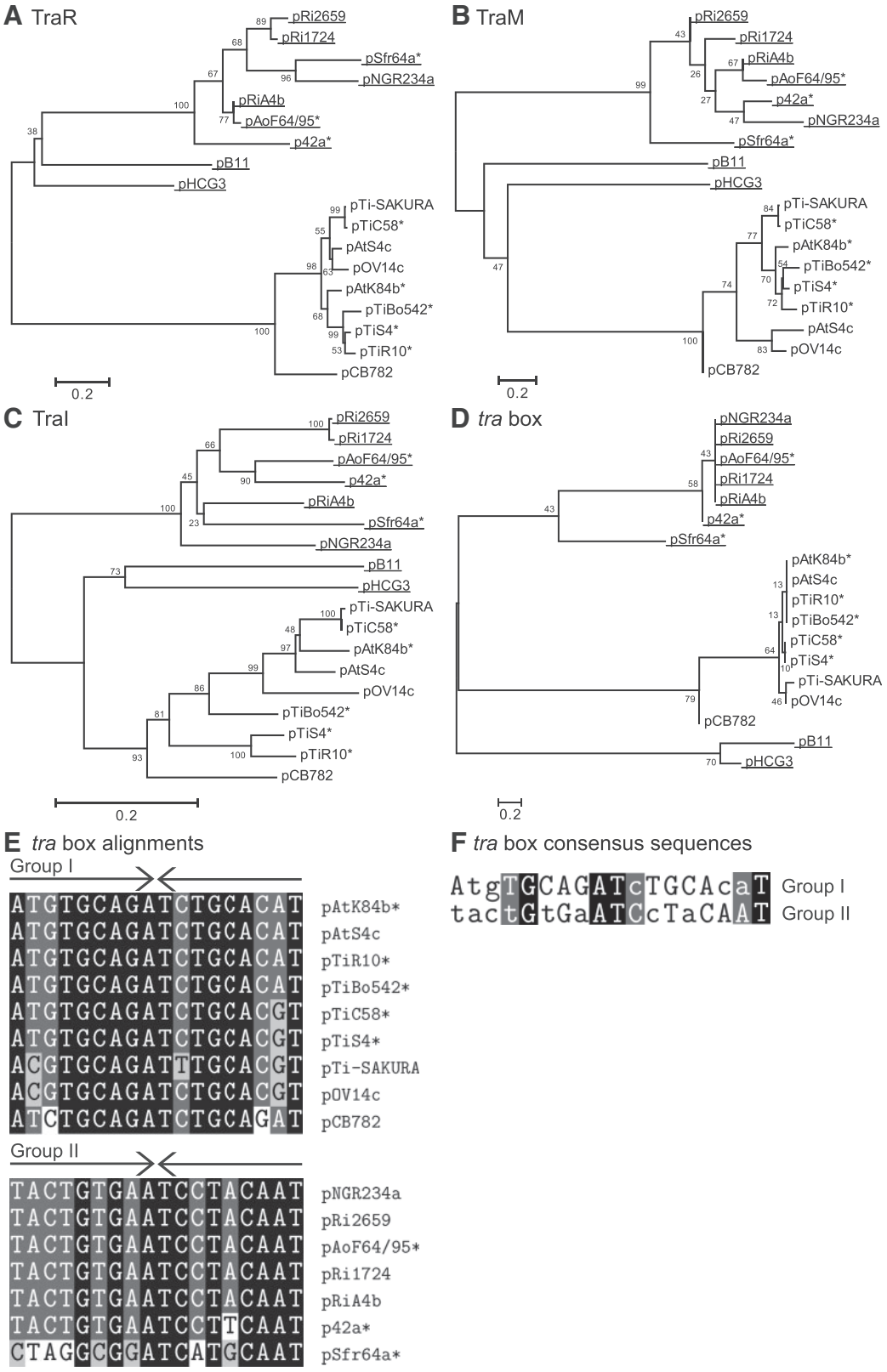

Fig. 2.—TraR, TraI and TraM, and the *tra* box I sequences divide into three clades. NJ trees constructed from MAFFT alignments for (*A*) TraR, (*B*) TraM, and (*C*). TraI proteins, and (*D*) *tra* box I nucleotide sequences from 18 plasmids in the Rhizobiales. Bootstrap values from 1,000 replicates are located next to

assessed by β-galactosidase activity in cultures grown with and without 3-oxo-C8-HSL (AAI).

In cells grown in the absence of the quormone, $TraR_{pTiC58}$ minimally activated only its cognate promoter (fig. 3*A*). When AAI was added to the medium, $TraR_{pTiC58}$ strongly activated only $ptraA_{pTiC58}$ (fig. 3*B*). No significant levels of β-galactosidase activity were detected in any growth condition from strains in which $TraR_{pTiC58}$ was paired with the *traA* promoter-reporters from pAoF64/95 or from pRi1724 (fig. 3*A* and *B*). In cells grown in the absence of AAI, $TraR_{pAoF64/95}$ failed to detectably activate expression of β-galactosidase from any of the three *ptraA* promoters (fig. 3*A*). However, when grown with the acyl-HSL, the strain expressing $TraR_{pAoF64/95}$ strongly activated its cognate promoter and to a lesser extent $ptraA_{pRi1724}$ (fig. 3*B*). $TraR_{pAoF64/95}$ also activated $ptraA_{pTiC58}$ in an AAI-dependent manner, but at only a barely detectible level (fig. 3*B*). Unexpectedly, when grown with AAI, $TraR_{pRi1724}$ did not activate any of the reporters, including its cognate promoter, $ptraA_{pRi1724}$ (fig. 3*B*). However, in strains grown without AAI, $TraR_{pRi1724}$ activated $ptraA_{pAoF64/95}$, and to a lesser extent its cognate promoter, and very minimally $ptraA_{pTiC58}$ (fig. 3*A*).

TraR Exhibits Dimerization Specificity Delineated by the Group I- and Group II-Like Clade Structure

The N-terminal region of TraR contains both an quormone-binding domain and the primary dimerization domain (Vannini et al. 2002; Zhang et al. 2002). We assessed whether the proteins from different clades can cross-dimerize by using TrlR, a dominant-negative mutant of TraR encoded by the octopine-type Ti plasmids pTi15955 and pTiR10 (Oger et al. 1998; Zhu and Winans 1998). TrlR has functional ligand-binding and dimerization domains, but lacks the C-terminal DNA-binding domain (Oger et al. 1998; Zhu and Winans 1998; Chai et al. 2001). When coexpressed TrlR strongly inhibits the activity of $TraR_{Octopine}$ by forming inactive heterodimers (Oger et al. 1998; Zhu and Winans 1998; Chai et al. 2001).

We assessed the ability of different TraR proteins to cross-dimerize indirectly by evaluating conjugative transfer frequencies of derivatives of *A. tumefaciens* strain NTL4 expressing both $TrlR_{pTi15955}$ and TraR of the native plasmid. We utilized transfer-constitutive (tra^C) strains (see supplementary table S1, Supplementary Material online) harboring either pTiC58Δ*traM* or pTiR10Δ*traM* as representatives of the Group I plasmids and pAoF64/95Δ*traM* as a representative of the Group II plasmids. We constructed these three strains to additionally carry a plasmid with either a cloned wild-type copy of *trlR* (pPOKK*trlR*A) or the empty vector, pKK38 (Oger et al. 1998). If transfer frequencies decrease in the tra^C strain harboring pPOKK*trlR*A in comparison to the tra^C strain harboring pKK38, then TrlR is functionally dimerizing with the native TraR thereby inhibiting activation of the *tra* regulon.

In the donor harboring pTiR10Δ*traM*, a Group I plasmid essentially identical to pTi15955, expression of TrlR inhibited the transfer frequencies of the megaplamid by about 5,500-fold in comparison to the same tra^C strain harboring the empty vector (table 2). Transfer frequency of the donor harboring pTiC58Δ*traM*, another Group I plasmid, decreased by approximately 140-fold when TrlR was expressed in comparison to the same strain without TrlR (table 2). In donors harboring the Group II plasmid pAoF64/95Δ*traM* expression of TrlR had only a negligible effect on the transfer frequency of the opine-catabolic plasmid when compared with the same strain without TrlR (table 2).

Structural Genes of the Conjugative Transfer System Divide into Group I and Group II Clades

To assess whether, like the QS proteins, the proteins of the conjugative transfer systems of Group I and II plasmids have divergently evolved we included in our study sequence comparisons of the products of three essential, and one nonessential transfer genes, two from the DNA metabolism (Dtr) operons and two from the Mpf operon (supplementary table S2, Supplementary Material online). TraA, an essential representative of the Dtr system, is the strand transferase responsible for nicking the *oriT* and is a component of the relaxosome (Cho and Winans 2007). The protein contains two conserved motifs, a tyrosine-type site-specific recombinase domain and a helicase domain (Farrand et al. 1996b). The second protein, TraG, is the coupling factor that is believed to interface the relaxosome with the T4SS (Hamilton et al. 2000). TraG, which is essential, is distantly related to VirD4, a component involved in plant virulence, and contains two potential nucleotide-binding domains (Farrand et al. 1996b; Hamilton et al. 2000; Cho and Winans 2007). TrbE, an essential representative of the Mpf system, contains a Walker type-A nucleoside triphosphate binding domain and is distantly related to VirB4 of the T4SS involved in plant virulence (Li et al. 1998, 1999), whereas TrbK, which is not essential for transfer, is a short protein and is involved in entry exclusion (Cho et al. 2009). TraA, TraG, and TrbE are large multidomain proteins making them excellent representatives

Fig. 2.—Continued

the branches. The evolutionary distances were computed with the Poisson correction. The asterisks (*) denote plasmids that are known to be conjugative, whereas the underscores indicate plasmids with a Group II organization as described in the text. (*E*) Alignment of *tra* box I nucleotide sequences from the 16 Group I and Group II plasmids. (*F*) The consensus sequences of *tra* box I from Group I and Group II plasmids, excluding the two plasmids from Clade III, pB11, and pHCG3. Capital letters in the consensus sequences indicate invariant bases, whereas lower case letters are conserved nucleotides. The black color indicates invariant bases and gray indicates conserved bases.

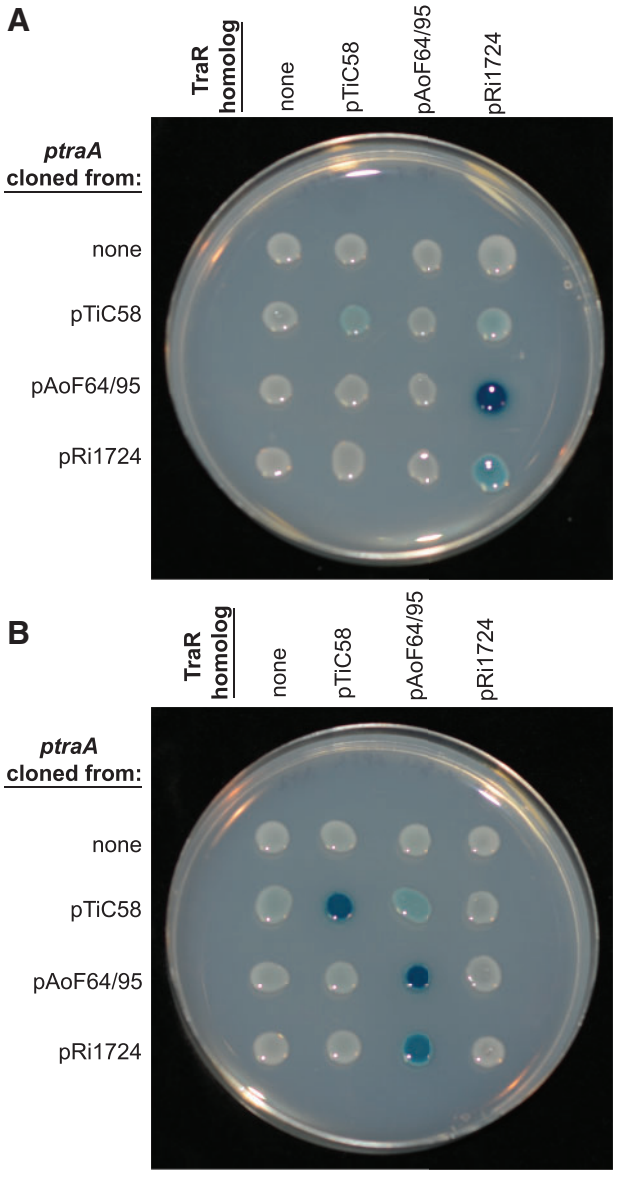

Fig. 3.—TraR activates transcription of the *traAFBH* promoters within but not between clades. Derivatives of strain NTL4 carrying a *ptraA::lacZ* transcriptional fusion from one of three plasmids (labeled on the vertical axis) and one of three TraR orthologs (labeled on the horizontal axis) were assessed for β-galactosidase activity on solid ABM media supplemented with X-gal and IPTG to induce expression of *traR*. The growth medium in plate (*B*) is additionally supplemented with AAI at 50 nM.

of the transfer system. Like those of the QS proteins, trees of TraA, TraG and TrbE divide into two major clades, precisely corresponding to the Group I and Group II organizational patterns (fig. 4*A*–*C*). Each of the three essential transfer proteins from *Oligotropha* and *Nitrobacter* are related to each other, but form a distinct third clade (fig. 4*A*–*C*). TrbK divides into two major clades that are mostly consistent with the two organizational types (fig. 4*D*). However, two TrbK sequences, those from pTiR10 and pTiS4, always group together and although they most often form a clade with other plasmids in Group I (fig. 4*D*), these two proteins occasionally form a separate branch depending upon which program was used to align the amino acid sequences and which program was used to construct the cladograms (data not shown). Additionally for all trees constructed with the TrbK alignments, pSfr64a always groups separately from the other Group I plasmids, and most often groups with the two bradyrhizobial plasmids, pB11 and pHCG3 (fig. 4*D*).

The Dtr Region of pTi-SAKURA Is Chimeric

Although the trees group the three QS proteins, TrbE, and TrbK of pTi-SAKURA and pTiC58 closely together (figs. 2*A*–*C*, 4*C*, and 4*D*), the cladograms constructed using TraA and TraG sequences indicate that these proteins of pTi-SAKURA are not most closely related to those of pTiC58 (fig. 4*A* and *B*). Considering the incongruous cladograms for TraA and TraG and the proximal location of the two genes encoding these protein to *traM* and *traR* (supplementary fig. S2*A*, Supplementary Material online), we hypothesized that this region of the two Ti plasmids differs in its phylogenies due to a recombination event. To examine this possibility, we constructed phylogenetic trees of the protein products of the remaining *tra* genes, *traF*, *traB*, *traH*, *traC,* and *traD* (supplementary table S2, Supplementary Material online). In comparing these proteins from the entire group of 18 plasmids, the trees for TraB, TraH, TraC, and TraF yielded Group I- and Group II-like clades (supplementary fig. S2*C*–*F*, Supplementary Material online). However, the cladogram for TraD exhibited anomalies (supplementary fig. S2*B*, Supplementary Material online); the tree did yield Group II-like clades, but a few members of the Group I-like clade, including those from pTi-SAKURA, pTiS4, pTiBo542, pCB782 and pOV14c were more distantly related to their orthologs in the Group I clade in comparison with the other proteins analyzed (supplementary fig. S2*B*, Supplementary Material online). Examining the trees for the two nopaline-type Ti plasmids, the cladograms for TraH and TraB, like the trees for TraM and TraR, group pTi-SAKURA with pTiC58 (supplementary fig. S2*E* and *F*, Supplementary Material online, and fig. 2*A* and *B*). However, the trees constructed for TraD and TraC, like the tree of TraG, indicate that these proteins from pTiC58 are most closely related to the orthologous proteins of pAtK84b, whereas those of pTi-SAKURA are on a separate branch (supplementary fig. S2*B* and *C*, Supplementary Material online, and fig. 4*B*). The cladogram constructed for TraF has the protein from pTiC58 grouping closely with pAtS4c and more distantly grouping with that of pTi-SAKURA (supplementary fig. S2*D*, Supplementary Material online).

Table 2
TrlR Dimerizes with TraR from Group I, but Not Group II Plasmids

Tra[C] Plasmid[a]	Plasmid[b]	*trlR*	Conjugation Frequency[c]	Fold Inhibition[d]
pTiR10Δ*traM*	None	–	1.5×10^{-4}	NA
pTiR10Δ*traM*	pKK38	–	8.2×10^{-5}	1.8
pTiR10Δ*traM*	pPOKK*trlRA*	+	1.5×10^{-8}	5,500
pTiC58Δ*traM*	pKK38	–	8.0×10^{-5}	NA
pTiC58Δ*traM*	pPOKK*trlRA*	+	5.8×10^{-7}	140
pAoF64/95Δ*traM*	pKK38	–	1.2×10^{-2}	NA
pAoF64/95Δ*traM*	pPOKK*trlRA*	+	7.7×10^{-3}	1.6

NOTE.—NA, not applicable.
[a]All in *Agrobacterium tumefaciens* strain NTL4.
[b]*trlR* was expressed from pPOKK*trlRA* (supplementary table S1, Supplementary Material online).
[c]Expressed as the number of transconjugates recovered per input donor cell. Each cross was performed in duplicate and the values presented are the average of each cross.
[d]Calculated by dividing the transfer frequency of the tra[C] strain harboring an empty vector by the transfer frequency of the donor harboring the *trlR* expressing vector or by dividing the tra[C] strain by the same tra[C] strain harboring an empty vector.

The *oriT* DNA Sequences Do Not Divide into Clades Based on Group I or Group II Plasmid Organization

TraA initiates conjugative transfer by introducing a single-strand nick (Cho and Winans 2007) at the origin of transfer (*oriT*) site (supplementary table S2, Supplementary Material online). This *cis*-acting nucleotide sequence is located between the *traAFBH* and *traCDG* operons, and the core DNA sequence (fig. 4*E*) is related to the *oriT* of the IncQ plasmid RSF1010 (Cook and Farrand 1992). An alignment of the nucleotide sequences of the known and putative *oriT* sites from the 18 plasmids shows conserved DNA sequences, especially at the putative *nic* site indicated by the open triangle in figure 4*E*. Although the cladal structure of the TraA, TraG, and TrbE proteins from the 18 plasmids corresponds precisely to their organizational group, the nucleotide sequences of the putative *oriT* sites do not; the sequences of Group I and Group II plasmids intermingle within the tree (fig. 4*F*).

Recombinant Plasmids with Cloned *oriT* Regions from Group I and Group II Plasmids Are Mobilizable by Strains Harboring Plasmids from Like and Unlike Groups

That the *oriT* sequences are highly conserved (fig. 4*E*) suggests that a given *cis*-acting site may be recognized by the relaxosome components from both clades. To test this hypothesis, we determined whether different *oriT* inserts could be processed by the Dtr systems from cognate and noncognate plasmids. The three pRG970b-based *tra* box I clones from pTiC58, pAoF64/95, and pRi1724 described previously also encode the respective *oriT* sites. The three *oriT* vectors were transformed into two transfer-constitutive donors, NTL4(pTiC58Δ*accR*) a representative of the Group I clade (Beck von Bodman et al. 1992), and NTL4(pAoF64/95Δ*mrtR*), from the Group II clade (Wetzel ME and Farrand SK, unpublished data). We then determined whether these transfer-constitutive donors could mobilize each *oriT* plasmid to the recipient strain, *A. tumefaciens* C58C1RS.

Both donors mobilized all three of the *oriT* vectors (table 3). Unsurprisingly, both mobilized the vectors containing their cognate *oriT* most efficiently (table 3). Moreover, the donor harboring pAoF64/95 mobilized the plasmid containing $oriT_{pTiC58}$ almost as efficiently as its cognate *oriT*. Although the Ao plasmid-containing donor mobilized the $oriT_{pRi1724}$ plasmid, the efficiency of transfer dropped by about 1 order of magnitude in comparison to transfer of a plasmid with its own *oriT* region (table 3). Donors harboring pTiC58 mobilized both $oriT_{pAoF64/95}$ and $oriT_{pRi1724}$ at similar efficiencies, but at frequencies approximately 70-fold lower in comparison to mobilization of a vector with the cognate *oriT* (table 3).

The RepABC Plasmid Replication and Partitioning System, while Adjacent to the *trb* Operon, Is More Highly Divergent than the QS and Transfer Systems

All 18 plasmids included in this study initiate replication from a *repABC* system that is conserved in the majority of large plasmids of the α-proteobacteria (Pinto et al. 2012). In the known Class I elements, the *repABC* operon is invariably adjacent and divergently oriented to the *traI/trb* operon (fig. 1 and supplementary fig. S1, Supplementary Material online). Moreover, in representatives from both organizational types transcription of the *repABC* operon, and concomitant increase in plasmid copy number, is positively controlled from the upstream *tra* box III by TraR in a quorum-dependent manner (Li and Farrand 2000; Pappas and Winans 2003; McAnulla et al. 2007). Given the influence of QS on expression of the *rep* genes (Li and Farrand 2000; Pappas and Winans 2003; McAnulla et al. 2007), and the location and importance of the *rep* system for stable maintenance of the plasmid, we assessed all three Rep proteins for evolutionary relatedness as we did with the proteins of the QS and conjugative transfer systems. Even though the *repABC* genes are adjacent to the conjugative transfer genes in these plasmids, the RepA, RepB, and RepC proteins do not divide into distinct clades corresponding to organizational Groups I or II (fig. 5*A*–*C*). Furthermore, although the cladograms of RepA and RepB are mostly congruous with each other, the tree for RepC is not (fig. 5). Despite their close linkage, the cladograms indicate that RepC is evolving independently from the RepAB pair.

The QS and Transfer Proteins Are Coevolving with Each Other but Separately from the RepA, RepB, and RepC Proteins

The QS proteins and the conjugative transfer proteins, with the possible exception of TrbK, fall into two major clades that are consistent with the type of plasmid organization. If the protein sets within each plasmid are coevolving, in comparison two proteins should occupy a similar position on the

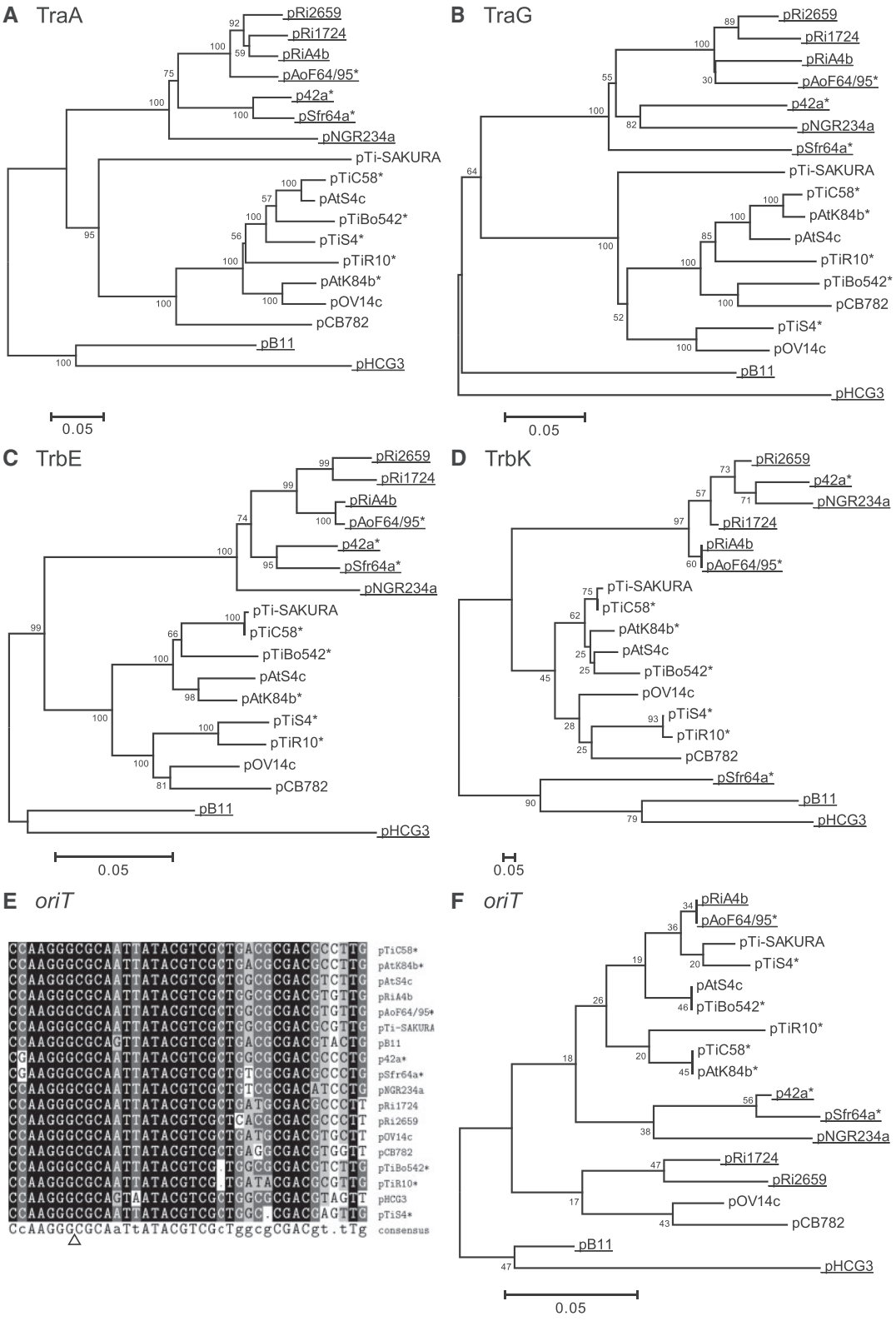

Fig. 4.—TraA, TraG, TrbE, and TrbK divide into two major clades, whereas the *oriT* sequences do not. NJ trees of the individual proteins from 18 plasmids constructed from MAFFT alignments are shown. Bootstrap values from 1,000 replicates are located next to the branches. The evolutionary distances were computed with the Poisson correction. (*A*) TraA, (*B*) TraG, (*C*) TrbE, and (*D*) TrbK. (*E*) The nucleotide sequences of the *oriT* region of the 18 selected Class I plasmids were aligned with MAFFT. The black color indicates invariant bases, whereas the bases colored dark gray are highly conserved. Nucleotides colored

Table 3

oriT Regions Are Recognized by Cognate and Noncognate *Trans*-Acting Transfer Functions

Tra[C] Plasmid[a]	*oriT*[b]	Mobilization Frequency[c]
pTiC58Δ*accR*	pTiC58	2.89×10^{-3}
pTiC58Δ*accR*	pRi1724	4.25×10^{-5}
pTiC58Δ*accR*	pAoF64/95	4.55×10^{-5}
pTiC58Δ*accR*	None	6.30×10^{-6}
pAoF64/95Δ*mrtR*	pTiC58	1.03×10^{-3}
pAoF64/95Δ*mrtR*	pRi1724	3.53×10^{-4}
pAoF64/95Δ*mrtR*	pAoF64/95	2.73×10^{-3}
pAoF64/95Δ*mrtR*	None	$<10^{-7}$

[a]All in *Agrobacterium tumefaciens* strain NTL4.

[b]The *traA-C* intergenic region, including the *oriT* region of three plasmids cloned in pRG970b and placed in *trans* to the tra[C] strain.

[c]Expressed as the number of transconjugates recovered per input donor cell. Each cross was performed in duplicate or triplicate and the values presented represent the average of each cross.

phylogenetic trees in relation to the other branches. However, if the two proteins in question occupy very different positions on the cladograms in relation to the other branches, then the two proteins from the same plasmid have not evolved together. We assessed the likelihood of how well the phylogenies agreed between protein sequences by comparing how well the data for one protein set would fit on a tree constructed with the data from a second protein set. Using the formula described in Materials and Methods, each comparison between protein data set and tree was given a score. If the trees are identical, the score is 1.00 and the heat map color is white (fig. 6*A*). The lower the level of relatedness between a protein data set and the tree for a different protein translates to a lower number and a darker color on the heat map.

The trees of the QS and transfer proteins are fairly compatible with each of the QS and transfer protein data sets (fig. 6*A*, columns). The data for TraR indicate that the trees and protein alignments are most compatible with the trees and alignments of TraM and TraI (fig. 6*A*, TraR rows and columns). Although the data for the TraM proteins can be fitted to the optimal trees constructed from each protein data set (fig. 6*A*, TraM row), this tree best fits with the TraR protein alignment (fig. 6*A*, TraM column). The TrbK protein alignment fits best with the tree constructed with TrbE (fig. 6*A*, TrbK row). The TrbK tree, however, fits best with TraR, TrbE and TraM, and slightly less well with the remaining QS and transfer proteins alignments (fig. 6*A*, TrbK column).

Unlike the QS and transfer proteins RepA and RepB share nearly identical optimal tree constructions but differ from any of the other trees. However, the RepC protein data set does not fit well with any of the other protein data sets, including RepA and RepB (fig. 6*A*, RepC row). Among all of the other proteins the RepC tree fits best with the RepB protein data set followed by the RepA protein alignment (fig. 6*A*, RepC column). Although the RepC protein alignment fits best with the RepA and RepB trees, the values are low (fig. 6*A*, RepC row).

Discussion

The QS and Transfer Proteins Overall Form Three Clades Associated with Plasmid Organization and the Proteins within a Particular Clade Are Coevolving

All of the 18 *repABC* plasmids in this study that contain Class I, QS-regulated conjugative transfer systems fall into two organizational types first noted by Moriguchi et al. (2001). Our phylogenetic analysis indicates that the evolution of the QS and transfer proteins correspond to organization; those from the Group I plasmids form one clade (Clade I), whereas the same proteins from plasmids with the Group II organization subdivide into two clades (figs. 1*A* and 6*B*), one from plasmids found in the family Rhizobiaceae (Clade II), the other from plasmids in the family Bradyrhizobiaceae (Clade III) (fig. 6*B*).

Remarkably, with respect to the virulence elements of *Agrobacterium*, the organizational split correlates with function. We know of no Ri plasmid with Group I organization. Nor do we know of any Ti plasmid with Group II organization. This observation implies that all of the Ti plasmids and all of the Ri plasmids, while sharing a common ancestor, have diverged and are evolving independently. However, among opine-catabolic plasmids, both organizational types are represented (see table 1). We conclude that this "cheater" class of opine catabolic plasmids can evolve from either Group I or Group II elements. The two Sym plasmids examined, pNGR234a and pCB782, divide between Groups I and II (table 1 and fig. 6*B*). This observation suggests that this functional group of plasmids is less constrained in its evolution as compared with the virulence elements of the agrobacteria.

The trees constructed from the proteins of pHCG3 and pB11 suggest that these two Group II bradyrhizobial plasmids have diverged from the other Group II elements and now constitute an independently evolving clade. Although only sparsely described, the existence of Class I *repABC* plasmids in these diverse and ecologically distinct bacteria indicates that members of this family of episomal elements have extended their genetic range beyond bacteria that interact with plants.

Fig. 4.—Continued

light gray and white show less conservation. Capital letters in the consensus sequences indicate invariant bases, whereas lower case letters are bases conserved in 50% or more of the sequences. The triangle beneath the consensus sequence indicates the location on the complementary strand of the *nic* site of the *oriT* of pTiR10 (Cho and Winans 2007). (*F*) An NJ tree constructed using the data from the MAFFT alignment of the *oriT*. Bootstrap values located next to the branches are from 1,000 replicates. The asterisks (*) indicate plasmids that are known to be conjugative. Plasmids with Group II organization of the *rep* and Class I transfer genes are underlined.

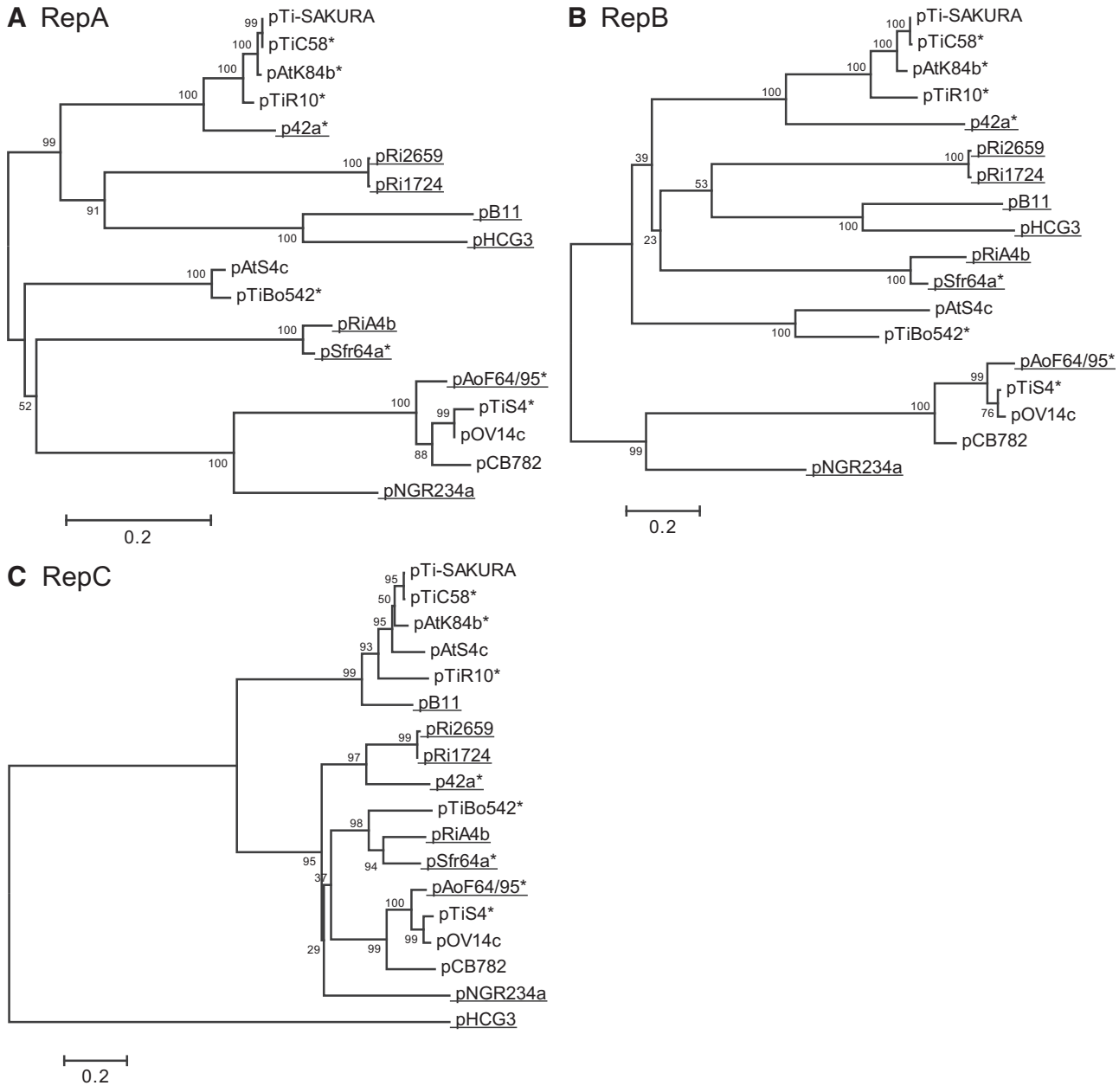

Fig. 5.—The Rep protein sequences do not divide based on plasmid organization. Trees were constructed from MAFFT alignments of (*A*) RepA, (*B*) RepB, and (*C*) RepC proteins using the NJ methodology. Evolutionary distances were estimated using the Poisson correction implemented in MEGA 6. Underlined plasmids have Group II organization. Asterisks (*) indicate plasmids that are known to be conjugative.

Although there is a clear phylogenic divergence of the QS and transfer proteins into three clades, consistency in the branch structures of the 12 QS and transfer proteins examined indicates that these proteins have not only evolved into separate clades but are also coevolving in their particular plasmids within a clade (figs. 2 and 4 and supplementary fig. S2, Supplementary Material online). The scores in the heat map analysis support this conclusion (fig. 6*A*). The robustness of these evolutionary patterns is further illustrated by the structure of a cladogram constructed using a concatenated alignment of all 12 protein sequences from the 18 plasmids examined (fig. 6*B*).

Overall, components of the QS and transfer systems that interact with each other or lie in the same operon show the highest degree of coevolution. For example, TraM and TraR interact with each other, and for both Group I and Group II plasmids, the *traR* and *traM* genes are always located near one another (fig. 1 and supplementary fig. S1, Supplementary

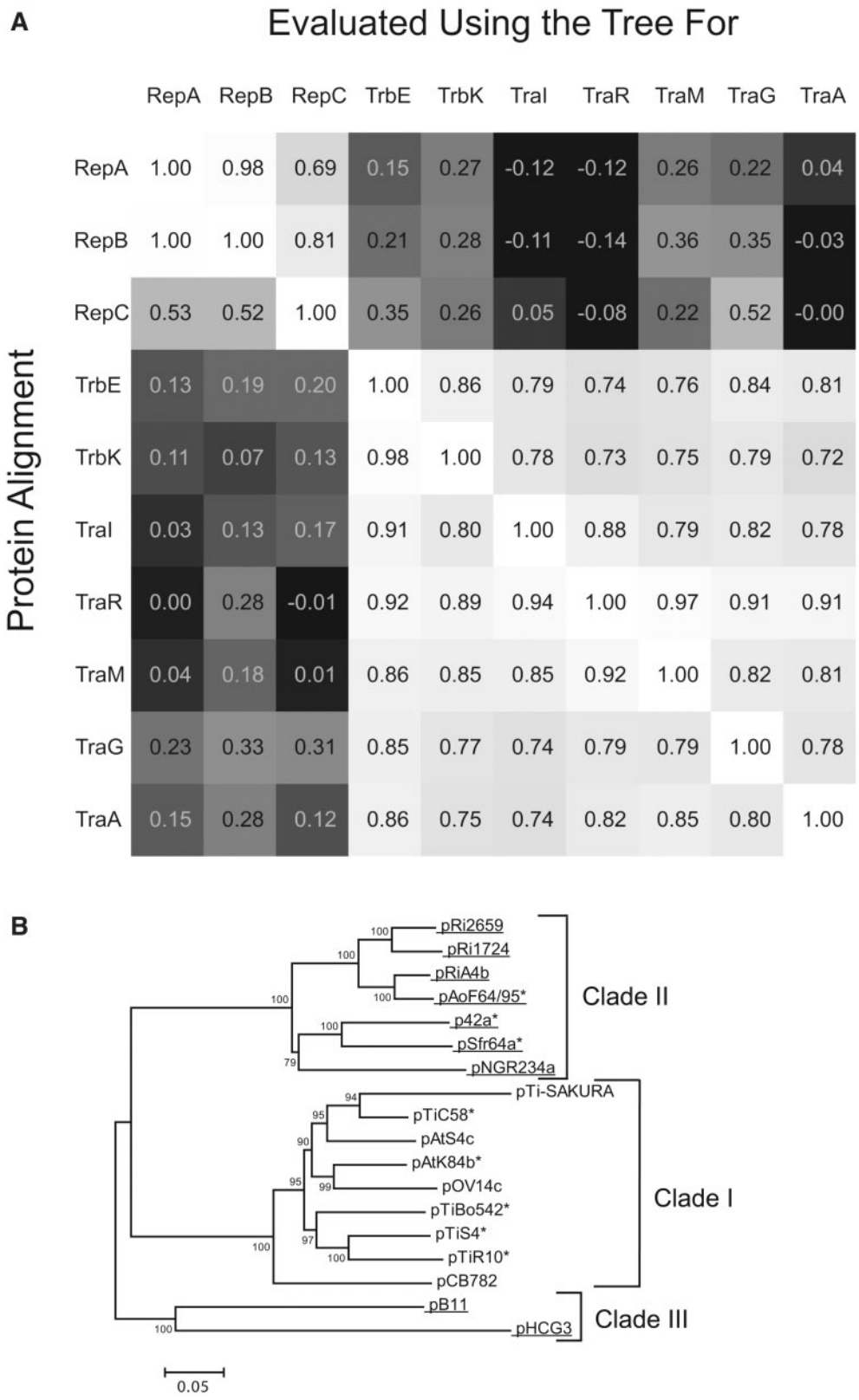

Fig. 6.—Overall, the QS and transfer proteins, but not the Rep proteins, are coevolving and divide into three separate clades. (*A*) A heat map in which alignments for ten of the core transfer, QS and Rep proteins (vertical axis) were given a score based on how well they fit onto the trees of each protein (horizontal axis) as described in Materials and Methods. The shading corresponds with the values; identical tree topologies give scores of 1.00 and are white,

Material online). Concomitantly, these two proteins exhibit largely similar trees and show a high degree of coevolution in the heat map (figs. 2 and 6*A*).

The Highly Conserved *oriT* Regions Can Be Recognized and Nicked by Components of the Relaxosome from Both Clades

That there is cross-cladal functioning between the *trans*-acting *tra* components and *cis*-acting *oriT* sites (table 3) raises the possibility that induction of transfer by a specific signal of one plasmid may stimulate transfer from that donor of a co-resident, noninduced plasmid with an appropriate *oriT* region. Although cotransfer of plasmids has been described in both *Rhizobium* and *Agrobacterium* species, most studies have concluded that this phenomenon is likely the result of plasmid co-integration (Petit et al. 1983; Brom et al. 2000). However, cotransfer of a noninduced megaplasmid by *trans*-mobilization has been reported (Torres Tejerizo et al. 2014). The range of mechanisms involved in cotransfer of these large plasmids merits further study.

TraR Proteins Have DNA Binding and Dimerization Functions that Are Conserved within, but Not between Clades

Our in vivo analysis showing that TraR-dependent promoters are activated by only cognate and closely related TraR orthologs is consistent with the in silico results that the QS and transfer proteins are coevolving with each other within, but not between clades (figs. 3 and 6*A*). Our results extend those of He et al. (2003) showing that TraR of pNGR234a, with a Group II organization, can activate transcription of at least two of its cognate promoters, but does not activate the same promoters from pTiR10, a Group I-type plasmid (He et al. 2003). Moreover, $TraR_{pTiR10}$ activates a cognate TraR-dependent promoter but not the orthologous promoter from pNGR234a (He et al. 2003). Coupled with our observation that TraR proteins from the two major clades cannot efficiently cross-dimerize, our results suggest that both DNA binding and dimerization domains of TraR proteins are structurally and functionally conserved within, but not between the clades.

$TraR_{pRi1724}$ Functionally Differs from Other TraR Proteins in the *repABC* Family of Plasmids with Class I Transfer Systems

Gene activation by TraR of pRi1724 differs remarkably from that of the archetypical system. The TraR proteins from pTiC58 and pAoF64/95 strongly activate their cognate *traA* promoters, but only in cells grown with the appropriate quormone (fig. 3*B*). TraR from pRi1724, on the other hand, activates its cognate *traA* promoter, and also that from pAoF64/95, but only in cells grown in the absence of the QS signal (fig. 3*A*). Three lines of evidence indicate that the anomalous behavior of $TraR_{pRi1724}$ is a function of the activator, and not of its *tra* box binding site. First, the *tra* box I sequences of pRi1724 and pAoF64/95 are identical (fig. 2*E*). Second, $TraR_{pAoF64/95}$ activates expression from the *traAFBH* promoter from both pAoF64/95 and pRi1724 in a quormone-dependent manner (fig. 3*B*). Third, $TraR_{pRi1724}$ activates transcription of these two promoters only in the absence of the autoinducer (fig. 3*A*). That the amino acid sequences of the recognition helices of $TraR_{pRi1724}$ and $TraR_{pAoF64/95}$ are strongly conserved suggests that the anomalous properties of $TraR_{pRi1724}$ are due to alterations in regions of the protein other than those involved in quormone binding or DNA site recognition.

The anomalous activity of $TraR_{pRi1724}$ could be a result of random mutational events or could represent a regulatory divergence where transfer is induced only when the acyl-HSL-producing population is low. Other LuxR homologs such as EsaR of *Pantoea* spp. and YenR of *Yersinia enterocolitia* bind to their target DNA sequence in the absence, and not the presence, of the acyl-HSL cognate to the system. Although apo-EsaR represses transcription of the target genes when the population density is low (Beck von Bodman et al. 1998; Minogue et al. 2002), apo-YenR activates transcription of a target gene, *yenS*, at low population densities, a regulatory strategy called quorum-hindering (Tsai and Winans 2011).

The RepA, RepB, and RepC Proteins Do Not Coevolve with the Transfer and QS Proteins

RepA, RepB, and RepC comprise a well-characterized family of replication proteins that, in the 18 Class I plasmids included in this study, are located adjacent and divergently oriented to the *trai/trb* operon. The fact that the arrangement of the *rep* genes and the *trai/trb* operon is obligatory in plasmids with the Class I transfer system raises the possibility that the *rep* system is coevolving with the genes of the adjacent transfer systems and perhaps even the QS regulatory system. This proved not to be the case; the trees constructed for the RepA, RepB, and RepC proteins do not divide based on plasmid organization (fig. 5). Moreover, none of the Rep proteins

Fig. 6.—Continued

whereas protein alignments that fit less well to a given tree and are not coevolving have lower numbers and are darker shades of gray or are black. (*B*) NJ tree constructed from MAFFT alignments from a concatenated sequence of the Tra, Trb, and QS proteins from the 18 plasmids. The concatenated protein contains the following proteins in order: TraR, TraM, TraI, TraA, TraG, TrbE, TrbK, TraC, TraD, TraF, TraH, and TraB. Bootstrap values from 1,000 replicates are located next to the branches. The evolutionary distances were computed in Mega 6 using the Poisson correction. The asterisks (*) denote plasmids that are known to be conjugative, whereas the plasmids with a Group II plasmid organization are underlined.

are coevolving with the *trall/trb* operon, despite their close physical linkage as divergently oriented operons (figs. 2, 4, and 5). This divergent evolution of the *rep* and transfer genes is especially striking considering that the *repABC* and *trall/trb* operons share a region containing promoters and *tra* boxes, allowing TraR-mediated regulation of transcription of not only the genes for transfer but also the gene system for plasmid replication and partitioning (Li and Farrand 2000; Pappas and Winans 2003; McAnulla et al. 2007).

Our analyses showing that RepA and RepB proteins coevolve with each other but do not strongly evolve with RepC confirm previous studies of the coevolutionary relationships of these three proteins (reviewed in Pinto et al. 2012). Although the RepABC proteins have been extensively studied, to our knowledge, ours is the first report comparing the phylogenies of the RepA, RepB, and RepC proteins with the proteins of the conjugative transfer and QS systems encoded by these plasmids. This independent evolution makes sense considering that genetically active components within the *repABC* operon determine plasmid incompatibility (Ramírez-Romero et al. 2000; Bartosik et al. 2001; Venkova-Canova et al. 2004; Chai and Winans 2005a, 2005b; MacLellan et al. 2005, 2006; Cervantes-Rivera et al. 2011). This property of plasmids is often determined by the degree of relatedness of the replication and partitioning functions of the elements (Novick and Hoppensteadt 1978). The problem of incompatibility is of considerable interest; some isolates, such as *Rhizobium etli* strain CFN 42, stably maintain as many as six different *repABC* replicons (González et al. 2006). Clearly the components of the replication system are actively evolving to generate different incompatibility properties. Uncoupling the evolution of the replication system from the transfer system provides these plasmids with the ability to transfer to and be stably maintained in a variety of strains, even those harboring other *repABC* replicons.

The three Rep proteins, as well as a small RNA and *cis*-acting sites within the operon are involved in plasmid incompatibility (Ramírez-Romero et al. 2000; Cervantes-Rivera et al. 2011; Pinto et al. 2012; Pérez-Oseguera and Cevallos 2013). pTiC58, pAtK84b, and pTiR10 all belong to the IncRh-1 group (Montoya et al. 1978; Hooykaas et al. 1980; Clare et al. 1990). Strikingly, for the trees constructed with the RepA, RepB and RepC protein sequences, these plasmids group within the same subclade (fig. 5). The tartrate-catabolism plasmid of *Agrobacterium vitis* strain S4, pAtS4c, belongs to the IncRh-2 class (Szegedi and Otten 1998), whereas the coresident Ti plasmid, pTiS4, belongs to the IncRh-4 class (Szegedi et al. 1996). Consistent with this compatibility, the RepA, RepB, and RepC proteins of pTiS4 and pAtS4c are members of separate subclades. These two examples provide phylogenetic support for the role of divergent evolution of these proteins in determining the ability of the *repABC* plasmids to coexist in the same host.

Some of These Plasmids Are Highly Homologous and Have Shared Evolutionary Histories

Several of the plasmids examined in this analysis share significant regions of synteny. For example, pRi1724 and pRi2659 are highly homologous over their entire lengths (Moriguchi et al. 2001). Both of these Ri plasmids were isolated from biovar 1 strains of *Agrobacterium rhizogenes*, one from a diseased melon plant in Japan and the other from a diseased cucumber plant in the United Kingdom. For nearly all of the 15 protein sequences analyzed, including RepA, RepB and RepC, the proteins of pRi1724 and pRi2659 are related most closely to each other. Considering that overall the Rep proteins evolve separately from the transfer and QS proteins, the fact that these two sets of proteins of pRi1724 and pRi2659 are evolving together indicates that the two plasmids share a very recent common ancestor.

Similarly, pTiC58 and pTi-SAKURA which are both nopaline–agrocinopine-type Ti plasmids share syntenic regions that are highly homologous (Suzuki et al. 2000; Goodner et al. 2001). Despite the apparent recombination-generated divergence in the region around *traF*, the QS, conjugative transfer, and Rep proteins of pTiC58 most closely resemble those of pTi-SAKURA, indicating that these two plasmids share a recent common ancestor. That such a recombination event has occurred is not surprising, plasmids in the *repABC* family are known to recombine, often by forming and resolving co-integrates (White and Nester 1980; Jouanin et al. 1986; Brom et al. 2000; Mavingui et al. 2002).

Evolution of the Class I-*repABC* Family of Plasmids Is Complex

The differences in organization of the conjugative transfer genes of the Group I and Group II plasmids and the chimeric nature of these large plasmids make the evolutionary study of these extrachromosomal elements intriguing and complex. The two types of plasmid organization suggest that expansion of a conceptual core plasmid containing the QS, *rep*, and transfer genes (pCORE; fig. 7) occurs by inserting new DNA primarily in two locations. The first and perhaps preferred location occurs between the *repABC* and *traCDG* operons (fig. 7 Site A). The second region (fig. 7 Site B) is located downstream of the *trb* genes and separates this operon from the *traR–traM–tra* cluster. Both Group I and II plasmids exhibit insertions in variable region A, whereas Group I plasmids have a second variable region inserted into Site B (fig. 7). Intriguingly, Ti plasmids, which all show Group I organization, differ with respect to which of the two variable regions encoding loci important for tumor induction are located. For example, the segment of DNA encoding the T-region and the *vir* regulon of the Octopine-type Ti plasmids is inserted into Site A, whereas the analogous segment of pTiC58 is inserted in Site B. Of significance to the evolution of regulation in the Group I Ti plasmids, insertions into the B site

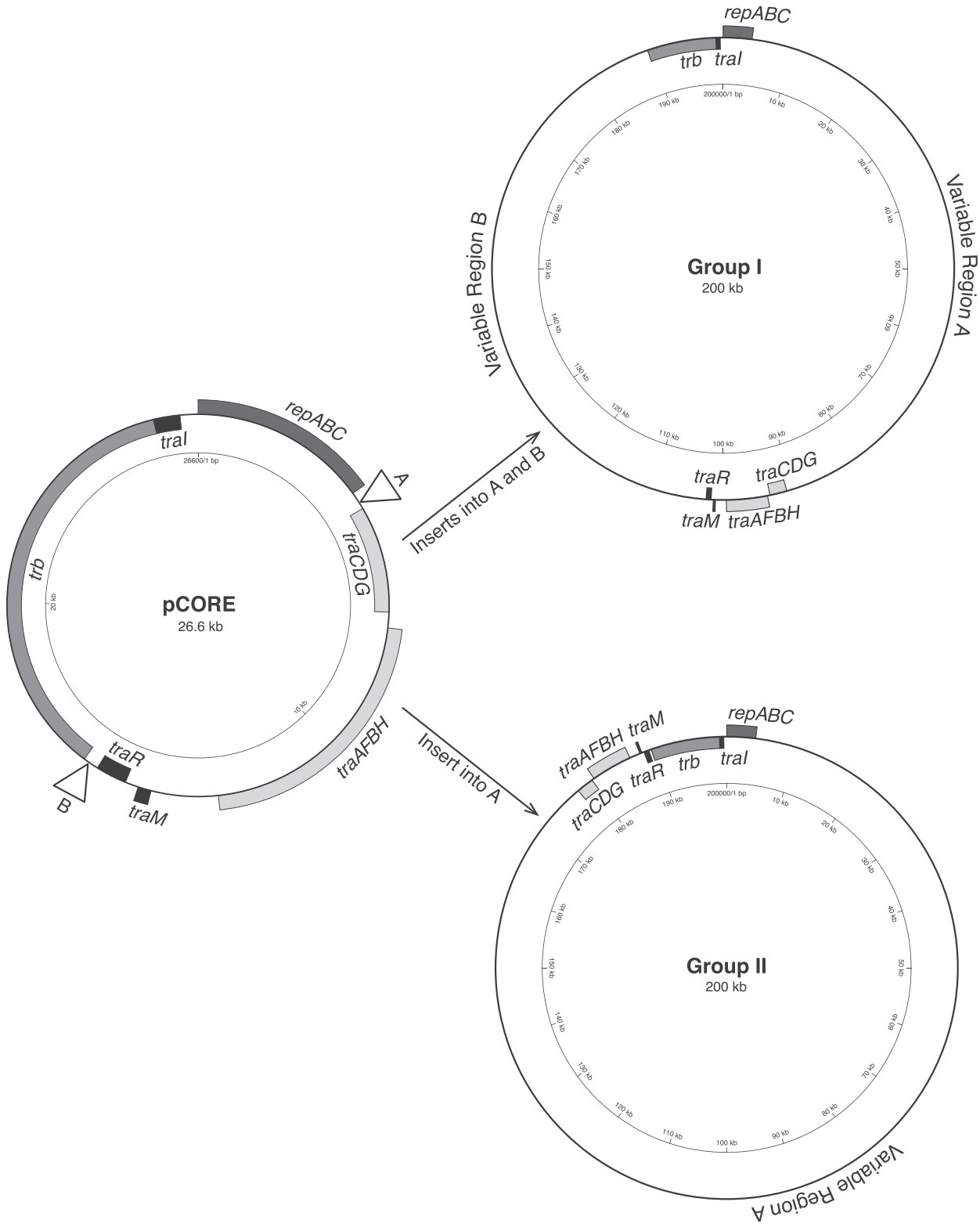

Fig. 7.—Acquisition of additional DNA by Class I-*repABC* plasmids occurs at two favored locations. A conceptual *repABC* plasmid, pCORE, containing the core Class I transfer genes (medium gray for Mpf genes and light gray for Dtr genes), the QS genes *traR*, *traM* and *traI* (all in black), and the *rep* operon (dark gray). The triangles labeled A and B mark the sites at which additional sequence is found in the two organizational types, Group I and Group II, of plasmids.

provide a mechanism by which *traR* can be fused to a gene system regulated by the conjugative signal. Insertions into the other two regions, between the two *tra* operons and between *traI/trb* and *repABC*, most likely is restricted by their richness in *cis*-acting promoter and regulatory sequences.

Despite the organizational differences and the role of recombination and horizontal gene transfer in increasing genetic diversity, overall our analysis indicates that the genes of the *tra* and *trb* regions of a particular plasmid are coevolving. Additionally, the orthologous genes shared by these plasmids within a clade coevolve and retain a level of functionality within, but not between clades. Furthermore, we conclude that the QS proteins overall are coevolving with the transfer system.

Although the evolution of these plasmids may be restricted to the two organizational groups, it is most probable that additional clades will emerge as more sequence becomes available. Even now, the tree structures of pHCG3 and pB11 suggest that these two plasmids have evolved into a new Group II clade (fig. 6*B*, Clade III). Evolution of these plasmids is of particular importance to the order Rhizobiales for several reasons. First, the *repABC* family of plasmids exhibits an extended conjugative and replicative host-range among the α-proteobacteria. Second, the *repABC* elements have a large genetic carrying capacity and the genes encoded on these replicons often confer defining characteristics to the bacteria that harbor them. Third, quite clearly some of the *repABC* plasmids are evolving into chromids in a range of genomes within the order Rhizobiales (Harrison et al. 2010; Bailly et al. 2011; Landeta et al. 2011). Considering the potential for conjugative transfer and dissemination of these Class I plasmids, the large genetic carrying capacity, their propensity to recombine, and the extended host range encoded by *repABC* lend importance to the study of the evolution and interactions of this group of plasmids, and the bacteria that harbor them.

Supplementary Material

Supplementary figures S1 and S2 and tables S1 and S2 are available at *Genome Biology and Evolution* online (http://www.gbe.oxfordjournals.org/).

Acknowledgments

This work was supported in part by National Institute of Health (grant number R01 GM52465 to S.K.F.), Syngenta Sponsored Research Agreement (grant number 2010-06329 to S.K.F.), the United States Department of Energy Office of Biological and Experimental Research (grant number SC0006642 to J. Sweedler, P. Bohn, and S.K.F.), and the United States Department of Energy Systems Biology Knowledgebase project (grant number DE-AC02-06CH11357 to G.J.O.]. The authors thank Clay Fuqua for providing them with the strain containing the *traM* mutant of pTiR10 and Rachel J. Whitaker for her input on this manuscript.

Literature Cited

Bailly X, et al. 2011. Population genomics of *Sinorhizobium medicae* based on low-coverage sequencing of sympatric isolates. ISME J. 5:1722–1734.

Bartosik D, Szymanik M, Wysocka E. 2001. Identification of the partitioning site within the *repABC*-type replicon of the composite *Paracoccus versutus* plasmid pTAV1. J Bacteriol. 183:6234–6243.

Beck von Bodman S, Hayman GT, Farrand SK. 1992. Opine catabolism and conjugal transfer of the nopaline Ti plasmid pTiC58 are coordinately regulated by a single repressor. Proc Natl Acad Sci U S A. 89:643–647.

Beck von Bodman S, Majerczak DR, Coplin DL. 1998. A negative regulator mediates quorum-sensing control of exopolysaccharide production in *Pantoea stewartii* subsp *stewartii*. Proc Natl Acad Sci U S A. 95:7687–7692.

Beitz E. 2000. TeXshade: shading and labeling of multiple sequence alignments using LaTeX2e. Bioinformatics 16:135–139.

Brom S, García-de los Santos A, Cervantes L, Palacios R, Romero D. 2000. In *Rhizobium etli* symbiotic plasmid transfer, nodulation competitivity and cellular growth require interaction among different replicons. Plasmid 44:34–43.

Cangelosi GA, Best EA, Martinetti G, Nester EW. 1991. Genetic analysis of *Agrobacterium*. In: Jeffrey HM, editor. Methods in enzymology. San Diego: Academic Press. p. 384–397.

Castillo-Ramirez S, Vazquez-Castellanos JF, Gonzalez V, Cevallos MA. 2009. Horizontal gene transfer and diverse functional constraints within a common replication-partitioning system in Alphaproteobacteria: the *repABC* operon. BMC Genomics 10:536.

Cervantes-Rivera R, Pedraza-Lopez F, Perez-Segura G, Cevallos MA. 2011. The replication origin of a *repABC* plasmid. BMC Microbiol. 11:158.

Cevallos MA. 2002. *Rhizobium etli* CFN42 contains at least three plasmids of the *repABC* family: a structural and evolutionary analysis. Plasmid 48:104–116.

Cevallos MA, Cervantes-Rivera R, Gutiérrez-Ríos RM. 2008. The *repABC* plasmid family. Plasmid 60:19–37.

Chai Y, Winans SC. 2005a. RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. Mol Microbiol. 58:1114–1129.

Chai Y, Winans SC. 2005b. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. Mol Microbiol. 56:1574–1585.

Chai Y, Zhu J, Winans SC. 2001. TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function *in vitro* by forming inactive TrlR:TraR dimers. Mol Microbiol. 40:414–421.

Chen L, Chen Y, Wood DW, Nester EW. 2002. A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. J Bacteriol. 184:4838–4845.

Cho H, Pinto UM, Winans SC. 2009. Transsexuality in the rhizosphere: quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. J Bacteriol. 191:3375–3383.

Cho H, Winans SC. 2007. TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of *Agrobacterium tumefaciens*. Mol Microbiol. 63:1769–1782.

Clare BG, Kerr A, Jones DA. 1990. Characteristics of the nopaline catabolic plasmid in *Agrobacterium* strains K84 and K1026 used for biological control of crown gall disease. Plasmid 23:126–137.

Cook DM, Farrand SK. 1992. The *oriT* region of the *Agrobacterium tumefaciens* Ti plasmid pTiC58 shares DNA sequence identity with the

transfer origins of RSF1010 and RK2/RP4 and with T-region borders. J Bacteriol. 174:6238–6246.
Danino VE, Wilkinson A, Edwards A, Downie JA. 2003. Recipient-induced transfer of the symbiotic plasmid pRL1JI in *Rhizobium leguminosarum* bv. *viciae* is regulated by a quorum-sensing relay. Mol Microbiol. 50:511–525.
Datsenko KA, Wanner BL. 2000. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proc Natl Acad Sci U S A. 97:6640–6645.
Dessaux Y, Petit A, Farrand SK, Murphy PJ. 1998. Opines and opine-like molecules involved in plant-Rhizobiaceae interactions. In: Spaink H, Kondorosi A, Hooykaas PJ, editors. The Rhizobiaceae. Dordrecht: Kluwer Academic Publishers. p. 173–197.
diCenzo G, Milunovic B, Cheng J, Finan TM. 2013. The $tRNA^{arg}$ gene and *engA* are essential genes on the 1.7-Mb pSymB megaplasmid of *Sinorhizobium meliloti* and were translocated together from the chromosome in an ancestral strain. J Bacteriol. 195:202–212.
Ding H, Hynes MF. 2009. Plasmid transfer systems in the rhizobia. Can J Microbiol. 55:917–927.
Edgar RC. 2004a. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics 5:113.
Edgar RC. 2004b. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792–1797.
Ellis JG, Murphy PJ, Kerr A. 1982. Isolation and properties of transfer regulatory mutants of the nopaline Ti-plasmid pTiC58. Mol Gen Genet. 186:275–281.
Farrand S, Hwang I, Cook D. 1996a. The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. J Bacteriol. 178:4233–4247.
Farrand SK. 1998. Conjugal plasmids and their transfer. In: Spaink H, Kondorosi A, Hooykaas PJ, editors. The Rhizobiaceae. Dordrecht: Kluwer Academic Publishers. p. 199–233.
Farrand SK, Hwang I, Cook DM. 1996b. The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. J Bacteriol. 178:4233–4247.
Farrand SK, Qin Y, Oger P. 2002. Quorum-sensing system of *Agrobacterium* plasmids: analysis and utility. Methods in enzymology. 358:452–484.
Felsenstein J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791.
Fuhrmann S, et al. 2003. Complete nucleotide sequence of the circular megaplasmid pHCG3 of *Oligotropha carboxidovorans:* function in the chemolithoautotrophic utilization of CO, H_2 and CO_2. Gene 322:67–75.
Fuqua C, Burbea M, Winans S. 1995. Activity of the *Agrobacterium* Ti plasmid conjugal transfer regulator TraR is inhibited by the product of the *traM* gene. J Bacteriol. 177:1367–1373.
Fuqua C, Winans SC. 1994. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. J Bacteriol. 176:2796–2806.
Fuqua C, Winans SC. 1996. Conserved *cis*-acting promoter elements are required for density-dependent transcription of *Agrobacterium tumefaciens* conjugal transfer genes. J Bacteriol. 178:435–440.
Galibert F, et al. 2001. The composite genome of the legume symbiont *Sinorhizobium meliloti*. Science 293:668–672.
Gérard J-C, Canaday J, Szegedi E, de la Salle H, Otten L. 1992. Physical map of the vitopine Ti plasmid pTiS4. Plasmid 28:146–156.
Giusti MdlÁ, et al. 2012. Genetic and functional characterization of a yet-unclassified rhizobial Dtr (DNA-transfer-and-replication) region from a ubiquitous plasmid conjugal system present in *Sinorhizobium meliloti*, in *Sinorhizobium medicae*, and in other nonrhizobial Gram-negative bacteria. Plasmid 67:199–210.
González V, et al. 2006. The partitioned *Rhizobium etli* genome: genetic and metabolic redundancy in seven interacting replicons. Proc Natl Acad Sci U S A. 103:3834–3839.
Goodner B, et al. 2001. Genome sequence of the plant pathogen and biotechnology agent *Agrobacterium tumefaciens* C58. Science 294:2323–2328.
Hamilton CM, et al. 2000. TraG from RP4 and TraG and VirD4 from Ti plasmids confer relaxosome specificity to the conjugal transfer system of pTiC58. J Bacteriol. 182:1541–1548.
Harrison PW, Lower RPJ, Kim NKD, Young JPW. 2010. Introducing the bacterial "chromid": not a chromosome, not a plasmid. Trends Microbiol. 18:141–148.
He X, et al. 2003. Quorum sensing in *Rhizobium* sp. strain NGR234 regulates conjugal transfer (*tra*) gene expression and influences growth rate. J Bacteriol. 185:809–822.
Hooykaas PJ, den Dulk-Ras H, Ooms G, Schilperoort RA. 1980. Interactions between octopine and nopaline plasmids in *Agrobacterium tumefaciens*. J Bacteriol. 143:1295–1306.
Hwang I, Cook D, Farrand S. 1995. A new regulatory element modulates homoserine lactone-mediated autoinduction of Ti plasmid conjugal transfer. J Bacteriol. 177:449–458.
Hwang I, et al. 1994. TraI, a LuxI homologue, is responsible for production of conjugation factor, the Ti plasmid N-acylhomoserine lactone autoinducer. Proc Natl Acad Sci U S A. 91:4639–4643.
Jouanin L, Tourneur J, Tourneur C, Casse-Delbart F. 1986. Restriction maps and homologies of the three plasmids of *Agrobacterium rhizogenes* strain A4. Plasmid 16:124–134.
Katoh K, Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief Bioinformatics 9:286–298.
Khan SR, Gaines J, Roop RM, Farrand SK. 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. Appl Environ Microbiol. 74:5053–5062.
Landeta C, et al. 2011. Plasmids with a chromosome-like role in rhizobia. J Bacteriol. 193:1317–1326.
Li PL, Everhart DM, Farrand SK. 1998. Genetic and sequence analysis of the pTiC58 *trb* locus, encoding a mating-pair formation system related to members of the type IV secretion family. J Bacteriol. 180:6164–6172.
Li PL, Farrand SK. 2000. The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. J Bacteriol. 182:179–188.
Li PL, Hwang I, Miyagi H, True H, Farrand SK. 1999. Essential components of the Ti plasmid *trb* system, a type IV macromolecular transporter. J Bacteriol. 181:5033–5041.
Luo ZQ, Clemente TE, Farrand SK. 2001. Construction of a derivative of *Agrobacterium tumefaciens* C58 that does not mutate to tetracycline resistance. Mol Plant Microbe Interact. 14:98–103.
Luo ZQ, Farrand SK. 1999. Signal-dependent DNA binding and functional domains of the quorum-sensing activator TraR as identified by repressor activity. Proc Natl Acad Sci U S A. 96:9009–9014.
Luo ZQ, Qin Y, Farrand SK. 2000. The antiactivator TraM interferes with the autoinducer-dependent binding of TraR to DNA by interacting with the C-terminal region of the quorum-sensing activator. J Biol Chem. 275:7713–7722.
MacLellan SR, Smallbone LA, Sibley CD, Finan TM. 2005. The expression of a novel antisense gene mediates incompatibility within the large *repABC* family of alpha-proteobacterial plasmids. Mol Microbiol. 55:611–623.
MacLellan SR, Zaheer R, Sartor AL, MacLean AM, Finan TM. 2006. Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. Mol Microbiol. 59:1559–1575.

Mavingui P, et al. 2002. Dynamics of genome architecture in *Rhizobium* sp. strain NGR234. J Bacteriol. 184:171–176.
McAnulla C, Edwards A, Sanchez-Contreras M, Sawers RG, Downie JA. 2007. Quorum-sensing-regulated transcriptional initiation of plasmid transfer and replication genes in *Rhizobium leguminosarum* biovar *viciae*. Microbiology 153:2074–2082.
Minogue TD, Wehland von Trebra M, Bernhard F, Beck von Bodman S. 2002. The autoregulatory role of EsaR, a quorum-sensing regulator in *Pantoea stewartii* ssp. *stewartii*: evidence for a repressor function. Mol Microbiol. 44:1625–1635.
Montoya AL, Moore LW, Gordon MP, Nester EW. 1978. Multiple genes coding for octopine-degrading enzymes in *Agrobacterium*. J Bacteriol. 136:909–915.
Moriguchi K, et al. 2001. The complete nucleotide sequence of a plant root-inducing (Ri) plasmid indicates its chimeric structure and evolutionary relationship between tumor-inducing (Ti) and symbiotic (Sym) plasmids in Rhizobiaceae. J Mol Biol. 307:771–784.
Nei M, Kumar S. 2000. Molecular evolution and phylogenetics. New York: Oxford University Press.
Novick RP, et al. 1976. Uniform nomenclature for bacterial plasmids: a proposal. Bacteriol Rev. 40:168–189.
Novick RP, Hoppensteadt FC. 1978. On plasmid incompatibility. Plasmid 1:421–434.
Oger P, Farrand SK. 2002. Two opines control conjugal transfer of an *Agrobacterium* plasmid by regulating expression of separate copies of the quorum-sensing activator gene *traR*. J Bacteriol. 184:1121–1131.
Oger P, Kim K-S, Sackett RL, Piper KR, Farrand SK. 1998. Octopine-type Ti plasmids code for a mannopine-inducible dominant-negative allele of *traR*, the quorum-sensing activator that regulates Ti plasmid conjugal transfer. Mol Microbiol. 27:277–288.
Pappas KM, Winans SC. 2003. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. Mol Microbiol. 48:1059–1073.
Pérez-Oseguera Á, Cevallos MA. 2013. RepA and RepB exert plasmid incompatibility repressing the transcription of the *repABC* operon. Plasmid 70:362–376.
Petit A, et al. 1983. Further extension of the opine concept: plasmids in *Agrobacterium rhizogenes* cooperate for opine degradation. Mol Gen Genet. 190:204–214.
Pinto UM, Pappas KM, Winans SC. 2012. The ABCs of plasmid replication and segregation. Nat Rev Microbiol. 10:755–765.
Piper KR, Beck von Bodman S, Farrand SK. 1993. Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. Nature 362:448–450.
Piper KR, Farrand SK. 2000. Quorum sensing but not autoinduction of Ti plasmid conjugal transfer requires control by the opine regulon and the antiactivator TraM. J Bacteriol. 182:1080–1088.
Qin Y, et al. 2000. Quorum-sensing signal binding results in dimerization of TraR and its release from membranes into the cytoplasm. EMBO J. 19:5212–5221.
Ramírez-Romero MA, Soberón N, Pérez-Oseguera A, Téllez-Sosa J, Cevallos MA. 2000. Structural elements required for replication and incompatibility of the *Rhizobium etli* symbiotic plasmid. J Bacteriol. 182:3117–3124.
Rzhetsky A, Nei M. 1992. A simple method for estimating and testing Minimum-Evolution trees. Mol Biol Evol. 9:945.
Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol. 4:406–425.
Sneath PHA, Sokal RR. 1973. Numerical taxonomy: the principles and practice of numerical classification. San Francisco: W. H. Freeman.
Subramaniam S. 1998. The biology workbench—a seamless database and analysis environment for the biologist. Proteins 32:1–2.
Suzuki K, et al. 2000. Complete nucleotide sequence of a plant tumor-inducing Ti plasmid. Gene 242:331–336.
Szegedi E, Czakó M, Otten L. 1996. Further evidence that the vitopine-type pTi's of *Agrobacterium vitis* represent a novel group of Ti plasmids. Mol Plant Microbe Interact. 9:139–143.
Szegedi E, Otten L. 1998. Incompatibility properties of tartrate utilization plasmids derived from *Agrobacterium vitis* strains. Plasmid 39:35–40.
Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol Biol Evol. 30:2725–2729.
Tempé J, Petit A, Farrand S. 1984. Induction of cell proliferation by *Agrobacterium tumefaciens* and *A. rhizogenes*: a parasite's point of view. In: Verma DPS, Hohn T, editors. Genes involved in microbe-plant interactions. Vienna (Austria): Springer. p. 271–286.
Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
Torres Tejerizo G, et al. 2014. Rhizobial plasmid pLPU83a is able to switch between different transfer machineries depending on its genomic background. FEMS Microbiol Ecol. 88:565–578.
Tsai CS, Winans SC. 2011. The quorum-hindered transcription factor YenR of *Yersinia enterocolitica* inhibits pheromone production and promotes motility via a small non-coding RNA. Mol Microbiol. 80:556–571.
Van den Eede G, Deblaere R, Goethals K, Van Montagu M, Holster M. 1992. Broad host range and promoter selection vectors for bacteria that interact with plants. Mol Plant Microbe Interact. 5:228–234.
Vannini A, et al. 2002. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. EMBO J. 21:4393–4401.
Venkova-Canova T, Soberón NE, Ramírez-Romero MA, Cevallos MA. 2004. Two discrete elements are required for the replication of a *repABC* plasmid: an antisense RNA and a stem-loop structure. Mol Microbiol. 54:1431–1444.
Wetzel ME, Kim KS, Miller M, Olsen GJ, Farrand SK. 2014. Quorum-dependent mannopine-inducible conjugative transfer of an *Agrobacterium* opine-catabolic plasmid. J Bacteriol. 196:1031–1044.
White CE, Winans SC. 2007. Cell-cell communication in the plant pathogen *Agrobacterium tumefaciens*. Philos Trans R Soc Lond B Biol Sci. 362:1135–1148.
White FF, Nester EW. 1980. Relationship of plasmids responsible for hairy root and crown gall tumorigenicity. J Bacteriol. 144:710–720.
Zhang L, Murphy PJ, Kerr A, Tate ME. 1993. *Agrobacterium* conjugation and gene regulation by N-acyl-$_L$-homoserine lactones. Nature 362:446–448.
Zhang RG, et al. 2002. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. Nature 417:971–974.
Zhu J, et al. 2000. The bases of crown gall tumorigenesis. J Bacteriol. 182:3885–3895.
Zhu J, Winans SC. 1998. Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated, dominant defective TraR-like protein. Mol Microbiol. 27:289–297.
Zhu J, Winans SC. 1999. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters *in vitro* and decreases TraR turnover rates in whole cells. Proc Natl Acad Sci U S A. 96:4832–4837.
Zhu J, Winans SC. 2001. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. Proc Natl Acad Sci U S A. 98:1507–1512.
Zuckerkandl E, Pauling L. 1965. Evolutionary divergence and convergence in proteins. In: Bryson V, Vogel HJ, editors. Evolution of genes and proteins. New York: Academic Press. p. 97–166.

Associate editor: Esther Angert