

TraA is required for megaplasmid conjugation in *Rhodococcus erythropolis* AN12

Joyce C. Yang^a, Philip A. Lessard^{a,b}, Neil Sengupta^b, Steven D. Windsor^b,
Xian M. O'Brien^a, Michael Bramucci^d, Jean-Francois Tomb^d,
Vasantha Nagarajan^d, Anthony J. Sinskey^{a,b,c,*}

^a Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^b Microbial Genetics Project Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^c Division of Health Sciences and Technology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^d DuPont Central Research and Development, Biochemical Sciences and Engineering, P.O. Box 80328, Wilmington, DE 19880-0328, USA

Received 16 June 2006, revised 1 August 2006

Available online 25 September 2006

Communicated by Ellen Zechner

Abstract

Pulsed-field gel electrophoresis (PFGE) revealed three previously uncharacterized megaplasms in the genome of *Rhodococcus erythropolis* AN12. These megaplasms, pREA400, pREA250, and pREA100, are approximately 400, 250, and 100 kb, respectively, based on their migration in pulsed-field gels. Genetic screening of an AN12 transposon insertion library showed that two megaplasms, pREA400, and pREA250, are conjugative. Mobilization frequencies of these AN12 megaplasms to recipient *R. erythropolis* SQ1 were determined to be approximately 7×10^{-4} and 5×10^{-4} events per recipient cell, respectively. It is known for other bacterial systems that a relaxase encoded by the *traA* gene is required to initiate DNA transfer during plasmid conjugation. Sequences adjacent to the transposon insertion in megaplasmid pREA400 revealed a putative *traA*-like open reading frame. A targeted gene disruption method was developed to generate a *traA* mutation in AN12, which allowed us to address the role of the *traA* gene product for *Rhodococcus* megaplasmid conjugation. We found that the AN12 *traA* mutant is no longer capable of transferring the pREA400 megaplasmid to SQ1. Furthermore, we confirmed that the conjugation defect was specifically due to the disruption of the *traA* gene, as pREA400 megaplasmid conjugation defect is restored with a complementing copy of the *traA* gene.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Megaplasmid; Conjugation; *Rhodococcus*; Relaxase; Nickase; Transfer; Gene targeting; Pulsed-field gel electrophoresis

1. Introduction

Rhodococcus sp. bacteria are Gram-positive actinomycetes that possess a variety of biochemical and

metabolic properties relevant to environmental and industrial microbiology (de Carvalho and da Fonseca, 2005; van der Geize and Dijkhuizen, 2004). Gene clusters that enable these unique enzymatic pathways often reside on large episomal elements called megaplasms. *Rhodococcus* megaplasms exhibit either circular or linear topology, and range

* Corresponding author. Fax: +1 617 253 8550.

E-mail address: asinskey@mit.edu (A.J. Sinskey).

from less than 50 kb to greater than one megabase-pairs in length (Priefert et al., 2004; Shimizu et al., 2001; Stecker et al., 2003; Trott et al., 2002). Though megaplasids are common within this genus of actinomycetes, this type of extrachromosomal replicon is not exclusive to rhodococci. Large plasmids (>50 kb) have been found in Gram-negative bacteria (Argandona et al., 2003; Pedraza and Diaz Ricci, 2002) and other types of Gram-positive bacteria (Kinashi et al., 1987; Scholle et al., 2003).

Molecular details of plasmid conjugation, involving both *cis*- and *trans*-acting elements, have been well-characterized for many Gram-negative bacteria (Lanka and Wilkins, 1995). For these plasmids, the DNA processing is initiated in *cis* at the origin of transfer (*oriT*) by a *trans*-acting protein complex called the relaxosome. The core enzyme of this complex is called a relaxase, and it cleaves one strand of the plasmid at the *nic* site within the *oriT* via a transesterification reaction. The nicked plasmid is then unidirectionally transferred to the recipient cell as a single-stranded DNA intermediate. In plasmid F, the relaxase is encoded by a gene called *traI* (Matson and Ragonese, 2005; Traxler and Minkley, 1988). This mechanism of plasmid conjugation appears to be conserved in many replicons isolated from Gram positive bacteria. In plasmid pIP501, which is found in the Gram-positive bacterium *Streptococcus agalactiae*, the relaxase is encoded by a gene called *traA* (Kopeck et al., 2005; Kurenbach et al., 2002).

Several different *Rhodococcus* megaplasids are known to be conjugative (Dabrock et al., 1994; Desomer et al., 1988; Kalkus et al., 1993; Priefert et al., 2004; Shimizu et al., 2001). The sequences and gene annotations of five rhodococcal megaplasids-p33701 from *Rhodococcus equi* ATCC33701 (Takai et al., 2000), pBD2 from *Rhodococcus erythropolis* BD2 (Stecker et al., 2003), pRHL3 from *Rhodococcus* sp. RHA1 (Warren et al., 2004), as well as pREC1 and pREL1 from *R. erythropolis* PR4 (Sekine et al., 2006)- were recently completed. Analysis of pREC1 (Genbank Accession No. NC_007486) revealed that pREC1 Orf16 encodes a putative relaxase (Sekine et al., 2006). This suggests that a single-stranded DNA transfer system similar to other bacterial conjugation systems may function in the transfer of *Rhodococcus* megaplasids. However, no functional analysis for this *Rhodococcus* TraA-like relaxase has been completed.

AN12 was first isolated and so named for its ability to use the aromatic compound, aniline, as a

carbon source. Besides the initial characterization of its small cryptic plasmid, pAN12 (Kostichka et al., 2003), it was not known whether the genome of AN12 harbored other extrachromosomal replicons, nor whether any of its replicons are transmissible via bacterial conjugation. We show in this study that AN12 possesses at least three distinctly migrating species of megaplasids, and that at least two of the AN12 megaplasids (pREA400 and pREA250) can be mobilized to a closely related *R. erythropolis* strain, SQ1. A putative *traA* open reading frame was found on the AN12 megaplasmid, pREA400. We demonstrate a role of this putative relaxase in *Rhodococcus* megaplasmid transfer, as a *traA* mutant was unable to transfer pREA400 to SQ1.

2. Materials and methods

2.1. Bacterial strains and culturing conditions

Bacterial strains, plasmids, and primers used in the present study are summarized in Table 1. Both strains of *R. erythropolis* AN12 and SQ1, as well as strains of *Escherichia coli*, were grown in LB liquid or on LB plates with 2% agar (Sambrook and Russell, 2001) supplemented with the following antibiotics purchased from Sigma-Aldrich (St. Louis, MO) as appropriate; gentamicin (Gm, 10 µg/ml), kanamycin (Km, 100 µg/ml), rifampicin (Rf, 20 µg/ml), and streptomycin (Sm, 150 µg/ml). *R. erythropolis* AN12 and SQ1 cells were cultivated at 30 °C, while *E. coli* strains were cultivated at 37 °C. All liquid culture flasks were shaken on an orbital shaker at 120 rpm, and small (<5 ml) volumes of liquid cultures were agitated using a roller drum. Frozen stocks of each strain were prepared by mixing equal volumes of saturated liquid cultures of bacteria and sterile 40% glycerol, then storing cells at –80 °C until use. All experiments were conducted with colonies that had been cultured for fewer than ten days from the initial frozen stock inoculum.

2.2. Preparation and standard transformation of electrocompetent AN12 cells

AN12 electrocompetent cells were prepared essentially as previously described (Kostichka et al., 2003), except cells were grown in NBYE/0.05% Tween 80 media in a 1 L baffled flask with shaking until OD₆₀₀ of about 0.5 was reached. Standard transformations with plasmids capable of replicating in *Rhodococcus* was achieved by incubating 0.5 µg of transforming DNA in 1× TE buffer with 100 to 150 µl competent AN12 cells for 5 min prior to electroporation. This suspension was placed in an ice-cold sterile 2 mm gapped electroporation cuvette. Electroporation was carried out at 2.5 kV, 25 µF, 400 Ω, with a Bio-Rad

Table 1
Bacterial strains, plasmids, and primers used in this study

Strain, plasmid, or primer name	Description or sequence	Source or reference
<i>Strains</i>		
DH5 α	<i>Escherichia coli</i> ; supE44 Δ lacU169 (ϕ 80 lacZ Δ M15) recA1 endA1 hsdR17 thi-1 gyrA96 relA1	Invitrogen (Carlsbad, CA)
EC100D pir-116	<i>Escherichia coli</i> ; F [−] mcrA Δ (mrr-hsdRMS-mcrBC) ϕ 80dlacZ Δ M15 Δ lacX74 recA1 endA1 araD139 Δ (ara, leu)7697 galU galK λ [−] rpsL nupG pir-116(DHFR)	Epicentre (Madison, WI)
TOP10	<i>Escherichia coli</i> ; F-mcrA Δ (mrr-hsdRMS-mcrBC) Φ 80 lacZ Δ M15 Δ lacX74 recA1 deoR araD139 Δ (ara-leu)7697 galU galK rpsL (Sm ^R) endA1 nupG	Invitrogen
AN12	Environmental isolate of <i>Rhodococcus erythropolis</i>	Kostichka et al. (2003)
SQ1	Environmental isolate of <i>R. erythropolis</i> ; Rf ^R Sm ^R	ATCC4277-1
AN12PL	AN12 derivative	This study
AN12PL-1F6	<i>Rhodococcus erythropolis</i> AN12PL derivative recovered from transformation with EZTn transposome; (pREA400::EZTn); Km ^R	This study
AN12-5A6	<i>Rhodococcus erythropolis</i> AN12 derivative recovered from transformation with EZTn transposome; (pREA250::EZTn); Km ^R	This study
JY524	<i>Rhodococcus erythropolis</i> SQ1 derived transconjugant recovered from mating with AN12-1F6; (pREA400::EZTn); Km ^R Rf ^R Sm ^R	This study
JY640	<i>Rhodococcus erythropolis</i> SQ1 derived transconjugant recovered from mating with AN12-5A6; (pREA250::EZTn); Km ^R Rf ^R Sm ^R	This study
JY825	<i>Rhodococcus erythropolis</i> AN12 transformed with pJY37 to disrupt the pREA400 megaplasmid encoded <i>traA</i> ; Km ^R	This study
JY893	JY825 derivative transformed with pJY48; Km ^R Gm ^R	This study
JY926	JY825 derivative transformed with pJY49B; Km ^R Gm ^R	This study
JY953	<i>Rhodococcus erythropolis</i> SQ1 derived transconjugant recovered from mating with JY926; Km ^R Gm ^R Rf ^R Sm ^R	This study
<i>Plasmids</i>		
pEZTn1F6	Plasmid containing the EZTn transposome and flanking AN12 genomic sequences isolated from AN12-1F6; Km ^R	This study
pEZTn5A6	Plasmid containing the EZTn transposome and flanking AN12 genomic sequences isolated from AN12-5A6; Km ^R	This study
pREA400	Endogenous AN12 megaplasmid	This study
pREA250	Endogenous AN12 megaplasmid	This study
pREA100	Endogenous AN12 megaplasmid	This study
pAL349	RP4 <i>mob</i> PCR product amplified from pSUP301 as a <i>MfeI</i> - <i>PstI</i> fragment, then ligated into pAL298 digested with <i>EcoRI</i> and <i>PstI</i>	Simon et al. (1983); Lessard et al. (2004); This study
pJY37	Plasmid used for <i>traA</i> targeted disruption; primers JYP560 and JYP561 were used to amplify a 706 bp fragment from AN12 genomic DNA, then cloned into pCR2.1Topo	This study
pJY48	<i>FspI</i> deletion of pAL349	This study
pJY49B	4900 bp <i>XmnI</i> fragment isolated from pEZTn1F6 containing the entire <i>traA</i> ORF cloned into pAL349 digested with <i>FspI</i>	This study
<i>Primers</i>		
JYP543	Reverse primer to amplify pCR2.1Topo sequences; 5'-ACGCCAAGCTTGGTACCGAGC-3'	This study
JYP557	Forward primer to amplify AN12 genomic DNA sequences; 5'-GTGGCCATGTGATGACCCTCC-3'	This study
JYP560	Forward primer to amplify AN12 genomic DNA sequences; 5'-GGCGATGGCTACGAGTACCTG-3'	This study
JYP561	Reverse primer to amplify AN12 genomic DNA sequences; 5'-CGGGTGAAGCACGCTTCCTTC-3'	This study
JYP576	Reverse primer to amplify AN12 genomic DNA sequences; 5'-CGCCACTTGCCGCTCTTTTCC-3'	This study

Gene Pulser (Hercules, CA). Immediately following pulsing, 200 μ l of LB was added to the cells. Cells were allowed to recover after electroporation at 30 °C for 2 h with gentle agitation prior to selection with appropriate antibiotics.

2.3. AN12 transposon library construction and screening

Two modifications to the above standard transformation method were made to generate the AN12 transposon

library. First, 20 ng (1 µl) of EZ::TN (R6Kγ/KAN-2) transposon (Epicentre, Madison, WI) was used per transformation. Second, cells were allowed to recover for 24 h prior to plating on selective media. Genomic DNA was prepared from a subset of kanamycin resistant AN12 mutants, which was screened using Southern blotting techniques to ensure that transposon insertion sites were relatively random (data not shown). The matings to *R. erythropolis* SQ1 were conducted by pooling 4 or 5 individual donor mutant cultures and incubating with an equivalent volume of SQ1 recipient culture, first on LB agar plates for 24 h, then on selective media. Few or no putative triply resistant (Km^R Rf^R Sm^R) transconjugants were detected when the majority of mutants were mated to SQ1. However, two pools of mutants from the library gave rise to >10,000 CFUs on the selective medium upon mating to SQ1. Independent matings of the individual mutants in these two pools revealed that a single-mutant donor in each case was responsible.

2.4. Site-specific gene disruption transformations of AN12

The *traA* disruption plasmid, pJY37, replicates in *E. coli* but not in *Rhodococcus*. Plasmid integration in *Rhodococcus* was selected using the plasmid-borne kanamycin resistance marker. pJY37 also harbors a 706 bp *traA* homology cassette starting from 25 bp of the coding sequence (nt. 2710 to 3416 of Genbank Accession No. DQ489716). Within this cassette, there is an endogenous *BspEI* site used to prelinearize the plasmid. Digestion of pJY37 with *BspEI* leaves 430 and 273 bp of *traA* homology upstream and downstream of the *BspEI* site, respectively. To generate the *traA* mutant, 2 µg of pJY37 was treated with *BspEI* in a 20 µl digestion reaction, then 1 µl of this reaction was used for each transformation of AN12. The transformation was then carried out as described previously except after electroporation, cells were allowed to recover for 24 h prior to plating on selective media.

2.5. *Rhodococcus* conjugation and megaplasmid mobilization frequency

Matings were carried out on solid LB agar surfaces as previously described (Lessard et al., 2004). The plasmid mobilization frequencies were determined by plating appropriate serial dilutions of the mating cell resuspensions (consisting of 1:1 mixtures of donor and recipient cultures) onto selective media, and plating control recipient culture dilutions to obtain recipient viable counts on non-selective media. Mating efficiencies were calculated as transconjugant CFU per recipient viable cell counts.

2.6. DNA manipulation and plasmid construction

Plasmids used in the present study, as well as a brief description of the cloning strategies used for the con-

struction of each, are summarized in Table 1, except for pEZTn1F6 and pEZTn5A6, which are described below. All DNA modifying enzymes and DNA size ladders were purchased from New England Biolabs (Beverly, MA) and used according to manufacturer's instructions. All PCR primers used in this study were purchased from Integrated DNA Technologies (Coralville, IA). *Rhodococcus* genomic DNA was prepared as previously described (Lessard et al., 2004). *Rhodococcus* plasmid DNA was isolated using the procedures previously described (Kostichka et al., 2003). pEZTn1F6 and pEZTn5A6 were generated via a plasmid rescue approach by digesting 2 µg total genomic DNA prepared from AN12PL-1F6 and AN12-5A6, respectively, with *EcoRI* for 2 h, followed by incubation at 65 °C for 30 min to abolish *EcoRI* activity. The DNA fragments were ligated with T4 DNA ligase, supplemented with 1 mM ATP (Sigma–Aldrich). 1 µl of the ligation reaction was used to transform EC100 *pir*-116 (Epicentre) cells. Transformed *E. coli* harboring recircularized plasmids containing transposons and associated AN12 sequences were selected with LB Km media.

2.7. Pulsed-field gel electrophoresis (PFGE)

Cell pellets from saturated overnight 2.5 ml LB cultures were frozen at –80 °C for >1 h, then resuspended in 1 ml of resuspension buffer (10 mM Tris, pH 8.0, 1 mM EDTA, and 10% Triton X-100). Resuspended cells were incubated at 30 °C for 2 h with gentle rocking, pelleted, washed with 1 ml wash buffer (0.2 M NaCl, 10 mM Tris, pH 8.0, and 100 mM EDTA), and pelleted again. To normalize, cell pellets were resuspended in the appropriate amount of wash buffer to a concentration of [200 µg/µl] wet cell weight per volume buffer. This suspension was pre-warmed at 42 °C before mixing with an equal volume of molten 1% Pulsed-Field Gel Certified Agarose (Bio-Rad) in 1× TBE, yielding a final cell concentration of 100 µg/µl. Eighty microliters of this mixture was quickly dispensed into a disposable plug mold (Bio-Rad). The 0.5% agarose plugs containing embedded bacterial cells were incubated at 4 °C for 15 min to ensure solidification. Then, individual plugs were incubated with 1 ml of enzyme lysis solution (10 mM Tris, pH 8.0, 50 mM NaCl, and 100 mM EDTA) containing fresh lysozyme [5 mg/ml] and mutanolysin [200 U/ml] (Sigma–Aldrich) for 2 h at 37 °C in microfuge tubes on a rocking platform. The supernatant was removed, and the plugs were treated with 1 ml of detergent lysis solution (10 mM Tris, pH 8.0, 50 mM NaCl, 100 mM EDTA, 4.8 mM sodium deoxycholate, and 1.7 mM *N*-lauryl sarcosine) overnight at 37 °C on a rocking platform. The supernatant was discarded and plugs washed once with washing buffer. Finally, the plugs were incubated with 1 ml of digestion buffer (10 mM Tris, pH 8.0, 50 mM EDTA, and 3.4 mM *N*-lauryl sarcosine) containing fresh proteinase K [0.5 mg/ml] overnight at 50 °C on a rocking platform. DNA species were resolved using the Bio-Rad

CHEF-DR II PFGE apparatus with 1% agarose gel (Bio-Rad) in $1\times$ TBE at 14°C , 5 V/cm, included angle of 120° for 24 h with 30 s initial and 60 s final switch times. The lambda ladder PFG marker (NEB) was used to estimate replicon size.

2.8. Southern blot analysis

For standard hybridizations, 2–5 μg genomic DNA was digested with appropriate restriction enzymes, then separated on agarose gels. Gels were subjected to depurination with 0.25 M HCl for 40 min, denaturation with 0.5 N NaOH for 30 min, and neutralized with Tris–Cl for 30 min. DNA was then transferred onto positively charged nylon membranes (Roche Diagnostics Corp., Indianapolis, IN) using $20\times$ SSC as the transfer buffer for 24 h. This step was extended to 48 h to accommodate for the transfer of the large DNA replicons resolved using PFGE. DIG-11 dUTP labeled probes were generated, and hybridizations were carried out by using reagents in the DIG-High Prime DNA Labeling and Detection Starter Kit per manufacturer's instruction. Hybridizing species were detected using Kodak Biomax Light scientific imaging film.

2.9. *Rhodococcus erythropolis* colony PCR

Approximately 100 μg of cells were collected using a pipet tip, then resuspended in a 50 μl PCR reaction using all of the reagents and protocols from the HotStarTaq Polymerase Kit (Qiagen Sciences, Valencia, CA). DNA was amplified in a PTC-200 Cycler (Bio-Rad, Waltham, MA) using cycling parameters of 1 cycle of 15 min at 94°C , followed by 30 cycles of 30 s at 94°C , 30 s at 55°C , and 1 min at 72°C , followed by an added extension time of 10 min at 72°C . Annealing temperatures and extension times were adjusted to optimize performance of each primer pair.

3. Results

3.1. Discovery of AN12 extrachromosomal elements

It was shown previously that *R. erythropolis* AN12 possesses at least one extrachromosomal circular replicon of ca. 6.3 kb, called pAN12 (Kostichka et al., 2003). Preliminary experiments in our lab suggested that at least two different species of megaplasmids also exist in the AN12 genome. Pulsed-field gel (PFG) electrophoresis parameters were further optimized to examine the megaplasmid species of AN12 and *R. erythropolis* SQ1. By modifying two conditions in the protocol, namely, the lysis step and the concentration of the embedding agarose, both the yield and resolution of these large replicons were significantly enhanced. An

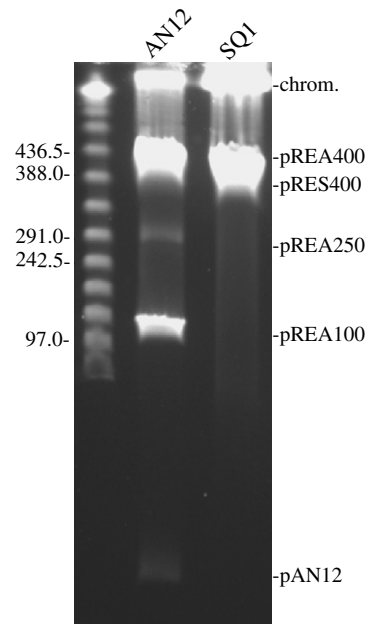


Fig. 1. Pulsed-field gel (PFG) profiles of wild-type AN12 and SQ1 replicons. Approximately 2 μg of the lambda ladder PFG marker were loaded to determine the migrating size of the megaplasmids. Chromosomal (chrom.) and distinguishable species of plasmids, as well as relevant ladder sizes (in kb) are labeled.

example of this is shown in Fig. 1. Three distinct megaplasmid species in AN12 can now be detected using these PFG conditions. These plasmids were named pREA400, pREA250, and pREA100, for *R. erythropolis* AN12, based on their approximate sizes in kb. One distinct megaplasmid migrating at 400 kb was observed in SQ1, consistent with what has been previously shown for this strain (Priefert et al., 2004). We propose to call this SQ1 megaplasmid, pRES400.

3.2. A genetic screen for transmissible elements in AN12

A few characterized *Rhodococcus* megaplasmids, such as pBD2 and pRHL2, have been shown to be conjugative at frequencies of approximately 1×10^{-4} event per recipient cell (Dabrock et al., 1994; Shimizu et al., 2001). To determine whether any AN12 replicons are transmissible, a commercially available transposon-mediated mutagenesis system (EZ::TN (R6K γ ori/Kan-2)) was used to tag genetic elements of AN12. By using the transposon-encoded kanamycin resistance marker, the selection of transconjugants need not rely on activities encoded on the DNA elements being transferred.

Thus theoretically, any cryptic but transmissible AN12 replicon should be detected using this library.

Two founding strains, AN12-wild-type (WT) and an AN12 derived strain, AN12PL, which lacks pREA100 and possibly pREA250, were used in the generation of a transposon library consisting of a total of 432 mutants. A subset of 119 mutants was then screened for the ability to mobilize the kanamycin resistance cassette and associated AN12 genomic DNA to SQ1, which is naturally resistant to rifampicin (Rf^R) and streptomycin (Sm^R) (Les-sard et al., 2004). Thus, SQ1-derived transconjugants can be selected on the basis of triple antibiotic resistance (Km^R Rf^R Sm^R). When used as donors, two mutants, AN12PL-1F6 and AN12-5A6, gave rise to Km^R Rf^R Sm^R transconjugants. Conjugal transfer of the tagged AN12PL-1F6 element was determined to occur at a frequency of 7.1×10^{-4} events per recipient SQ1 cell, and the tagged AN12-

5A6 element at 5.2×10^{-4} events per recipient cell (Table 2), comparable to the frequencies determined for pBD2 and pRHL2 (Dabrock et al., 1994; Shimizu et al., 2001).

3.3. AN12PL-1F6 and AN12-5A6 mutants bear transposon insertions in transmissible megaplasmids

Southern blots were used to verify that AN12-specific DNA was transferred from the AN12PL-1F6 and AN12-5A6 mutants to SQ1. DNA adjacent to each transposon insertion site was rescued (see Section 2) from AN12PL-1F6 and AN12-5A6 in the form of plasmids that we named pEZTn1F6 (Fig. 2A) and pEZTn5A6 (data not shown), respectively. These plasmids were used as templates to synthesize probes that were hybridized to fragmented genomic DNA from each donor, recipient and transconjugant. Results indicated that transposons and associated AN12 sequences were indeed derived from the donor strains, and mobilized to SQ1. Furthermore, no homologous sequences were detected in the “wild-type” SQ1 parent strain (data not shown).

Pulsed-field gel (PFG) electrophoresis was used as the initial approach to address whether the AN12PL-1F6 and AN12-5A6 mutants bear transposon insertions in megaplasmids. As shown in Fig. 2B, the transconjugant (JY640) from the AN12-

Table 2
Summary of mobilization frequency of AN12 megaplasmids to SQ1

Donor strain	No. trials	Weighted average frequency per SQ1 CFU	Standard deviation
AN12PL-1F6	5	7.1×10^{-4}	9.9×10^{-5}
AN12-5A6	4	5.2×10^{-4}	1.3×10^{-4}
JY825	4	2.2×10^{-8}	3.8×10^{-9}
JY893	3	3.2×10^{-8}	1.9×10^{-8}
JY926	4	1.4×10^{-5}	2.0×10^{-6}

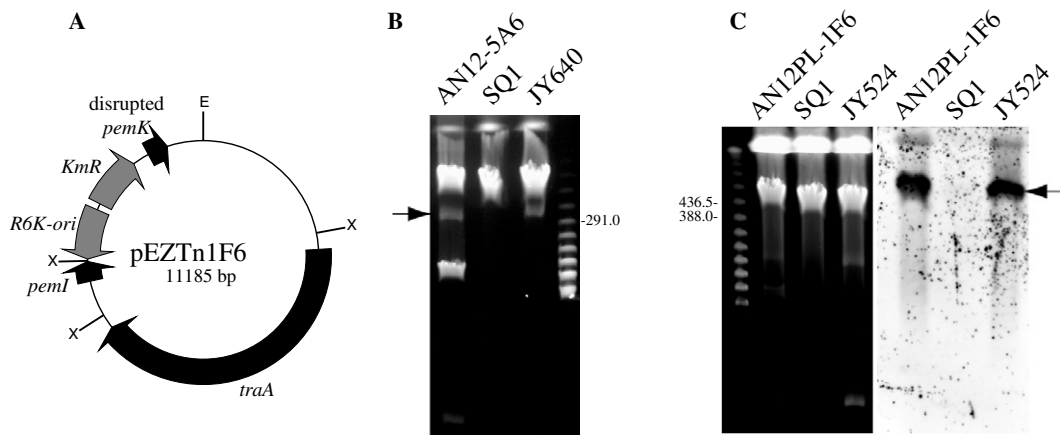


Fig. 2. Transposon-tagged megaplasmids can be transferred to SQ1. (A) Schematic diagram of the plasmid rescue product from recircularized genomic DNA isolated from AN12PL-1F6. Gray or black coloring is used to differentiate between ORFs encoded by the transposon and those derived from *Rhodococcus* megaplasmid, pREA400, respectively. The *EcoRI* site (E) used to recircularize this plasmid, as well as *XmnI* sites (X) used in the construction of a *traA* complementation plasmid, pJY49B, are indicated. (B) Pulsed-field gel (PFG) profiles of donor strain, AN12-5A6, recipient strain, SQ1, and SQ1-derived transconjugant, JY640. The AN12 replicon, pREA250, that has been mobilized to SQ1 is labeled with an arrow. (C) PFG (left) and subsequent Southern blotting (right) of extrachromosomal replicons. An arrow indicates hybridizing replicons from lysed AN12PL-1F6, SQ1, and JY524 (SQ1 transconjugant) cells. The surrounding sequences adjacent to the AN12PL-1F6 transposon insertion in the form of pEZTn1F6 served as the template for the DIG-labeled Southern probe. Relevant ladder bands have also been indicated with sizes in kbs.

5A6 \times SQ1 mating gained the pREA250 megaplasmid. PFG analysis showed that all other transconjugants isolated from this mating also gained the pREA250 element (data not shown). These data suggest that the AN12-5A6 insertion is in pREA250, and that pREA250 can be mobilized to SQ1.

PFG analysis alone could not discern mobilization of a megaplasmid from AN12PL-1F6 to SQ1. It was likely the transposon insertion of AN12PL-1F6 lies in pREA400, whose presence in the transconjugant was obscured by the endogenous SQ1 megaplasmid, pRES400. To determine whether this was so, a Southern blot of a PFG was performed to compare AN12PL-1F6, SQ1, and the JY524 transconjugant. Using the same pEZTn1F6-derived probe generated earlier, it was determined that, indeed, the probe hybridized strongly to pREA400 in the AN12PL-1F6 lane (Fig. 2C). No probe hybridized to DNA in the SQ1 lane. Moreover, a hybridizing megaplasmid of the same size was present in the JY524 transconjugant. These data strongly suggests that AN12PL-1F6 harbors a transposon insertion in the pREA400 megaplasmid, and that pREA400 can be mobilized to SQ1.

3.4. The AN12PL-1F6 transposon insertion is in *pemK*

Plasmid “addiction” systems ensure the stable inheritance of a plasmid through rounds of cell division by killing off cells that have lost the plasmid (Engelberg-Kulka and Glaser, 1999). Generally, a long-lived toxin protein is paired with a short-lived antitoxin (protein or mRNA). The antitoxin not only acts to repress toxin expression, but its binding to the toxin protein precludes toxin binding to essential cellular proteins. Loss of plasmid after cell division is concurrent with loss of antitoxin production, and the toxin then induces cell arrest. An ORF encoding a PemK-like toxin has been identified on the *Rhodococcus* megaplasmid, pBD2, suggesting that an analogous post-segregational killing mechanism may function in megaplasmid inheritance in actinomycetes (Stecker et al., 2003).

Analysis of pEZTn1F6 revealed that the transposon pREA400 inserted one basepair downstream of the start codon of a *pemK*-like toxin gene, likely disrupting its expression. We looked for the potential instability of the transposon tagged pREA400 megaplasmid in AN12PL-1F6. However, our preliminary evidence suggests that even without antibiotic selection for 10 consecutive days on LB plates,

all thirty independent AN12PL-1F6 single-colony isolates retained the ability to subsequently thrive on LB plates supplemented with kanamycin, suggesting that the tagged megaplasmid is inherited stably in each case. This may mean that the *Rhodococcus* PemK-like protein is not required for megaplasmid maintenance, or perhaps other selective pressures are at play, obscuring the potential activity of the PemK-like protein.

3.5. Sequence analysis of pREA400 encoded *traA*

One aim of this project was to find determinants of *Rhodococcus* megaplasmid conjugation. A candidate gene, *traA*, adjacent to the AN12PL-1F6 transposon insertion whose gene product was likely to be involved in this process was identified. Sequence analysis indicated that the pREA400 megaplasmid *traA* encodes a protein of 1506 amino acids (a.a.). Proteins putatively involved in plasmid transfer isolated from Gram-positive actinomycetes were among those most similar to the pREA400 TraA. Not surprisingly, the putative transfer protein encoded by *orf16* (Genbank Accession No. BAE46257) from the pREC1 megaplasmid in *R. erythropolis* PR4 (Sekine et al., 2006) is one of the closest matches to pREA400 TraA with 34% identity over 1049 residues. The TraA (Genbank Accession No. CAE09129) from *Gordonia westfalica* is encoded by a 101 kb megaplasmid called pKB1 (Broker et al., 2004), and is 31% identical to pREA400 TraA over 1079 residues. It is unclear whether the *traA* gene product (Genbank Accession No. AAS20144) from *Arthrobacter aureus* also resides on a megaplasmid (Sajjaphan et al., 2004), however, this protein is 29% identical to pREA400 TraA over 1005 residues. The *R. equi* ATCC33701 *traA* gene product (Genbank Accession No. AAG21733) encoded by the circular p33701 megaplasmid (Takai et al., 2000) is 33% identical to pREA400 TraA over 972 amino acid residues. The gene products of the following exhibit less homology to pREA400 TraA. The *tra* gene product (Genbank Accession No. NP_863184) from *Corynebacterium diphtheriae* is encoded by the 12 kb circular plasmid, pNG2 (Tauch et al., 2003), and aligns to two distinct portions of pREA400 TraA from a.a. 190 to 419 and from a.a. 590 to 976 with an overall 33% identity over 633 residues. The putative relaxase from *Streptomyces clavuligerus* (Genbank Accession No. AAQ93521) is encoded by pSCL2- a linear 200 kb megaplasmid (Wu et al., 2006) and

aligns to only the N-terminal portion of pREA400 TraA from a.a. 1 to 422 with 29% sequence identity.

A conserved domain (CD) search revealed a pREA400 TraA region (a.a. 545 to 964) that is similar to ATP-dependent RecD type helicases (Marchler-Bauer and Bryant, 2004). In addition, this region is 27% identical to TraA relaxase/helicase encoded by the *Agrobacterium tumefaciens* pTiC58 plasmid (Farrand et al., 1996), and 22% identical to the TraI relaxase/helicase of the *E. coli* F factor (Abdel-Monem et al., 1983; Larkin et al., 2005; Street et al., 2003). Alignments to the consensus helicase motifs (Hall and Matson, 1999) revealed that pREA400 TraA and its actinomycete relatives belong to the superfamily I (SFI) type of helicases (Fig. 3A). Nearly all of the residues are conserved except for some differences in SFI helicase motif III and IV. In these alignments (where +, o, and x denote hydrophobic, hydrophilic and any residues, respectively), the second hydrophobic residue in conserved motif III (++++GDxoQ) appears to be replaced by a hydrophilic arginine residue in most of these actinomycete helicases, and the third, hydrophobic residue in consensus motif IV (xx+xxoxR) has been replaced with a hydrophilic threonine residue.

Helicase activities are sometimes paired with relaxase activities in plasmid transfer proteins. It was not clear by performing BLAST and CD analysis over the entire length of the pREA400 TraA protein whether it encodes relaxase function. Functional studies and sequence analysis indicate that at least four families of relaxases-exemplified by founding members from plasmids F, RP4, RSF1010, and pMV158- exist that differ in amino acid identity over three well-conserved relaxase motifs (Zechner et al., 2000).

Based on what is known for these Gram-negative relaxases/helicases, we anticipated that the pREA400 TraA relaxase domain may be located at the N-terminus. Therefore, a partial alignment of pREA400 TraA a.a. 1 to 312 to sequences from the *Gordonia*, *Arthrobacter*, and *Streptomyces* TraA's using ClustalW analysis (Thompson et al., 1994) was completed (Figs. 3B and C). We found that pREA400 TraA is 43, 42, and 35% identical to pKB1 TraA, *Arthrobacter* TraA, and pSCL2 TraA, respectively, over this N-terminal region. Analysis of the first 60 amino acids of these proteins shows that these actinomycetes proteins most closely resemble each other, yielding new consensus motifs I (+T+Hx+x+GxGYxY+xxox+xxD) and IA (LxoYYxxxGxxPGxWxGxG+xx+) (Fig. 3B). More-

over, alignments to F TraI (Genbank Accession No. AAC44186), R388 TrwC (Genbank Accession No. CAA44853), RP4 TraI (Genbank Accession No. CAA38336), and pMV159 MobM (Genbank Accession No. CAA33713) reveal that, although the pREA400 TraA and related relaxases show certain similarities to RP4 TraI and pMV158 MobM (Fig. 3B), they more closely resemble F-factor TraI and R388 TrwC relaxases, especially with respect to relaxase motifs II and III (Fig. 3C). Recently, a crystal structure of the relaxase domain of F factor TraI protein was reported, and key residues (Y16, D81, H146, H157, and H159) of the active site were determined (Larkin et al., 2005). Importantly, all of these catalytic residues of F factor TraI, in addition to the arginine (R150) residue required for substrate binding, are absolutely conserved in these actinomycete relaxases (Figs. 3B and C).

Taken with the alignment to helicases, we propose that pREA400 TraA, like the F factor TraI, is a consolidated relaxase/helicase conjugation initiating protein. While this combination of relaxase/helicase appears to be rare for Gram-negative transfer proteins (Byrd et al., 2002), it appears to be the rule in actinomycetes since the *R. erythropolis* Orf16, *R. equi* TraA, *Arthrobacter* TraA, *Corynebacterium* Tra, and *Gordonia* TraA proteins all appear to be consolidated relaxase/helicases, with the *Streptomyces* TraA being an exception as it appears to only encode relaxase activity.

3.6. Targeted gene disruption of *traA*

A gene disruption strategy based on double-stranded DNA break repair for homologous recombination was developed to address the function of pREA400 *traA* in megaplasmid conjugation. It is well known that double-stranded DNA breaks can trigger homologous recombination at the site of the lesion (Cox, 1999; van den Bosch et al., 2002). This phenomenon was first exploited to achieve genetic transformations of *Saccharomyces cerevisiae* (Orr-Weaver et al., 1981). The approach we adopted (see Section 2) to generate a *traA* disruption is similar to the yeast model, and is shown in Fig. 4. A colony PCR assay was developed to quickly screen whether the drug resistant transformants arose by homologous recombination at the *traA* locus or illegitimately elsewhere (Fig. 4). Only recombinants that integrated homologously at the *traA* locus should give the expected 800 bp PCR product. Of the one hundred transformants analyzed with colony PCR,

A	R. eryth. pREA400_TraA	610	VGVPAGTGKT	620	637	VIALAPSA	644	691	LLVDEAS	697	718	VRLLGDPNQ	726
	R. equi p33701_TraA	537	VAVGAAGAGKT	547	564	VVALGPSA	571	609	LLVDEAG	615	636	VRLLGDPQQ	644
	R. eryth. pREC1_Orf16	600	VAVGAAGAGKT	610	627	VIALGPSA	634	672	LLVDEAG	678	699	VRLLGDPQLQ	707
	A. aures. TraA-like	595	VGIGAAGTGKS	605	622	VIGLAPSA	629	677	VIVDEAG	683	704	VRLIGDDRQ	712
	G. westf. pKB1_TraA	627	LVLAPAGTGKT	637	654	LIGLAPTA	661	745	VVVDEAG	751	772	VRLVGDDQQ	780
	C. diph. pNG2_Tra	466	VVVAPAGAGKT	476	493	VIGLAPTG	500	545	VVVDEAG	551	572	VVFVGDPPEQ	580
	CONSENSUS SFI Helicase	(I)	++xGxAGoGKS	(IA)	xx+xxxoo	(II)	+++DExo	(III)	+++GDxoo				
			A P T										
	R. eryth. pREA400_TraA	814	LSTRDDVR	821	922	AATVHRSQGIVTDS	936	951	VATTRGRT	958			
	R. equi p33701_TraA	731	APTNDLVR	738	839	ATTVHRSQGMTVGSC	853	868	VAMTRGKN	875			
	R. eryth. pREC1_Orf16	794	APTNDLVR	801	902	AATVHRSQGMTVGAC	916	931	VAMTRGKQ	938			
	A. aures. TraA-like	802	APTNESSQ	809	908	ALTTHRAQGMTRDAG	922	937	VAATRGRD	944			
	G. westf. pKB1_TraA	867	AATRDVTN	874	972	ASTIHAAQGMTADTC	986	1001	VAMTRGRH	1008			
	C. diph. pNG2_Tra	673	AATRDQVT	680	778	AATGHSSQCATVDVA	792	808	VPMTRGRE	815			
	CONSENSUS SFI Helicase	(IV)	xx+xooxR	(V)	xxT+xxxQG+o+ooV	(VI)	VA+TRxoo						
B	pREA400_TraA	1	---MTLHVLHAGDGYEYLTSQVATADREREHGQELTDYYSAHGTPPGTWFGGGLDDLARM	58									
	A.aures._TraA-like	1	---MMSLHVLHAGDGYEYLTSQVATADREREHGQELTDYYSAHGTPPGTWFGGGLDDLARM	58									
	pKB1_TraA	1	---MTATIHKLTAGDGYEYLTSQVATADREREHGQELTDYYSAHGTPPGTWFGGGLDDLARM	60									
	pSCL2_TraA-like	1	---MISMSPPVRPGSGWRYLFRGVGMAGDGHRAPTSLRAAQDEAGVPPGVWKGGLAAVAVG---	56									
	CONSENSUS Relaxase	(I)	+T+Hx+x+GxGYxY+xoxx+xxD	(IA)	LxoYY+xxGxxPGWxGxG+xx+								
			S W										
	F_TraI	1	---MMSIAQVRSAGSAGNYYTD-----KDNYY-VLGSNGERWAGRGAEQLGLQG	45									
	R388_TrwC	1	MLSHMVLTRQDIGRAASYYED-----GADYY-AKGDGASEWQGGAEELGLSG	48									
	RP4_TraI	1	---MIAKHV..FAELVKYITD-----	25									
	pMV158_MobM	19	---AFKHNERV...LNYELT-----	47									
C	pREA400_TraA	195	GFDLVFTPPKSVSTMWALADDDLRRQIERIHHTVKDTLGWIEKEACFTRTGATSQEHQDT	255									
	A.aures._TraA-like	186	GYDLTFPPVKISIVLWALGDADTRRLVEDAQQAALNDSIEYLETHALATRLGTNGIAQSSV	246									
	pKB1_TraA	206	GYDMTFSPVKVSALWAVPLPMAEKIEAAHRQAVADALAFIEREACLSRLGTDGIAQVDT	266									
	pSCL2_TraA-like	126	AFDLVFRPSTAHIAWALMDDETTRVLEECQDTARDKTLAWLEESVAQIRWGSNGKHKRPV	186									
	F_TraI	79	GYDLTFSPKSVSMAMLG---DKRLIDAHNQAVDFAVRQVEALA-STRVMTDGGQSETVL	135									
	R388_TrwC	83	GLDLTFSPKSVSLQALVAG---DAEIIKAHDRAVARTLEQAEARA-QARQKIQKGTRIET	139									
C	pREA400_TraA	256	TG-VVATLYDHYDSRAGDPNLHTHAVLSVKVCTEKDG---KWRALFSSTLHRYGVSPASQRY	312									
	A.aures._TraA-like	247	KGGLTATAFRHHDSRLGDPNLHTHVVSNNKVQDLAG---NWKSIDGKLLHRSVAVVSEHY	303									
	pKB1_TraA	267	DG-LIAAAFTHRDSRAGDPDLHTHVAISNKVVRDAAITRWMLDGTPLFKATVSASEVY	326									
	pSCL2_TraA-like	187	RDGLIVTVFRHYESRAGQPQLHDAHVVSVIRARRPDDG---TWGNLSADSLANIVAADTLY	244									
	F_TraI	136	TGNLVMALFNHDSRQDEPQLHTHAVVANVTQHNG-----EWKTLSSDKVG-KTGFIEVNY	190									
	R388_TrwC	140	TGNLVIGKFRHETSRRERDPQLHTHAVILNMTKRSDG---QWRALKNDEIVKATRYLGAVY	196									

Fig. 3. Protein alignments of DNA helicase and relaxase domains of pREA400 TraA. (A) Alignments of actinomycete transfer proteins from the indicated strains or plasmids to the consensus motifs (I, IA, II, III, IV, V, and VI) of superfamily I-type helicases (Hall and Matson, 1999). Consensus hydrophobic, hydrophilic and non-specific helicase residues are indicated with +, o, and x, respectively. Common variants to the conserved residue are listed directly below the consensus motif. Aligning residues are indicated (gray boxes). (B) ClustalW alignments between the N-terminal amino acids from 1 to 312 of pREA400 TraA and related gene products are shown. The consensus relaxase motif I and motif IA derived from these alignments are shown below these proteins. These motifs were then compared and aligned to relaxases from F, R388, RP4 and pMV158, whose absolutely conserved residues (Zechner et al., 2000) are indicated (bold). Aligning amino acids residues to the actinomycete consensus motifs I and IA are highlighted with gray boxes. The active site tyrosine residues are indicated (*). (C) Alignments of various relaxase motifs II and III. Crystallographically determined (Larkin et al., 2005) active site residues of the relaxase domain of F plasmid (Y16, D81, H146, H157, and H159), and aligning residues have been indicated (bold). Also in bold are residues that align with R150 of the F plasmid, which is known to be important in DNA substrate binding (Harley et al., 2002). In all of these alignments, single a.a gaps are indicated with dashes, and larger gaps (> 10 a.a.) are indicated with periods. The numbers at the left and right of each line correspond to residue numbers of the first and last amino acids of that line in that particular sequence, respectively. GenBank accession numbers for these proteins are as follows; *R. erythropolis* pREA400 TraA (ABF48485), *R. equi* p33701 TraA (BAB16639), *R. erythropolis* pREC1 Orf16 (YP_345077), *A. aurescens* TraA (AAS20144), *G. westfalica* TraA (CAE09129), *C. diphtheriae* pNG2 Tra (NP_863184), *S. clavuligerus* pSCL2 TraA (AAQ93521), *E. coli* F TraI (AAC44186), *E. coli* R388 TrwC (CAA44853), *E. coli* RP4 TraI (CAA38336), and *Streptococcus agalactiae* pMV158 MobM (CAA33713).

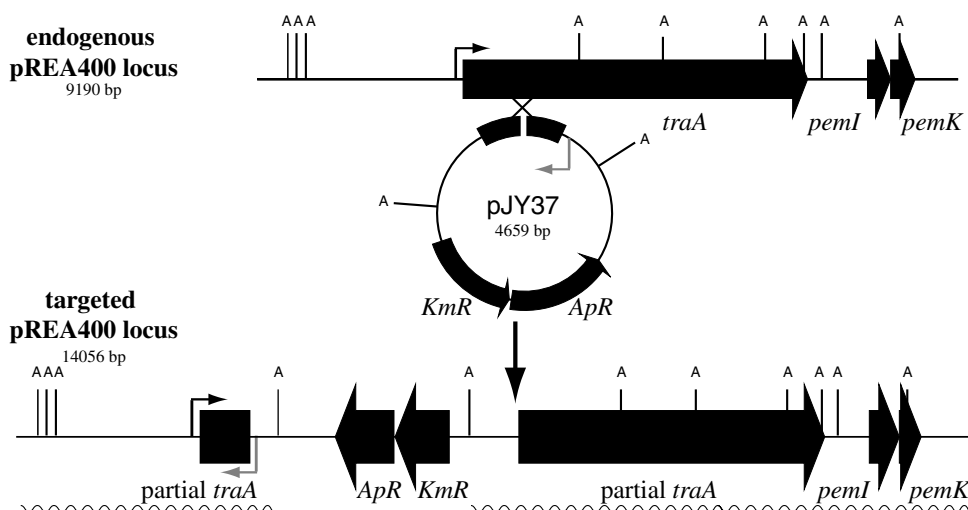


Fig. 4. Schematic diagram of the *Rhodococcus* targeted gene-disruption strategy. The *EcoRI* fragment (9190 bp) from pREA400 is shown with the *traA* ORF, *pemI* ORF, and *pemK* ORF. The integration plasmid, pJY37, was prelinearized at the unique *BspEI* site to induce homologous recombination. The targeted pREA400 locus (14,056 bp) is depicted following the integration of pJY37. Sites corresponding to *AflIII* recognition sequences are labeled. Sites where colony PCR primers (JYP557 and JYP543) anneal (A) are indicated (bent black and gray arrows, respectively). DNA species predicted to anneal to a Southern probe generated with pEZTn1F6 template DNA are indicated (wavy lines).

one gave a positive result (Fig. 5A). We named this candidate mutant strain, JY825.

We expected the gene disruption strategy to yield a partially duplicated *traA* locus, although neither of the partial *traA* fragments should express full-length protein. The first *traA* would contain the upstream regulatory sequence, as well as the start codon. However, the expression of this *traA* should yield a carboxy-terminally truncated protein, the length of which is predicted to be less than 250 a.a.. The second *traA* includes most of the ORF, but no start codon nor promoter, and should not be expressed. Southern blot analysis was used to verify the *traA* targeted locus. Genomic DNA samples from AN12 WT, JY825, AN12PL-1F6, and SQ1 were examined (Fig. 5B). We conclude that pREA400 *traA* is indeed disrupted as expected in JY825, because annealing fragments of the predicted sizes were detected in all sample lanes.

3.7. Phenotypic analysis of the *traA* mutant and genetic complementation of the *traA* defect

Either JY825 or AN12PL-1F6 (as positive control) was mated to SQ1 to determine whether the *traA* disruption affected the conjugal transfer of the tagged megaplasmid, pREA400. Triple drug (Km^R Rf^R Sm^R) resistances can be used to positively select for SQ1-derived transconjugants from both JY825

and AN12PL-1F6 donor strains, as both strains bear insertions of Km^R cassettes in pREA400. While the AN12PL-1F6 donor strain frequently gives rise to greater than 200,000 transconjugants per mating, JY825 gives rise to less than 30 transconjugants per mating (Table 2, Fig. 6C), a roughly 10,000-fold reduction. This strongly suggests that *traA* gene function is required for pREA400 megaplasmid conjugation.

To make certain that the defect in pREA400 transfer was specifically due to *traA* disruption and not due to the disruption of neighboring genes, such as the *pemI/pemK* locus, a complementation plasmid was constructed and introduced to the JY825 mutant strain. The pAL349 parent plasmid contains a *ColEI* replication system and the *aacCI* gene encoding gentamicin resistance (Gm^R). The entire *traA* ORF and flanking sequences was subcloned into pAL349, resulting in the complementation plasmid pJY49B. A modified pAL349 control plasmid, pJY48, was also generated at the same time (Table 1).

Both pJY48 and pJY49B were introduced to JY825 *traA* mutant and transformants were selected following a 24 hour recovery period. Since neither pJY48 nor pJY49B can replicate in *Rhodococcus*, the only way the cells could express the gentamicin resistance is if pJY48 or pJY49B had illegitimately or legitimately recombined into the genome, respec-

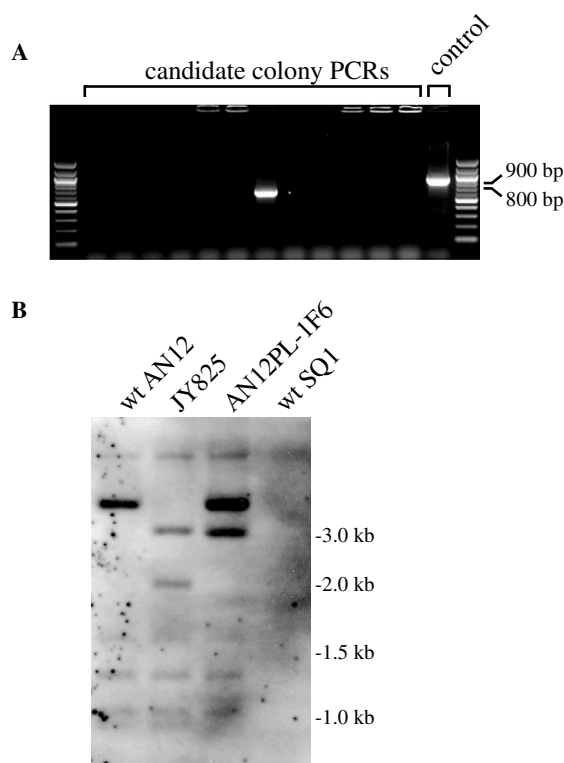


Fig. 5. Identification and verification of the candidate *traA* disruption mutant strain, JY825. (A) Colony PCR-based screening using primers JYP557 and JYP543 should yield an approximately 800 bp product if the *traA* locus is disrupted as expected. Each lane represents a kanamycin resistant AN12(pJY37) transformant. The control colony PCR reaction was done using primers JYP557 and JYP576 on wild-type AN12. This reaction is expected to yield ca. 900 bp product. The flanking lanes each contain 0.5 μ g 100 bp DNA ladder. (B) Southern blot. Genomic DNA fragments digested with *Afl*III from the indicated strains were hybridized to DIG-labeled probe made from pEZTn1F6 plasmid template. Expected hybridizing fragment sizes for DNA from each of the following strains are as follows: wild-type AN12—3588, 1342, 1100, and 1021 bp; JY825—3210, 2227, 1342, 1100, and 1021 bp; AN12PL-1F6—3588, 3022, 1342, and 1100 bp. No hybridizing bands were expected in wild-type SQ1.

tively. Greater than 500 Gm^R CFUs were observed when pJY49B was used as the transforming DNA, whereas only 3 Gm^R CFUs were observed when pJY48 was used. We named one non-specific pJY48 integrant JY893. We suspected that pJY49B would recombine preferentially at the second partial *traA* locus, as this region offered close to 4 kb of sequence homology (Fig. 6A). This was indeed the case for one *traA* complementing transformant, JY926, as confirmed by plasmid rescue experiments (data not shown) and Southern blot analysis (Fig. 6B).

The *traA* complementing strain, JY926, and the control strain, JY893, were tested for the ability to

give rise to triply resistant (Km^R Rf^R Sm^R) trans-conjugants when mated to SQ1 recipients. We do not yet know on what replicon the pJY48 plasmid integrated in the JY825 genome, however, it was clear that the empty vector could not restore pREA400 megaplasmid transmission (Fig. 6C). In contrast, the full-copy of *traA* in the JY926 strain does indeed complement the original *traA* disruption in JY825. The efficiency of JY926 to transfer pREA400 is about an order of magnitude less than wild-type, but about three orders of magnitude more than the negative control (Table 2). From these data, we conclude that the pREA400 encoded *traA* gene function is required for megaplasmid transfer in *R. erythropolis* AN12.

4. Discussion

Common modes by which bacteria exchange genetic information include phage-mediated transduction and plasmid conjugation. Exploitation and further understanding of the latter method should prove most beneficial for the industrially relevant *Rhodococcus*, since plasmids have the largest capacity for genetic information after the chromosome. The conjugative properties of these megaplasmids are of particular interest from the metabolic engineering point of view, because traditional piecemeal cloning methods required to reconstitute multi-enzymatic pathways could prove extremely difficult. Indeed, one could imagine engineering a single strain of *Rhodococcus* with many desirable activities simply by shuffling together megaplasmids via conjugation.

This study has shown that the genome of *R. erythropolis* AN12 harbors at least three megaplasmids in addition to the previously characterized small cryptic plasmid pAN12. Two megaplasmids, pREA400 and pREA100, appear brighter on pulsed-field gels than pREA250, suggesting that they are more abundant. Quantitative PCR techniques can be used in future experiments to determine the relative copy numbers of these megaplasmids in relationship to the AN12 chromosome.

Conjugative transfer frequencies of pREA400 and pREA250 were determined to be more than 1 event per 10,000 recipient SQ1 cells, which is very comparable to what has already been shown for other rhodococcal megaplasmids. Interestingly, plasmid size is apparently not a factor in determining efficiency of plasmid transfer for rhodococci,

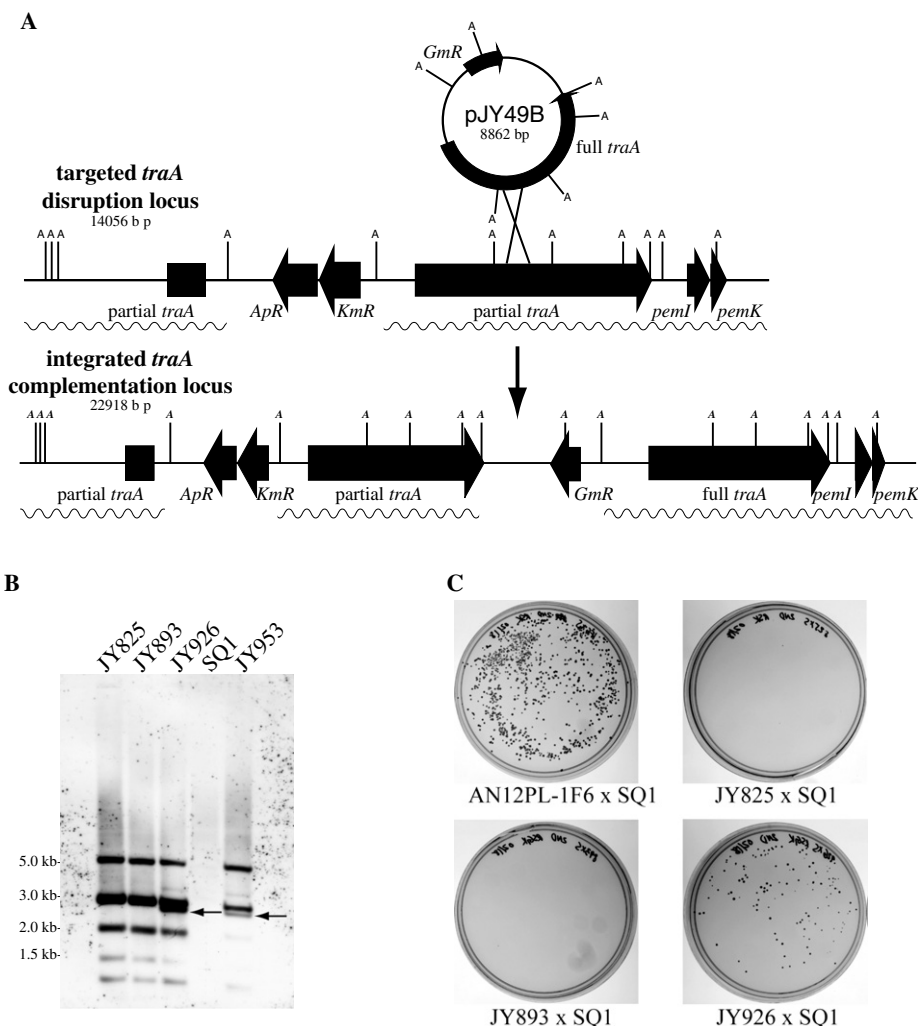


Fig. 6. Genetic and phenotypic complementation of the *traA* disruption mutant. (A) The integration of the complementation plasmid, pJY49B at the disrupted *traA* locus is depicted. pJY49B contains 4900 bp homologous sequence corresponding to the entire *traA* ORF, 221 bp of upstream sequence, as well as 156 bp downstream sequence. The exact site of pJY49B integration is unknown. Sites corresponding to *Afl*III recognition sequences are labeled, as are the regions which anneal to Southern probes generated from pEZTn1F6 template (indicated with wavy lines). (B) Southern blot confirmation. Genomic DNA from each sample was linearized with *Afl*III, then probed with digoxigenin labeled probe synthesized with pEZTn1F6 template DNA. Expected annealing DNA fragments are as follows: JY825 and JY893—3210, 2227, 1342, and 1100 bp; JY926 and JY953(JY926 × SQ1 transconjugant)—3210, 2855, 2227, 1342, and 1100 bp. The arrows point to novel 2,855 bp hybridizing bands in JY926 and JY953 that differentiate these strains from JY925 and JY893 *traA* mutants. No DNA from SQ1 was expected to hybridize to the probe. (C) Phenotypic analysis of the *traA* disruption mutant. Donor strains AN12PL-1F6, JY825, JY893, or JY926 were mated to SQ1 recipients (marked as donor × recipient), and a 1:200 dilution of the each mating cell suspension were plated out onto LB with triple (Rf Sm Km) drug selection. Photographs were taken 5 days after plating.

since pREA400 and pREA250 exhibit essentially the same mobilization frequency.

PFG analysis shows that AN12PL-1F6 was apparently cured of megaplasmid pREA100, and possibly pREA250. One alternative explanation for this observation might have been that the AN12PL-1F6 mutant is actually an SQ1 contaminant, since SQ1 carries a megaplasmid of approximately 400 kb. However, four lines of evidence (data not shown)

allowed us to rule out the possibility. First, sequence analysis and subsequent Southern blotting showed that the rescued plasmid containing the transposon and flanking AN12PL-1F6 genomic DNA, pEZ-Tn1F6, harbored AN12-specific sequences not found in SQ1. Second, pAN12 can be purified from AN12PL-1F6 but is never found in wild-type SQ1 cells. Third, unlike SQ1, AN12PL-1F6 is sensitive to the antibiotics, rifampicin and streptomycin. Lastly,

AN12PL-1F6 growth rate is virtually identical to AN12 WT, whereas SQ1 cells grow more slowly. These data strongly suggest that this strain is an AN12 derivative. We reasoned that the curing of the megaplasms in AN12PL-1F6 was a spontaneous event and not related to electroporation or the presence of the EZ::Tn transposon, since the AN12-5A6 mutant generated in the same way appears to have all of its megaplasms. Furthermore, all other cells tested from the same culture that was used to generate AN12PL-1F6 also lacked pREA100 (and possibly pREA250), indicating that curing of these megaplasms was a common feature of the founding culture.

At this time, without sequence information for pREA100, we cannot be certain whether it has been completely lost from the AN12PL-1F6 genome. It is possible that the copy number of pREA100 in AN12PL-1F6 is below our limits of detection. It is also possible that pREA100 integrated into the genome. This has been shown to be the case for the large plasmids of *Streptomyces*, which exhibit remarkable fluidity and recombination potential (Bey et al., 2000). It would be very interesting to determine whether this plasticity is the case for *rhodococci* as well.

Four methods of targeted gene disruption that rely on double cross-over (ends-out) recombinations have been described for *Rhodococcus* genetic analysis prior to this work (Desomer et al., 1991; Patrauchan et al., 2005; Plaggenborg et al., 2006; van der Geize et al., 2001). One recent study of carotenoid production in *R. erythropolis* ATCC47072 successfully employed a simple single cross-over (ends-in) recombination strategy with a targeting cassette of 800 bp homologous sequence to generate a gene knockout (Tao and Cheng, 2004). We devised an ends-in targeted gene disruption strategy to study *Rhodococcus* gene function, based on similar strategies employed in eukaryotic systems. Our preliminary results indicate that the DNA lesion within the homologous sequence is critical to generate the desired targeted gene disruption/recombination event, as we could not achieve the *traA* mutation with an intact pJY37 plasmid, nor with pJY37 pre-linearized at a restriction site outside of the homology cassette (data not shown). Evidence suggests that larger targeting cassettes facilitate the recombination in *R. erythropolis* AN12, as the integration event of the pJY39 rescue construct was readily generated (>500 Gm^R CFU with 0.1 µg DNA), whereas the *traA* gene disruption event was rare. Currently,

the gene-disruption frequency of our method (1%) is poor compared to the 12% success rate reported by the λ-red-mediated gene disruption method (Patrauchan et al., 2005). In the future, we seek to optimize our procedure by exploring the effects of increasing lengths of homology, the restriction site used, and by improving transformation efficiencies. Although we have used this gene disruption strategy to study the effects of a *traA* loss-of-function mutation, simple modifications to this homologous recombination driven strategy should allow one to generate epitope-tagged fusion proteins *in vivo*, and to study the regulation of gene expression *in vivo* by integrating reporter genes, such as *lacZ* or *GFP*, downstream of promoters of interest. Alternatively, a modification of this strategy could be used to replace endogenous promoters to alter gene expression.

Though a putative relaxase function has been predicted by a recent *Rhodococcus* megaplasmid sequencing effort (Sekine et al., 2006), the present study is the first genetic analysis of a determinant of megaplasmid conjugation in *Rhodococcus*. It is known that the transesterification reaction carried out by all relaxases characterized thus far is dependent on the action and proximity of 3 or 4 key residues at the active site, one of which is the absolutely conserved catalytic tyrosine residue required for the formation of a covalent phosphotyrosyl bond between DNA and protein (Larkin et al., 2005; Pansegrau et al., 1994; Scherzinger et al., 1993; Street et al., 2003). It is expected that the relaxase encoded by *R. erythropolis* AN12 pREA400 *traA* will function similarly because these residues appear to be conserved (Figs. 3B and C). Mutagenesis of each of these residues, and genetic analysis of the ability of the mutant to restore megaplasmid transfer should give us additional insights as to the conservation of the relaxase active site in *Rhodococcus*.

While relaxase/helicase functions of the F factor *traI* and similar genes found on other Gram-negative plasmid have been well-characterized over the past two decades, only three Gram-positive plasmid encoded relaxases have been examined in detail; the *traA* gene product from the 30 kb *Streptococcus* plasmid, pIP501, the *nes* gene product of the 50 kb *Staphylococcus* pGO1, and the *mobM* gene product from 5 kb *Streptococcus* plasmid, pMV158 (Climo et al., 1996; de Antonio et al., 2004; Grohmann et al., 1999; Guzman and Espinosa, 1997; Kopec et al., 2005; Kurenbach et al., 2002). None of these previously characterized Gram-positive relaxases also encode helicase activities. In contrast, our study

indicates in actinomycetes, relaxase activities are often associated with helicase activities. In fact, based on sequence similarities of pREA400 TraA to its most closely related proteins, we propose that pREA400 TraA may be the founding member of a new family of relaxases since its relaxase motif I seems to diverge considerably from consensus motif I sequences of other relaxases. Alternatively, these actinomycete relaxases may be a subfamily of the *IncF/W* relaxases, such as F TraI and R388 TrwC, as they align well-over relaxase motifs IA, II and III. Particularly interesting is the conservation of the two tyrosine residue (Y23 and Y24 of F TraI, and Y35 and Y36 in pREA400 TraA) in the motif IA, with the gap between these and the active site tyrosine (Y16 of F TraI and Y14 of pREA400 TraA) being slightly longer in the actinomycete relaxases. Some evidence suggests that the first of these two tyrosine residues in R388 TrwC (Y26) can participate in DNA cleavage reactions *in vitro* and may be important for conjugation termination (Grandoso et al., 2000). Further experimentation will show whether this holds true for the analogous Y35 residue in pREA400 TraA.

Our study suggests that the basic machinery required to initiate plasmid transfer appears to be evolutionarily conserved from previously well-characterized plasmids, such as F, pTiC58, and RP4 to a newly characterized actinomycete megaplasmid, pREA400. Based on the conservation of the pREA400 *traA* relaxase with the F plasmid relaxase, we speculate that pREA400 may transfer as a single-stranded DNA. Further characterization of the *traA* function, regulation, gene product localization and protein–protein interactions will elucidate conserved and novel aspects of actinomycete megaplasmid conjugation.

Acknowledgments

We thank J. Parker, L. Willis, P. Boccazzi, J. VanEssendelft, J. Nordman, and V. Losick for critical reading of this manuscript. We also thank the insightful comments from two anonymous reviewers and an editor on an earlier draft of this article. This work was funded by the DuPont-MIT Alliance, and J. Yang was partially supported by a graduate fellowship award through the MIT Department of Biology. N. Sengupta and S.D. Windsor were supported by HHMI Undergraduate Biological Sciences Education Program Grant #771195-501306.

References

- Abdel-Monem, M., Taucher-Scholz, G., Klinkert, M.Q., 1983. Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. *Proc. Natl. Acad. Sci. USA* 80, 4659–4663.
- Argandona, M., Martinez-Checa, F., Llamas, I., Quesada, E., del Moral, A., 2003. Megaplasms in Gram-negative, moderately halophilic bacteria. *FEMS. Microbiol. Lett.* 227, 81–86.
- Bey, S.J., Tsou, M.F., Huang, C.H., Yang, C.C., Chen, C.W., 2000. The homologous terminal sequence of the *Streptomyces lividans* chromosome and SLP2 plasmid. *Microbiology* 146 (Pt 4), 911–922.
- Broker, D., Arenskotter, M., Legatzki, A., Nies, D.H., Steinbuechel, A., 2004. Characterization of the 101-kilobase-pair megaplasmid pKB1, isolated from the rubber-degrading bacterium *Gordonia westfalica* Kb1. *J. Bacteriol.* 186, 212–225.
- Byrd, D.R., Sampson, J.K., Ragonese, H.M., Matson, S.W., 2002. Structure–function analysis of *Escherichia coli* DNA helicase I reveals non-overlapping transesterase and helicase domains. *J. Biol. Chem.* 277, 42645–42653.
- Climo, M.W., Sharma, V.K., Archer, G.L., 1996. Identification and characterization of the origin of conjugative transfer (*oriT*) and a gene (*nes*) encoding a single-stranded endonuclease on the staphylococcal plasmid pGO1. *J. Bacteriol.* 178, 4975–4983.
- Cox, M.M., 1999. Recombinational DNA repair in bacteria and the RecA protein. *Prog. Nucleic Acid Res. Mol. Biol.* 63, 311–366.
- Dabrock, B., Kessler, M., Averhoff, B., Gottschalk, G., 1994. Identification and characterization of a transmissible linear plasmid from *Rhodococcus erythropolis* BD2 that encodes isopropylbenzene and trichloroethene catabolism. *Appl. Environ. Microbiol.* 60, 853–860.
- de Antonio, C., Farias, M.E., de Lacoba, M.G., Espinosa, M., 2004. Features of the plasmid pMV158-encoded MobM, a protein involved in its mobilization. *J. Mol. Biol.* 335, 733–743.
- de Carvalho, C.C., da Fonseca, M.M., 2005. The remarkable *Rhodococcus erythropolis*. *Appl. Microbiol. Biotechnol.* 67, 715–726.
- Desomer, J., Dhaese, P., Van Montagu, M., 1988. Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. *J. Bacteriol.* 170, 2401–2405.
- Desomer, J., Crespi, M., Van Montagu, M., 1991. Illegitimate integration of non-replicative vectors in the genome of *Rhodococcus fascians* upon electrotransformation as an insertional mutagenesis system. *Mol. Microbiol.* 5, 2115–2124.
- Engelberg-Kulka, H., Glaser, G., 1999. Addition modules and programmed cell death and antideath in bacterial cultures. *Annu. Rev. Microbiol.* 53, 43–70.
- Farrand, S.K., Hwang, I., Cook, D.M., 1996. The *tra* region of the nopaline-type Ti plasmid is a chimera with elements related to the transfer systems of RSF1010, RP4, and F. *J. Bacteriol.* 178, 4233–4247.
- Grandoso, G., Avila, P., Cayon, A., Hernando, M.A., Llosa, M., de la Cruz, F., 2000. Two active-site tyrosyl residues of protein TrwC act sequentially at the origin of transfer during plasmid R388 conjugation. *J. Mol. Biol.* 295, 1163–1172.
- Grohmann, E., Guzman, L.M., Espinosa, M., 1999. Mobilisation of the streptococcal plasmid pMV158: interactions of MobM protein with its cognate *oriT* DNA region. *Mol. Gen. Genet.* 261, 707–715.

- Guzman, L.M., Espinosa, M., 1997. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid oriT. *J. Mol. Biol.* 266, 688–702.
- Hall, M.C., Matson, S.W., 1999. Helicase motifs: the engine that powers DNA unwinding. *Mol. Microbiol.* 34, 867–877.
- Harley, M.J., Toptygin, D., Troxler, T., Schildbach, J.F., 2002. R150A mutant of F TraI relaxase domain: reduced affinity and specificity for single-stranded DNA and altered fluorescence anisotropy of a bound labeled oligonucleotide. *Biochemistry* 41, 6460–6468.
- Kalkus, J., Dorrie, C., Fischer, D., Reh, M., Schlegel, H.G., 1993. The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini. *J. Gen. Microbiol.* 139, 2055–2065.
- Kinashi, H., Shimaji, M., Sakai, A., 1987. Giant linear plasmids in *Streptomyces* which code for antibiotic biosynthesis genes. *Nature* 328, 454–456.
- Kopeck, J., Bergmann, A., Fritz, G., Grohmann, E., Keller, W., 2005. TraA and its N-terminal relaxase domain of the Gram-positive plasmid pIP501 show specific *oriT* binding and behave as dimers in solution. *Biochem. J.* 387, 401–409.
- Kostichka, K., Tao, L., Bramucci, M., Tomb, J.F., Nagarajan, V., Cheng, Q., 2003. A small cryptic plasmid from *Rhodococcus erythropolis*: characterization and utility for gene expression. *Appl. Microbiol. Biotechnol.* 62, 61–68.
- Kurenbach, B., Grothe, D., Farias, M.E., Szwed, U., Grohmann, E., 2002. The *tra* region of the conjugative plasmid pIP501 is organized in an operon with the first gene encoding the relaxase. *J. Bacteriol.* 184, 1801–1805.
- Lanka, E., Wilkins, B.M., 1995. DNA processing reactions in bacterial conjugation. *Annu. Rev. Biochem.* 64, 141–169.
- Larkin, C., Datta, S., Harley, M.J., Anderson, B.J., Ebie, A., Hargreaves, V., Schildbach, J.F., 2005. Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. *Structure* 13, 1533–1544.
- Lessard, P.A., O'Brien, X.M., Currie, D.H., Sinskey, A.J., 2004. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. *BMC Microbiol.* 4, 15.
- Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–W331.
- Matson, S.W., Ragonese, H., 2005. The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. *J. Bacteriol.* 187, 697–706.
- Orr-Weaver, T.L., Szostak, J.W., Rothstein, R.J., 1981. Yeast transformation: a model system for the study of recombination. *Proc. Natl. Acad. Sci. USA* 78, 6354–6358.
- Pansegrau, W., Schroder, W., Lanka, E., 1994. Concerted action of three distinct domains in the DNA cleaving-joining reaction catalyzed by relaxase (TraI) of conjugative plasmid RP4. *J. Biol. Chem.* 269, 2782–2789.
- Patrauchan, M.A., Florizone, C., Dosanjh, M., Mohn, W.W., Davies, J., Eltis, L.D., 2005. Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. *J. Bacteriol.* 187, 4050–4063.
- Pedraza, R.O., Diaz Ricci, J.C., 2002. In-well cell lysis technique reveals two new megaplasmids of 103.0 and 212.6 MDa in the multiple plasmid-containing strain V517 of *Escherichia coli*. *Lett. Appl. Microbiol.* 34, 130–133.
- Plaggenborg, R., Overhage, J., Loos, A., Archer, J.A., Lessard, P., Sinskey, A.J., Steinbuchel, A., Priefert, H., 2006. Potential of *Rhodococcus* strains for biotechnological vanillin production from ferulic acid and eugenol. *Appl. Microbiol. Biotechnol.* 1–11.
- Priefert, H., O'Brien, X.M., Lessard, P.A., Dexter, A.F., Choi, E.E., Tomic, S., Nagpal, G., Cho, J.J., Agosto, M., Yang, L., Treadway, S.L., Tamashiro, L., Wallace, M., Sinskey, A.J., 2004. Indene bioconversion by a toluene inducible dioxygenase of *Rhodococcus* sp. I 24. *Appl. Microbiol. Biotechnol.* 65, 168–176.
- Sajjaphan, K., Shapir, N., Wackett, L.P., Palmer, M., Blackmon, B., Tomkins, J., Sadowsky, M.J., 2004. *Arthrobacter aureus* TC1 atrazine catabolism genes *trzN*, *atzB*, and *atzC* are linked on a 160-kilobase region and are functional in *Escherichia coli*. *Appl. Environ. Microbiol.* 70, 4402–4407.
- Sambrook, J., Russell, D.W., 2001. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
- Scherzinger, E., Kruff, V., Otto, S., 1993. Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleaving/DNA-joining activity. *Eur. J. Biochem.* 217, 929–938.
- Scholle, M.D., White, C.A., Kunnimalaiyaan, M., Vary, P.S., 2003. Sequencing and characterization of pBM400 from *Bacillus megaterium* QM B1551. *Appl. Environ. Microbiol.* 69, 6888–6898.
- Sekine, M., Tanikawa, S., Omata, S., Saito, M., Fujisawa, T., Tsukatani, N., Tajima, T., Sekigawa, T., Kosugi, H., Matsuo, Y., Nishiko, R., Imamura, K., Ito, M., Narita, H., Tago, S., Fujita, N., Harayama, S., 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. *Environ. Microbiol.* 8, 334–346.
- Shimizu, S., Kobayashi, H., Masai, E., Fukuda, M., 2001. Characterization of the 450-kb linear plasmid in a polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. *Appl. Environ. Microbiol.* 67, 2021–2028.
- Simon, R., Priefer, U., Puhler, A., 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagenesis in gram negative bacteria. *Biotechnology* 1, 784–791.
- Stecker, C., Johann, A., Herzberg, C., Averhoff, B., Gottschalk, G., 2003. Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. *J. Bacteriol.* 185, 5269–5274.
- Street, L.M., Harley, M.J., Stern, J.C., Larkin, C., Williams, S.L., Miller, D.L., Dohm, J.A., Rodgers, M.E., Schildbach, J.F., 2003. Subdomain organization and catalytic residues of the F factor TraI relaxase domain. *Biochim. Biophys. Acta* 1646, 86–99.
- Takai, S., Hines, S.A., Sekizaki, T., Nicholson, V.M., Alperin, D.A., Osaki, M., Takamatsu, D., Nakamura, M., Suzuki, K., Ogino, N., Kakuda, T., Dan, H., Prescott, J.F., 2000. DNA sequence and comparison of virulence plasmids from *Rhodococcus equi* ATCC 33701 and 103. *Infect. Immun.* 68, 6840–6847.
- Tao, L., Cheng, Q., 2004. Novel beta-carotene ketolases from non-photosynthetic bacteria for canthaxanthin synthesis. *Mol. Genet. Genomics* 272, 530–537.
- Tauch, A., Bischoff, N., Brune, I., Kalinowski, J., 2003. Insights into the genetic organization of the *Corynebacterium diphtheriae* erythromycin resistance plasmid pNG2 deduced from its complete nucleotide sequence. *Plasmid* 49, 63–74.
- Thompson, J.D., Higgins, D.G., Gibson, T.J., 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.

- Traxler, B.A., Minkley Jr., E.G., 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. *J. Mol. Biol.* 204, 205–209.
- Trott, S., Burger, S., Calaminus, C., Stolz, A., 2002. Cloning and heterologous expression of an enantioselective amidase from *Rhodococcus erythropolis* strain MP50. *Appl. Environ. Microbiol.* 68, 3279–3286.
- van den Bosch, M., Lohman, P.H., Pastink, A., 2002. DNA double-strand break repair by homologous recombination. *Biol. Chem.* 383, 873–892.
- van der Geize, R., Hessels, G.I., van Gerwen, R., van der Meijden, P., Dijkhuizen, L., 2001. Unmarked gene deletion mutagenesis of *kstD*, encoding 3-ketosteroid Delta1-dehydrogenase, in *Rhodococcus erythropolis* SQ1 using *sacB* as counter-selectable marker. *FEMS Microbiol. Lett.* 205, 197–202.
- van der Geize, R., Dijkhuizen, L., 2004. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. *Curr. Opin. Microbiol.* 7, 255–261.
- Warren, R., Hsiao, W.W., Kudo, H., Myhre, M., Dosanjh, M., Petrescu, A., Kobayashi, H., Shimizu, S., Miyauchi, K., Masai, E., Yang, G., Stott, J.M., Schein, J.E., Shin, H., Khattra, J., Smailus, D., Butterfield, Y.S., Siddiqui, A., Holt, R., Marra, M.A., Jones, S.J., Mohn, W.W., Brinkman, F.S., Fukuda, M., Davies, J., Eltis, L.D., 2004. Functional characterization of a catabolic plasmid from polychlorinated-biphenyl-degrading *Rhodococcus* sp. strain RHA1. *J. Bacteriol.* 186, 7783–7795.
- Wu, W., Leblanc, S.K., Piktet, J., Jensen, S.E., Roy, K.L., 2006. Prediction and functional analysis of the replication origin of the linear plasmid pSCL2 in *Streptomyces clavuligerus*. *Can. J. Microbiol.* 52, 293–300.
- Zechner, E.L., de la Cruz, F., Eisenbrandt, R., Grahn, A.M., Koraimann, G., Lanka, E., Muth, G., Pansegrau, W., Thomas, C.M., Wilkins, B.M., Zatyka, M., 2000. Conjugative-DNA transfer processes. In: Thomas, C.M. (Ed.), *The Horizontal Gene Pool: Bacterial plasmids and Gene Spread*. Harwood Academic, Amsterdam. pp. xxiii, 419 p.