Conjugation of Extrachromosomal Replicons of *Rhodococcus erythropolis* AN12

by

Joyce Chun-Yi Yang

B.A., Biochemistry
Rutgers University, 1998

SUBMITTED TO THE DEPARTMENT OF BIOLOGY IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

AT THE

MASSACHUETTS INSTITUTE OF TECHNOLOGY

MAY 2006
[June 2006]

Signature of Author: _____
Department of Biology
May 15, 2006

Certified by: _____
Anthony J. Sinskey
Professor of Microbiology
Thesis Supervisor

Accepted by: _____
Stephen Bell
Professor of Biology
Chairman, Committee for Graduate Students

Conjugation of Extrachromosomal Replicons of *Rhodococcus erythropolis* AN12

by

Joyce C. Yang

Submitted to the Department of Biology
On May 15, 2006 in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy in Biology

Abstract.

Bacteria belonging to the Gram-positive actinomycete species, *Rhodococcus erythropolis*, are diverse not only in terms of metabolic potentials but the plasmids they encode. Pulsed-field gel electrophoresis (PFGE) revealed three previously uncharacterized megaplasmids in the genome of *Rhodococcus erythropolis* AN12. These megaplasmids, pREA400, pREA250 and pREA100, migrate at approximately 400 kb, 250 kb and 100 kb, respectively. Genetic screening of an AN12 transposon insertion library showed that two megaplasmids, pREA400 and pREA250, are conjugative. 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 novel site-specific gene disruption method was developed to generate a *traA* mutation in AN12, which allowed us to address the role of the *traA* gene for *Rhodococcus* megaplasmid conjugation. We found that the AN12 *traA* mutant is no longer capable of transferring the pREA400 megaplasmid to *Rhodococcus erythropolis* SQ1.

It was shown previously that the *R. erythropolis* AN12 genome harbors a 6.3 kb cryptic plasmid called pAN12. Here we show that pAN12 is conjugatively mobilizable into other rhodococcal strains. A series of plasmid deletion constructs were tested for loss of mobility to identify the pAN12 *cis*-acting conjugation requirement. In this way, an approximately 700 bp region was found to be required for plasmid transmission. A small 61 bp element within this region exhibited sequence similarity to the minimal 54 bp *clt* region known to be required for the conjugation of the streptomycete plasmid, pIJ101. The functionality of these *cis*-acting elements appears to be conserved, as the addition of this pAN12 *clt*-like region confers mobility to an otherwise non-conjugative plasmid. However, unlike pJI101 which encodes all necessary factors for transfer, pAN12 mobility is dependent on the presence of the AN12 megaplasmid, pREA400.

For my father
(1939-1998)

from whom I've inherited
a strict work ethic and
a sometimes churlish stubborness,
which in its more noble form
as perserverance
has served as my greatest asset on this journey

and

For my mother,

the bravest woman I know
who taught me
(in additional to great phone manners)
that integrity is the key to successful friendships and careers.

ACKNOWLEDGEMENTS

First and foremost, this work would not have been possible without my thesis advisor, Dr. Sinskey. Certainly, had Tony not taken a chance on me about two and a half years ago, I would have written a much different thesis. Not only has Tony built a rich working environment, but his encouragements and candid critiques both gave me faith in my own abilities and taught me how to be even better. The Sinskey lab's quest to understand basic biological phenomenons in the context of industrial metabolic engineering has rekindled in me the desire to focus on application-based research. I've also benefited greatly from Tony's unique perspective in his dual roles as professor and entrepreneur. Given the long history of MIT 1K/50K Entrepreneurship Competition winners from our lab, one can only conclude that others have felt that way too. Special thanks go to my MIT thesis committee members, Dr. RajBhandary, Dr. Walker for their time and guidance, and most especially to Dr. Orr-Weaver for her unwavering support on my journey through three different laboratories. I am also indebted to Drs. Garrity and Solomon for their mentorship. Across the ocean at the other Cambridge, I am very grateful to my outside thesis examiner, Dr. Archer, who not only gave me insightful comments but went out of his way to attend my defense, well above and beyond the duties of being our collaborator.

I would not have pursued a scientific research path had it not been for the guidance of a Whippany Park High School physics and chemistry teacher, Mr. Altenderfer, and the endless patience of my undergraduate advisor, Dr. Huibregtse, while attending Rutgers University. Thanks also to my collaborators at DuPont de Nemours, especially Drs. Nagarajan, Tomb, and Bramucci, who provided reagents, sequence and advice on working with rhodococci. I'm grateful for the financial support from the NIH Pre-Doctoral Training Grant, the MIT Dept. Biology, the Cancer Center Anna Fuller Fellowship, and the DuPont-MIT Alliance, throughout my graduate career.

I would like to thank the members of the Sinskey lab- especially Fen Tung, Dr. Vu Bui, Dr. Paolo Boccazzi, Dr. Jefferson Parker, Jes and Wil VanEssendelft, Joerg Schoenheit, Jae-Yeun Song, and Tim Ohlrich- for friendship, discussions, insights, and support. I've learned so much from all of you. Special thanks go to Xian O'Brien for optimizing pulsed-field gels, and undergraduates, Devin Currie, Neil Sengupta, and Steven Windsor for help in generating the transposon library. I am especially indebted to Dr. Philip Lessard for not only spearheading this and other *Rhodococcus* projects, but his mentorship, indomitable spirit, encyclopedic knowledge, and tireless reading of thesis chapters and manuscripts. Phil is also an excellent role model of how to juggle a family and a science career successfully. I don't know how he fits 48 hours into a day, but I'm not ruling out magic. Thanks also to lab/teammates- Phil , Charles Budde and Nina Kshetry- for the unforgettable MIT 1K Microbefuel/Nashville experience. At the Cancer Center, I thank Dr. Margaret Magendantz for teaching me how to do impeccable experiments, and Michael Moran, who rescued so many of my Macintosh computers from the brink of death. At the Garrity lab, I'd like to thank Monique Brouillette for help with the UAS-Ptpmeg constructs, and Dr. Jessica Whited, for being a vociferous ally, dear friend and terrific colleague.

I have been blessed by friendships with wonderful individuals, both at and outside of MIT. Some were members of the Class of 98 (Philina Lee, Arv Govindarajan, Suzanne Nguyen, Sandra Luikenhuis, Aron Eklund, John Doll, Jen Saionz, Brianna Burton, Tim Tayler, and Soni Lacefield). Others were found through WMBR (Rachel Grubb, Ted Young, Gene Fierro, Tani Chen, and Keith Sawyer). Yet more from the Rainbow Coffeehouse (Kevin Choi, Ajit Dash, Carlos Pacheco, Annie Frazer, An Nguyen, and Albert Chan), and through the Arts Scholars Program (Michele Oshima and Christa Starr). Also Elissa Lei, Stephanie Richard, Andrea Vala, Jared Nordman, Tim Montminy, and Mike Fisher at Harvard and Tufts for dating advice, friendships, and support. Thank you all for helping me cope with the challenging times, and sharing laughs through the good times.

Lastly, I'd like to acknowledge my family and extended family including; 1) my parents and my sister, Amy, whose unfailing love, faith and values are my ammunitions against any challenge, 2) my parents-in-law, Jan and Rich, for their unwavering support, generosity, and wisdom on life and microbiology, 3) Ga and Steve, for stimulating conversations (scientific, literary, and otherwise) and for cultivating an unlikely appreciation for wine despite of my Asian Flush Syndrome, and finally, 4) my best friend, colleague, and wife Vicki, who has taught me everything else I needed to know.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

Introduction.

1.1. Background and motivation.

1.1.1. Introduction to the biology and versatility of *Rhodococcus*. Bacteria of the genus *Rhodococcus* are Gram-positive, GC-rich, aerobic, non-motile members of the order *Actinomycetales*. As such, they are related to clinically and industrially important actinomycetes, such as *Mycobacterium tuberculosis*, the causative agent of tuberculosis, and various strains of *Streptomyces*, well known as producers of antibiotics. Actinomyces literally means "ray-fungus" in Greek, as their sometime filamentous morphologies closely resemble small fungi. Indeed, the classification of rhodococci as a distinct genera historically has been difficult due to their pleomorphic nature, such that within a single isogenic strain, cells may appear as cocci, short rods, filaments, or take on elaborately branched morphotypes (Figure 1-1A). In addition, colony morphologies between different strains can vary; colonies may be smooth or rough edged, mucoid or dry, exhibiting colors from white to deep orange-red (Figure 1-1B).

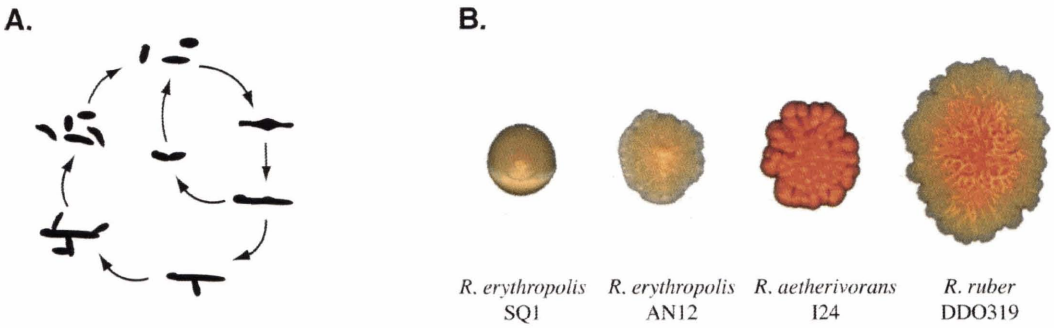


Figure 1.1. Cell and colony morphologies of *Rhodococcus* bacteria. (A) Morphologies that rhodococci cells can adopt during cycles of division and growth, adapted from an insert on pg. 91 in the *Biology of Actinomycetes* (49). (B) Variance of colony morphology between indicated strains of *Rhodococcus*. Photos taken (SQ1 through DDO319) after 1 month, 13 days, 1 month, and 8 days cultivation on LB agar at 30°C, respectively.

Rhodococci belong to the nocardioform family of bacteria, which are differentiated from other actinomycetes based on cell wall composition. Specifically, these bacteria possess chemotype IV cell walls, consisting of tiered meshwork of peptidoglycans, arabinogalactans, and lipoarabinomannans, atop basal lipid bilayers (92, 93, 125). Also found in these cell walls are mycolic acids, or large branched fatty acids between 20 and 90 carbons long, existing in both bound and free forms (151). Mycolic

acids are produced by other bacterial genera, including *Corynebacteria,* and *Mycobacteria*; collectively, this subgroup of bacteria is referred to as the mycolata. It is thought that the perpendicular arrangement of mycolic acids with respect to the cell's lipid bilayer results in a chemically recalcitrant permeability barrier (112, 113). Consistent with this theory, rhodococci as a group are especially noted for their solvent tolerance, being frequently isolated from contaminated soil and aquatic environments.

Though most *Rhodococcus* species are benign microbes, *Rhodococcus equi* has been found to be a pathogen of foals (102) and immune-compromised human (140). A few rhodococci have been found to be symbionts of termites (82, 101) and the insect, *Rhodnius prolixus*, associated with Chagas' disease (7, 41). Strains of *Rhodococcus fascians* were discovered to cause tumor-like growths (leafy galls) near the base of the sweet pea plants (10, 159). This pathogenistic process, known as fasciation, is due to the presence of three genetic loci- *fas, att,* and *hyp-* on a 200 kb linear megaplasmid, pFiD188 (18). Over 30 species of rhodococci have been taxonomically classified to date since the discovery of the type species, *Rhodococcus rhodochrous*, (aka *Rhodococcus roseus*) in 1891 (172).

1.1.2. Importance of *Rhodococcus* in applied microbiology. The veritable treasure trove of enzymatic pathways rhodococci encode, along with natural recalcitrance to organic solvents, makes this group of bacteria a particularly desirable resource for biotransformations, and leading candidate in considerations for industrial bioprocesses. Specifically, enzymes from rhodococci are capable of chiral/enantiomeric bioconversions of such exquisite specificity- such as the conversion of indene to *cis*-1,2-indiandiol (12, 162) and the resolution of racemic methyl nonactate to pure enantiomers (123)- that they could eventually augment or replace difficult chemical synthesis methods in the production of pharmaceuticals. *Rhodococcus* species are also masters of biodegradation, acting upon substrates ranging from fuels (26, 96) to halogenated organic compounds, in the form of chlorinated alkanes (20, 85, 141) or polycyclic aromatics (polychlorinated biphenyls or PCB) (5, 91, 107, 121, 146). As such, *Rhodococcus* bacteria have emerged as important microbes for bioremediation of oil spills and xenobiotic pollutants.

It is impossible to fully review the myriad of metabolic pathways documented for rhodococci; literally hundreds of scholarly articles have been published on this subject.

Instead, the discussion in this section will focus on the dominant roles that *Rhodococcus* plays in the biodesulfurization of fossil fuels, and the industrial production of acrylamide, as emerging and established bioprocesses, respectively. A more detailed list of plasmid-encoded enzymatic reactions in rhodococci and the possibility of using bacterial conjugation as the basis of metabolic engineering- serving as the impetus for this thesis research- will also be covered in a later section of this chapter. Readers are referred to excellent reviews for detailed discussions of other metabolic pathways these microbes possess (4, 27, 44, 90, 169).

1.1.2.1. *Rhodococcus* and biodesulfurization. This past decade bore witness to a major surge in transportation fuel prices, from an average retail price of $1.10/gallon in May of 1996 to the current price of $2.60/gallon (www.eia.doe.gov), an increase well above the concurrent 27% inflation rate (www.bis.gov). Though this is largely due to the rising demand for petroleum, increased refinery costs to desulfurize fuel to the 30 ppm requirement mandated by the Clean Air Act of 1990, has also contributed to this price gain at the pump. Crude oil typically contains between 0.01% to 5% sulfur (or 1000 to 30,000 ppm by weight), which largely exists in the form of an organic compound called dibenzothiophene, or DBT (53, 72, 114). The release of sulfur from fuel combustion, and subsequent formation of sulphur oxides in the atmosphere results in smog and acid rain, the latter of which not only causes damage to fisheries, forests, and buildings/structures, but also negatively impacts human health (www.epa.gov). Removal of DBT and other sulfur compounds from crude oil is currently achieved by an energy intensive and costly process called hydrodesulfurization, or HDS (130).

According to a 2004 U.S. Department of Energy (DOE) fact sheet, biodesulfurization (BDS) is a promising alternative to HDS which reduces capital and operating costs, as well as greenhouse gas emissions. BDS utilizes microbes to metabolize the sulfur compounds in reactions that leave the hydrocarbons intact. Several strains, *Rhodococcus erythropolis* IGTS8 (31, 32, 48, 51, 99, 134), *R. erythropolis* KA2-5-1 (58, 78, 117, 120, 156), *R. erythropolis* D-1 (64, 109, 128), *Rhodococcus* sp. ECRD-1 (53, 54, 136), and *Rhodococcus* sp. KT462 (155) have been identified with the desirable properties for BDS.

Most of these bacteria possess enzymes in the so-called 4S pathway (Figure 1.2A), in which a series of monoxygenases and a desulfinase act to transform DBT to 2-hydroxybiphenyl (HBP) and a sulfite ion. The genes encoding these activities are arranged in an operon. They were initially discovered by mutagenizing *Rhodococcus* sp. IGTS8 with shortwave UV, then screening for mutants that no longer secreted HBP, a fluorescent compound whose presence can easily be detected visually following excitation with UV (32). A mutant from this screen, UV1, was then transformed with a cosmid genomic library, and clones containing three open reading frames (*soxABC*) spanning a 10 kb genomic region (Figure 1.2B) restored the ability to produce HBP (32). Interestingly, subsequent analysis of the UV1 mutant by pulsed-field gel electrophoresis (PFGE) demonstrated the nature of the *sox* mutation was due to the curing of a 120 kb megaplasmid named pSOX (31). That the *soxABC* locus is plasmid-borne reflects a reccurring theme in a number of *Rhodococcus* metabolic pathways to be described later in this chapter.

A paper which appeared a few months following Denome et al. (129) by an independent group described the cloning and characterization of the same operon in *Rhodococcus* IGTS8, in which the *sox* genes were renamed *dszABC* (134). The second nomenclature has since been adopted to avoid confusion with previously identified mammalian SOX genes. At about the same time, it was determined that molecular oxygen, the NADH reduced pyridine nucleotide, and the FMN flavin, along with the actions of a flavin oxidoreductase (DszD) are required in the reactions catalyzed by the DszC and DszA monoxygenases (51, 127, 129). Furthermore, an *in vitro* desulfurization system using purified enzymes allowed for detailed kinetic studies of this process, which showed that the reaction catalyzed by the desulfinase in the last step of this pathway is rate-limiting (51). Later studies which elucidated the gene regulatory region upstream of the *dsz* operon (99), and analyzed the desulfurization potential of modified strains that overexpressed *dsz* genes aimed at optimizing this reaction for commercial purposes (46).

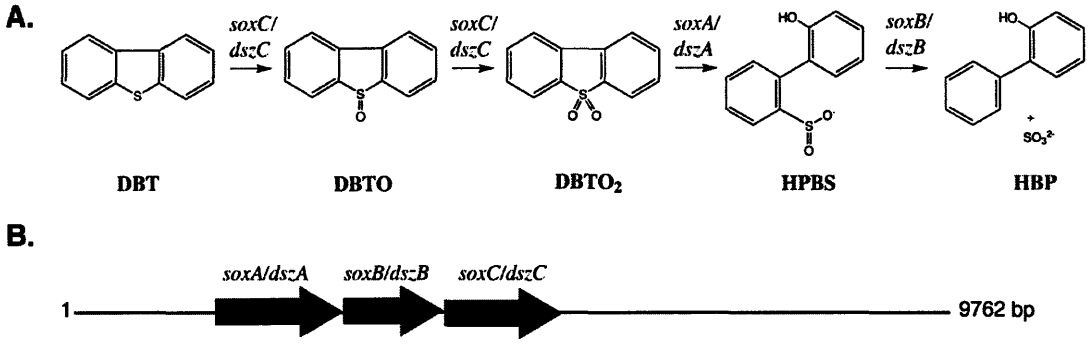


Figure 1.2. The proposed *Rhodococcus* enzymatic pathway for biodesulfurization of fossil fuels. (A) A simplified schematic of the pathway for bioconversion of dibenzothiophene (DBT) to 2-hydroxybiphenyl (HBP) adapted from Folsom et al. (1999). DBT is first converted to intermediates, dibenzothiophene sulfoxide (DBTO) and dibenzothiophene sulfone ($DBTO_2$) by the *soxC/dszC* encoded monoxygenase. $DBTO_2$ is then converted by the *soxA/dszA* monoxygenase to hydroxyphenyl sulfonate (HPBS), a reaction which opens the thiophenic ring. HPBS is then acted upon by the *soxB/dszB* gene product, a desulfinase, to release 2-hydroxybiphenyl back into the oil, and sulfite ion for cellular metabolism. (B) Operon structure of the central *Rhodococcus* desulfurization enzymes. Drawing represents the first cloned and sequenced desulfurization (*sox*) gene locus (Genbank No. U08850) found on a megaplasmid of *Rhodococcus erythropolis* IGTS8. *soxA, soxB, and soxC* are 1,362 bp, 1,098 bp, and 1,254 bp, respectively. These genes were subsequently renamed the *dsz* cluster.

The idea to use microbes for desulfurization reactions is not new. A study published in 1961 examined the use of *Thiobacillus ferrooxidans* to oxidize iron pyrite, FeS_2, from coal (148). Despite this early discovery, it took roughly 40 years for the first microbially catalyzed biodesulfurization patent to be filed (24). Certain strains of *Rhodococcus* appear well-poised to be the most ideal biocatalysts in fossil fuel biodesulfurization (114). It is anticipated that the significant challenges related to poor reaction kinetics and regeneration of the biocatalysts, also pointed out in this review, will be overcome in the near future to make *Rhodococcus*-mediated BDS platform a reality.

1.1.2.2. Use of *Rhodococcus rhodochrous* J1 in acrylamide production. Chemical production of the acrylamide and its versatile polymer, polyacrylamide, began in the 1950's, and now exceeds 400,000 metric tons per year (62). Because polyacrylamides can be made cationic, anionic, or neutral in charge, they are central to many industrial applications, the largest uses of which include the potable/sewage water treatments and the manufacturing of paper. Microbial synthesis of acrylamide was

prompted by the discovery of two different microbial pathways (Figure 1.3) to degrade nitriles (6, 13, 62, 76). The first pathway is mediated by a nitrilase which directly breaks down nitriles (R-CN) to carboxylic acids (R-COOH) and ammonia (NH_3). The second pathway is mediated stepwise by a nitrile hydratase (NHase), which converts R-CN to amides ($R\text{-}CONH_2$), followed by the action of an amidase, which converts amides to R-COOH and NH_3. The latter pathway has been investigated more thoroughly since nitrilases were found to be more labile.

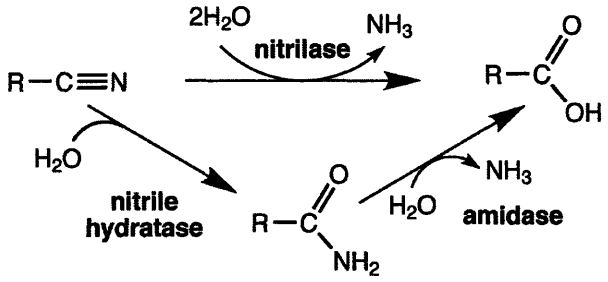


Figure 1.3. The two microbial nitrile degradation pathways schematic, adapted from Warhurst and Fewson (1994).

The *Rhodococcus* (previously known as *Brevibacterium*) strain R312 was initially considered for the microbial synthesis of acrylamide (17) since it yields 2.3M acrylamide every 30 minutes at room temperature. However, this enzymatic conversion was severely inhibited once the acrylamide concentration exceeded 2.8M. The nitrile hydratase of *Rhodococcus rhodochrous* J1 was found to be superior since it did not show significant inhibition by product concentration; indeed, it could produce up to 9.2M of acrylamide within 72 hours, and was thermostable up to 50°C (76). Furthermore, the amidase activity was found to be very low, thus contaminating quantities of acrylic acid was not an issue for this bioprocess.

The *R. rhodochrous* J1 nitrile hydratase activity can be attributed to two cobalt containing metalloenzymes; one of a lower molecular mass (L-NHase) (77), and one of a higher molecular mass (H-NHase) (115). Both enzymes are hetero-oligomers composed of two subunits named α and β (77), encoded by genes *nhhA*, *nhhB*, *nhlA* and *nhlB*. The crystal structure of a related *Rhodococcus* sp. R312 nitrile hydratase was later determined to 2.65 Å resolution (61). This study showed that the α and β subunits of the R312 NHase form a tight dimer. The catalysis occurs within a cavity, mediated by an

activating metal ion bound by thiol groups of three cysteines found in the α subunit. In the case of R312 the metal ion is iron, but evidence suggests that cobalt is the equivalent metal in *R. rhodochrous* J1 (9). Huang et al. proposed that the metal ion acts as a Lewis acid, activating the carbon in nitrile to form a hydroxyl bond.

Because many Fe- and Co-type nitrile hydratase activities have been detected in diverse bacterial genera, a group of investigators recently completed a study examining the nitrile hydratase gene clusters in two different strains of *Rhodococcus erythropolis* (AJ270 and AJ300) and a strain of *Microbacterium* (AJ115) to address the possibility of horizontal gene transfer of these loci in the environment (124). They found that the nitrile hydratase gene clusters are indeed identical in these three bacteria. While linear and circular plasmid were detected, unlike the megaplasmid-encoded *dsz* genes discussed previously, the NHase genes in AJ270, AJ300 and AJ115 are chromosomally located. O'Mahony et al. did determine that a complete copy of transposable element, IS*1166*, lies upstream of the NHase cluster, and concluded these genes most likely spread via transposition. Ironically, the IS*1166* element was identified originally within the *soxABC* operon of desulfurizing bacteria, *Rhodococcus erythropolis* IGTS8 (33).

1.1.3. Genetic analysis of *Rhodococcus* and tools development. Bacterial genetics begins with mutagenesis and recombination. For rhodococci, this was initiated by Adams et al. in the 1960's when UV-induced auxotrophic mutants- deficient for the production of various amino acids- of two different strains of *Rhodococcus erythropolis* (originally called *Nocardia erythropolis* and *Nocardia canicruria*) were mated to each other and prototrophic progenies were observed (1, 2). Importantly, this series of work demonstrated both that genetic material can be transferred between *Rhodococcus* strains, and that irradiation can be used for mutagenesis. However, it was puzzling that unlike *Escherichia coli* in which homologous recombination occurs readily between cells of the same strain, recombination in rhodococci occurs readily only between different strains (1). It was suggested later that this phenomenon of self-incompatibility may be due to conjugation determinants located on plasmids (50, 89). Complications in the genetic analysis of rhodococci have been compounded by the facts that: 1) no isogenic strain has been recognized as a type or reference strain; and 2) to date, there is no publicly available database with the complete and annotated genome sequence of any *Rhodococcus*, though

encouragingly, several groups are nearing this goal (J.-F. Tomb, J. Davies, and J. Archer, personal communication). Nevertheless, early genetic linkage maps of 65 traits were assembled for strains of *R. erythropolis* (11).

Besides UV mutagenesis, which was used to generate the auxotrophs above, and the *dsz* mutants mentioned previously, chemical mutagenesis has been successfully employed to study rhodococci (55, 134, 168). However, by far the most popular way to generate *Rhodococcus* mutants for genetic analysis has been transposition. One of the first reports of transposon-based mutagenesis used an endogenous insertion sequence element, IS*1415*, found proximal to the cobalamin synthesis genetic loci in *R. erythropolis* N186/21 (116). Insertion sequence (IS) elements were initially identified in *Rhodococcus fascians* using an entrapment system (65), in which the *Bacillus subtilis sac*B gene encoding sucrose sensitivity can be inactivated via an IS transposition event. Subsequently, they have been found in other *Rhodococcus* species either by accident through the study of nearby genes, as in the case of IS*1166* (33), or by intention via the *sacB* method (97). The IS*1415* or the resulting Tn*5561*, was found to have little, if any, sequence bias; however, *Tn5561* transposition appeared to be infrequent (1 event per μg vector DNA) and was not pursued further for mutagenesis. Since then, other transposon-based systems, commercial and otherwise, have been used successfully to generate mutants in a variety of *Rhodococcus* strains (3, 43, 103, 139, 156).

Study and utilization of bacterial plasmids, through the advent of recombinant DNA technology (14), have enabled many modern molecular biology techniques. Almost all strains of rhodococci harbor at least one plasmid (89), and study of their replication regions led to the development of shuttle vectors, generally chimeric plasmids that contain both an *E. coli* and a *Rhodococcus* replicon. The first of such, pMVS301, was generated by Singer and Finnerty in 1988 (149), using a 3.8 kb region of the endogenous pMVS300 plasmid found in *Rhodococcus* sp. strain H13-A. It carries a thiostrepton resistance cassette cloned from *Streptomycetes* (157), and was found to be replicate stably as an independent replicon in *Rhodococcus equi, Rhodococcus globulerus,* and *Rhodococcus erythropolis*. Since then, a dozen or so *Rhodococcus* cryptic plasmids and derivative shuttle vectors have been constructed (21, 29, 30, 39, 56, 69, 83, 98, 134, 147, 167), some of which include novel features such as; 1) the ability to

be mobilized directly from *E. coli* to *Rhodococcus*, bypassing often problematic rhodococci transformations (167); and 2) selective curing or integration of the plasmid based on replicon temperature sensitivity (98).

Post-genomic era tools for global *Rhodococcus* gene and protein expression analysis include DNA microarrays and proteomics (comparison of the 2-D mobility of expressed proteins). Studies using DNA microarrays to address questions in *Rhodococcus* biology have not yet been published, though our lab is currently addressing aromatic compound metabolism of *Rhodococcus aetherivorans* I24 by this method (J. Parker et al., unpublished). Reasons why *Rhodococcus* DNA microarray analyses have been so sparse include the lack of published genomic sequence, and difficulties in RNA isolation from rhodococci (89). In contrast, protocols for the isolation of proteins- in particular the purification of desirable enzymes- are well established for this group of bacteria. One of the latest articles reports the purification of an alcohol dehydrogenase from *Rhodococcus ruber* that is stable in the presence of 50% acetone and 80% isopropanol (81). A proteomics approach was taken very recently to identify genes specifically involved in phthalate metabolism in *Rhodococcus* sp. strain TFB (160). This method can prove to be a powerful analytic tool, when combined with genomic sequence information, to decipher the multitudes of seemingly redundant gene functions encoded in some *Rhodococcus* genomes, as is the case with biphenyl metabolism in *Rhodococcus* sp. RHA1 (132).

1.2. Bacterial plasmids and conjugation.

1.2.1. Biology of *Rhodococcus* extrachromosomal replicons. The term, plasmid, was coined by Joshua Lederberg in 1952 to describe an extrachromosomal genetic element that can replicate autonomously (94). Plasmids can adopt either circular or linear topologies, in turn leading to differences in how they replicate and transfer. Plasmids can also vary in size. Miniplasmids can be engineered to carry only the essential replication region and selectible marker in less than 1,000 bp of DNA sequence. Naturally occuring megaplasmids can exceed 1 megabasepairs in size (16, 100), comparable to the gene coding capacity of a yeast chromosome. Plasmids profoundly influence the fitness of their host organisms, both negatively and positively. Uncontrolled plasmid proliferation

can be detrimental to the host, as plasmid replication is dependent on host factors (15, 165, 166). Conversely, the presence of plasmid-encoded enzymatic pathways often confer advantages to the host organisms, enabling their survival in otherwise inhospitable environments. In essence, bacteria that harbor plasmids exhibit dynamic genomes, which allows for rapid adaptation under selective pressures.

It would not be an overstatement to say that horizontal gene transfer (HGT) is a major facilitator of bacterial evolution. Nearly all bacterial species examined to date are capable of at least one of the following modes of exogenous DNA exchange, 1) uptake and direct transformation, 2) transposition via mobile genetic elements like transposons, 3) transduction involving phage-intermediates, and 4) conjugation, referring to the exchange of genetic material most commonly in the form of plasmids. The exchange of plasmids through bacterial conjugation has been, and still is, one of the intensely studied forms of HGT with implications in human health, both negatively associated with plasmid-borne antibiotic resistances and virulence factors, and positively associated with plasmid-borne antibiotic production and pathways for bioremediation of xenobiotics.

Rhodococci are hosts to one of the most diverse plasmid-based horizontal gene pool. Tables 1.1. and 1.2 summarize all plasmids and associated properties found in rhodococci; these tables have been modified from an earlier table from the Larkin review (1998) to include recent findings. Some of these plasmids have been found to be conjugative, however, determinants governing their transmission are largely unknown. As this thesis directly addresses determinants of *Rhodococcus* plasmid conjugation, it will be prefaced here by what is currently known about the conjugation of a well-characterized Gram-negative bacterial plasmid, the F factor, and how Gram-positive conjugation, especially plasmids encoded by *Actinomycetales*, differ from this paradigm.

Table 1.1. Smaller (<20 kb) *Rhodococcus* plasmids and associated properties.

Strain	Plasmid[a]	Size-kb	Description	Reference
R. opacus MR11	pHG31-b	17	Cryptic circular plasmid	(145)
Rhodococcus sp. H13-1A	pMVS200	20	Cryptic circular plasmid	(149)
Rhodococcus sp. H13-1A	pMVS300	13	Cryptic circular plasmid	(149)
R. erythropolis NI86/21	pFAJ2600	6	Cryptic circular plasmid; rolling-circle replicase; sequenced	(29)

continued on next page

R. rhodochrous NCIMB13064	pKA22 pKA4	5 9	Cryptic circular plasmid; theta replicase; sequenced (pKA22); cryptic circular plasmid (pKA4)	(84)
R. erythropolis AN12	**pAN12**	6	Cryptic circular plasmids; rolling-circle replicase; sequenced	(83)
R. rhodochrous B264	**pB264-1**	5	Cryptic circular plasmid; theta replicase; highly similar to pKA22; sequenced; mobilizable	(98)
R. erythropolis JCM2895	pRE2895	≥ 1.9	Cryptic circular plasmid; theta replicase	(119)
R. erythropolis DSM8424	pRE8424	6	Cryptic circular plasmid; rolling-circle replicase; 91% identical to pAN12; sequenced	(118)
R. erythropolis PR4	pREC2	4	Cryptic circular plasmid; theta replicase; sequenced	(144)

a. Plasmids that can be mobilized are indicated (**bold**). Plasmids not indicated as such could be mobilizable though this property may not have been assayed.

Table 1.1. Larger (>20 kb) *Rhodococcus* plasmids and associated properties.

Strain	Plasmid(s)[a]	Size-kb	Description	Reference
R. opacus MR11	pHG31-a	140	Cryptic; circular	(145)
R. opacus MR22	pHG33	110	Cryptic; circular	(145)
Rhodococcus sp. TE1	**pTE1**	77	Atrazine and EPTC degradation; circular	(8, 154)
R. fascians NCPPB10675	**pRF2**	>100	Chloramphenicol and cadmium heavy metal resistance; circular	(34)
R. erythropolis ATCC12674	pDA20	>100	Arsenic and cadmium resistance; linear	(22)
R. opacus MR11	**pHG201**, pGH302, pGH203	270, 400, 420	Conjugative hydrogen autotrophy on pHG201; linear	(68)
R. opacus MR22	pGH204, **pGH205**, pGH206	180, 280, 510	Conjugative hydrogen autotrophy on pGH205; linear	(68)
R. equi ATCC33701	pREAT701	85	Virulence genes; circular	(152, 153)
R. equi L1	pREL1	90	Virulence genes; circular	(152, 153)
R. rhodochrous CTM	pTC1	111	2-methylaniline degradation; circular	(142)
R. rhodochrous CTM	pTC2	20	Cryptic; circular	(142)
R. fascians D188	**pD188** **pFiD188**	138; >200	Cadmium resistance (pD188); leafy gall formation and fasciation loci; linear (pFiD188)	(18, 19, 34)
R. erythropolis IGTS8	pSOX	150	Biodesulfurization genes (*dsz/sox*); circular	(31)
R. erythropolis BD2	**pBD2**	210	Isopropylbenzene metabolism; linear; conjugative; completely sequenced	(23, 150)
R. rhodochrous NCIMB13064	**pRTL1**, pRTL2	100, 80	Circular and chloroalkane degradation (pRTL1); cryptic and circular (pRTL2)	(86)
R. equi 103	pOTS	85	virulence genes; circular plasmid	(28)

continued on next page

Rhodococcus sp. RHA1	pRHA1, pRHA2, pRHA3	1,100, 390, 280	All linear plasmids; biphenyl metabolism (pRHA1 and pRHA2), cryptic and sequenced (pRHA3)	(106, 170)
R. erythropolis TA421	pTA421	500	Biphenyl degradation; linear	(82)
R. corallinus B-276	pNC10, pNC20, pNC30, pNC40	70, 85, 185, 235	All linear; pNC30 associated with propene metabolism	(138)
Rhodococcus sp. MP50	pMP50-40	40	Aromatic nitrile degradation, circular	(163)
Rhodococcus sp. DK17	pDK1, pDK2	380, 330	Topology and conjugativity not determined; alkylbenzene metabolism (pDK2)	(74)
R. opacus 1CP	p1CP	740	Linear; chlorocatechol catabolism	(79)
R. aetherivorans I24	**pNid**, pI24	50 340	Circular and naphthalene degradation (pNid); linear and toluene metabolism (pI24)	(135)
R. erythropolis PR4	pREL1, **pREC1**, pREC2	271, 104, 4	Linear and associated with alkane metabolism (pREL1); circular and cryptic (pREC1); circular and cryptic (pREC2)	(144)

a. Plasmids determined to be conjugative have been highlighted (**bold**). Plasmids not indicated as such, may be conjugative, though this property may not have been assayed.

1.2.2. The F-factor, a paradigm in bacterial conjugation. Bacterial conjugation between cells of the *Escherichia coli* strain K-12 was discovered by J. Lederberg and E. Tatum in 1946 (95). It was shown a few years later by B. Davis that physical contact between the cells is necessary for bacterial mating, a core principle that extends to virtually all genera of bacteria (25). Significantly, in 1952, W. Hayes demonstrated another conjugation principle- that DNA transfer process proceeds unidirectionally from a donor to a recipient cells, and is mediated by a sex factor named F, for fertility (57). Lederberg then proposed that the F-sex factor might be a previously unrecognized form of extrachromosomal genetic material, which he termed a plasmid (94), and this was conclusively shown to be the case by Marmur et al. in 1961 (105).

Since these early discoveries, researchers have shown that one-third of the genes (the *tra* and *trb* loci) on the 99 kb F plasmid encode two types of transfer functions- those involved in initiating and stabilizing donor-recipient contact, or mating pair formation (Mpf), and those involved in DNA transfer and replication (Dtr) (45). Mating begins when a thin flexible extracellular filament, called a pilus, from the (F+) donor cell makes contact with the (F-) recipient. The flexible nature of the F pilus allows for *E. coli* to mate in liquid suspension. As part of the Mpf, the pilus is a polymer of homogenous subunits called pilins, which are encoded by the F-plasmid *traA* gene (111). A number of

the *tra* genes encode pilin processing, translocation, and assembly functions (47). Though it has been proposed in the past that the pilus is the physical bridge through which the plasmid DNA traverses, it is now thought that the pilus brings about the physical contact of the two cells by retraction, mediated by the depolymerization of its subunits. The actual DNA transfer may occur through electron-dense regions of the cell envelopes, known as conjugation junctions, established when the intimate contact of the donor and recipient cells is achieved (40). This mating pair stability is known to involve at least the *traN* gene product found at the outer membrane of the donor cell, and its association with recipient OmpA (outer membrane protein) protein (75), as well as the inner membrane protein, TraG (104).

F is transferred as a single-stranded (ssDNA) molecule in the 5' to 3' orientation (63, 126), and is re-established as a double-stranded DNA (dsDNA) in the recipient cell following rolling-circle like DNA replication (88). The DNA processing is initiated at a region of the transferring strand called the origin of transfer, or *ori*T (158). The *ori*T sequence is bound by the relaxosome, a protein complex consisting of the plasmid-encoded *traI* relaxase/helicase (108, 137) and *traY* DNA binding protein (161), as well as a host-encoded histone-like protein called IHF, for integration host factor (164). It is the TraI relaxase/helicase which first recognizes a conserved 19 bp sequence (TTTGCGTGGGGTGTGGTGC) called the *nic* site within *ori*T, then catalyses the scission reaction at the bolded thymine residue, the molecular mechanism of which will be discussed in greater detail in Chapter 3. TraY and IHF were found to stimulate the nicking activity of the TraI protein by creating DNA bending (122), and are preloaded onto the *ori*T before TraI (60). The TraD inner membrane gene product is a coupling protein and DNA transporter (131), linking the relaxosome-*ori*T nucleoprotein complex (Dtr) to the Mpf via its interaction with the TraM DNA binding protein, which also binds to the *ori*T though it is not thought to be part of the relaxosome (35). Much is known about the how the replacement strand is synthesized in the recipient, as well as the organization and expression of the "early" genes located at the leader region (the first portion of the plasmid transferred to the recipient), which aid in plasmid establishment. For details on these processes, the readers are referred to the excellent review by Zechner et al. (171).

1.2.3. Important differences in Gram-positive plasmid conjugation: Actinomycetes appear to be exceptions to the rule.

It is generally thought that the majority of Gram-positive plasmids transfer in a mechanistically conserved fashion as their Gram-negative counterpart with respect to the DNA processing and transport (Dtr) and the need for cell-to-cell contact (52). Indeed, plasmid-encoded components such as the relaxase and the *ori*T are found to be conserved not only in sequence but in function for Gram-positive plasmids from *Streptococcus agalactiae* pIP501 (80, 87) and *Enterococcus faecalis* pRE25 (143). However, since no gene encoding for a pilus component has been found in Gram-positive bacteria, it is assumed that different mechanisms than the Mpf exist to bring cells together. Consistent with the lack of pilus, most Gram-positive microorganisms require a solid surface for plasmid conjugation. At least two different methods to establish cell contact are beginning to be elucidated for Gram-positive bacteria; the pheromone-induced conjugation system exemplified by the enterococci plasmid, pAD1 (38), and the aggregation-based conjugation systems of *Bacillus thuringienis* pXO16 (66) and *Lactococcus lactis* pRS01 (110).

For the Gram-positive bacteria of the genus *Streptomyces*, genetic recombination has been known for almost as long as that of *E. coli* (59). Conjugative transfer of various streptomycete plasmids, such as pIJ101, from isolated donors to a lawn of recipients, visibly results in the formation of pocks (73). Unlike other small mobilizable plasmids which depend on host or other plasmid encoded genes for transmission, pIJ101 does not appear to be dependent on the action of a relaxosome. In fact, only the function of one plasmid encoded gene called *tra* is required for pIJ101 conjugation (70, 71, 133). The pIJ101 Tra protein is a membrane protein not related to relaxases/helicases described earlier; rather, it resembles the DNA translocators FtsK/SpoIIIE, which function in chromosome segregation in bacilli and *E. coli*, respectively (42). Nor does the pIJ101 *cis*-acting site appear to be a substrate for nicking activity (36, 37). Instead, it has been proposed that these *Streptomyces* plasmids are unique in the transfer of plasmid DNA to the recipients in a dsDNA form (52, 171).

1.3. Thesis objectives and chapter outline.

The existence of various rhodococci megaplasmids and associated properties suggests that it may be possible to engineer strains which carry combinatorial metabolic pathways through bacterial conjugation. This would aid in combining genetic traits associated with large or spatially separated loci that would otherwise be difficult to clone piecemeal. We have termed this concept "megapathway shuffling", illustrated below in Figure 1.4.

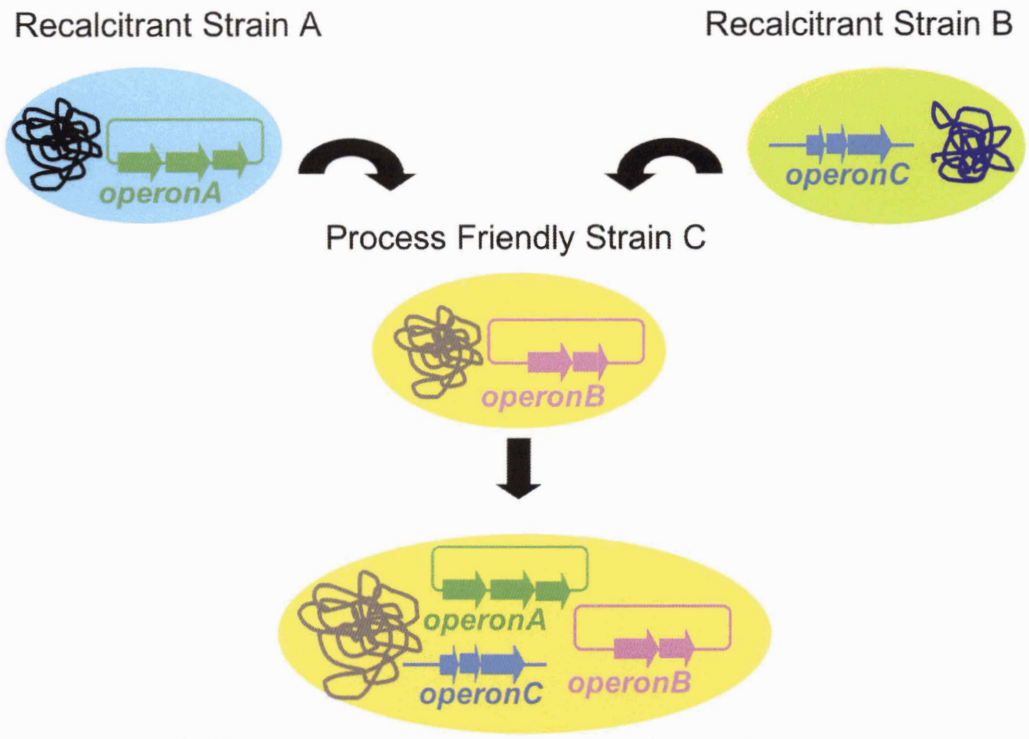


Figure 1.4. Metabolic engineering of rhodococci through megaplasmid conjugation.

With this ultimate objective in mind, we have examined the plasmid array and conjugative properties of the *Rhodococcus erythropolis* strain AN12, originally isolated from industrial wastewater sludge. AN12 utilizes the aromatic compound aniline as its sole carbon source, and was found to encode a small cryptic plasmid called pAN12 (67). It was not known whether the *R. erythropolis* AN12 genome encodes other plasmids, nor whether any of its plasmids might be conjugative. Through characterization of AN12 plasmids, we endeavored also to understand how might the conjugation system of

rhodococci be similar or dissimilar to those defined for Gram-negative and positive plasmids, such as F and pIJ101, described earlier. The chapters of this thesis addresses the following specific aims: 1) Determining whether *R. erythropolis* AN12 encode other plasmids, and if so, 2) determining whether any AN12 plasmid is conjugative, and 3) defining *cis*- and/or *trans*-acting determinants for conjugative plasmids. Chapter 2 discusses the discovery of two conjugative AN12 megaplasmids, pREA400 and pREA250. Chapter 3 discusses generating a transfer-deficient mutant of AN12 using a novel site specific targeted gene disruption technique. Chapter 4 discusses the mobilization of the small cryptic pAN12 and examines its dependence on one of the megaplasmids for its transfer. Chapter 5 outlines future experiments.

1.4. References.

1. **Adams, J., and S. Bradley.** 1963. Recombination events in the bacterial genus *Nocardia*. Science **140:**1392-1394.
2. **Adams, J. N.** 1964. Recombination between *Nocardia Erythropolis* and *Nocardia Canicruria*. J Bacteriol **88:**865-76.
3. **Ashour, J., and M. K. Hondalus.** 2003. Phenotypic mutants of the intracellular actinomycete *Rhodococcus equi* created by in vivo *Himar1* transposon mutagenesis. J Bacteriol **185:**2644-52.
4. **Ashraf, W., A. Mihdhir, and J. C. Murrell.** 1994. Bacterial oxidation of propane. FEMS Microbiol Lett **122:**1-6.
5. **Asturias, J. A., and K. N. Timmis.** 1993. Three different 2,3-dihydroxybiphenyl-1,2-dioxygenase genes in the gram-positive polychlorobiphenyl-degrading bacterium *Rhodococcus globerulus* P6. J Bacteriol **175:**4631-40.
6. **Banerjee, A., R. Sharma, and U. C. Banerjee.** 2002. The nitrile-degrading enzymes: current status and future prospects. Appl Microbiol Biotechnol **60:**33-44.
7. **Beard, C. B., E. M. Dotson, P. M. Pennington, S. Eichler, C. Cordon-Rosales, and R. V. Durvasula.** 2001. Bacterial symbiosis and paratransgenic control of vector-borne Chagas disease. Int J Parasitol **31:**621-7.
8. **Behki, R., E. Topp, W. Dick, and P. Germon.** 1993. Metabolism of the herbicide atrazine by *Rhodococcus* strains. Appl Environ Microbiol **59:**1955-9.
9. **Brennan, B. A., J. G. Cummings, D. B. Chase, I. M. Turner, Jr., and M. J. Nelson.** 1996. Resonance Raman spectroscopy of nitrile hydratase, a novel iron-sulfur enzyme. Biochemistry **35:**10068-77.
10. **Brown, N. A.** 1926. Sweet pea fasciation, a form of crowngall. Phytopathology **17:**29-30.
11. **Brownell, G. H., and K. Denniston.** 1984. Genetics of the nocardioform

bacteria., p. 201-228. *In* M. Goodfellow, M. Mordarski, and S. T. Williams (ed.), The biology of the actinomycetes. Academic Press, London.

12. **Buckland, B. C., S. W. Drew, N. C. Connors, M. M. Chartrain, C. Lee, P. M. Salmon, K. Gbewonyo, W. Zhou, P. Gailliot, R. Singhvi, R. C. Olewinski, Jr., W. J. Sun, J. Reddy, J. Zhang, B. A. Jackey, C. Taylor, K. E. Goklen, B. Junker, and R. L. Greasham.** 1999. Microbial conversion of indene to indandiol: a key intermediate in the synthesis of CRIXIVAN. Metab Eng **1:**63-74.
13. **Bunch, A. W.** 1998. Biotransformation of nitriles by rhodococci. Antonie Van Leeuwenhoek **74:**89-97.
14. **Chang, A. C., and S. N. Cohen.** 1974. Genome construction between bacterial species in vitro: replication and expression of *Staphylococcus* plasmid genes in *Escherichia coli*. Proc Natl Acad Sci U S A **71:**1030-4.
15. **Chittenden, T., A. Frey, and A. J. Levine.** 1991. Regulated replication of an episomal simian virus 40 origin plasmid in COS7 cells. J Virol **65:**5944-51.
16. **Coenye, T., and P. Vandamme.** 2003. Simple sequence repeats and compositional bias in the bipartite *Ralstonia solanacearum* GMI1000 genome. BMC Genomics **4:**10.
17. **Commeyras, A., A. Arnaud, P. Galzy, and J.-C. Jallageas.** 1976. Process for the preparation of amides by biological hydrolysis. US patent 4,001,081.
18. **Crespi, M., E. Messens, A. B. Caplan, M. van Montagu, and J. Desomer.** 1992. Fasciation induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. Embo J **11:**795-804.
19. **Crespi, M., D. Vereecke, W. Temmerman, M. Van Montagu, and J. Desomer.** 1994. The *fas* operon of *Rhodococcus fascians* encodes new genes required for efficient fasciation of host plants. J Bacteriol **176:**2492-501.
20. **Curragh, H., O. Flynn, M. J. Larkin, T. M. Stafford, J. T. Hamilton, and D. B. Harper.** 1994. Haloalkane degradation and assimilation by *Rhodococcus rhodochrous* NCIMB 13064. Microbiology **140 (Pt 6):**1433-42.
21. **Dabbs, E. R., B. Gowan, and S. J. Andersen.** 1990. Nocardioform arsenic resistance plasmids and construction of *Rhodococcus* cloning vectors. Plasmid **23:**242-7.
22. **Dabbs, E. R., and G. J. Sole.** 1988. Plasmid-borne resistance to arsenate, arsenite, cadmium, and chloramphenicol in a *Rhodococcus* species. Mol Gen Genet **211:**148-54.
23. **Dabrock, B., M. Kesseler, B. Averhoff, and G. Gottschalk.** 1994. Identification and characterization of a transmissible linear plasmid from *Rhodococcus erythropolis* BD2 that encodes isopropylbenzene and trichloroethene catabolism. Appl Environ Microbiol **60:**853-60.
24. **Darzins, A., L. Xi, J. Childs, D. Monticello, and C. Squires.** 1999. DSZ gene expression in *Pseudomonas* hosts. US patent 5952208.
25. **Davis, B. D.** 1950. Nonfiltrability of the agents of genetic recombination in *Escherichia coli*. J Bacteriol **60:**507-8.
26. **de Carvalho, C. C., and M. M. da Fonseca.** 2005. Degradation of hydrocarbons and alcohols at different temperatures and salinities by *Rhodococcus erythropolis* DCL14. FEMS Microbiol Ecol **51:**389-99.
27. **de Carvalho, C. C., and M. M. da Fonseca.** 2005. The remarkable *Rhodococcus*

erythropolis. Appl Microbiol Biotechnol **67:**715-26.

28. **de la Pena-Moctezuma, A., and J. F. Prescott.** 1995. A physical map of the 85 kb virulence plasmid of *Rhodococcus equi* 103. Can J Vet Res **59:**229-31.
29. **De Mot, R., I. Nagy, A. De Schrijver, P. Pattanapipitpaisal, G. Schoofs, and J. Vanderleyden.** 1997. Structural analysis of the 6 kb cryptic plasmid pFAJ2600 from *Rhodococcus erythropolis* NI86/21 and construction of *Escherichia coli-Rhodococcus* shuttle vectors. Microbiology **143 (Pt 10):**3137-47.
30. **Denis-Larose, C., H. Bergeron, D. Labbe, C. W. Greer, J. Hawari, M. J. Grossman, B. M. Sankey, and P. C. Lau.** 1998. Characterization of the basic replicon of *Rhodococcus* plasmid pSOX and development of a *Rhodococcus-Escherichia coli* shuttle vector. Appl Environ Microbiol **64:**4363-7.
31. **Denome, S. A., C. Oldfield, L. J. Nash, and K. D. Young.** 1994. Characterization of the desulfurization genes from *Rhodococcus* sp. strain IGTS8. J Bacteriol **176:**6707-16.
32. **Denome, S. A., E. S. Olson, and K. D. Young.** 1993. Identification and Cloning of Genes Involved in Specific Desulfurization of Dibenzothiophene by *Rhodococcus* sp. Strain IGTS8. Appl Environ Microbiol **59:**2837-2843.
33. **Denome, S. A., and K. D. Young.** 1995. Identification and activity of two insertion sequence elements in *Rhodococcus* sp. strain IGTS8. Gene **161:**33-8.
34. **Desomer, J., P. Dhaese, and M. Van Montagu.** 1988. Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. J Bacteriol **170:**2401-5.
35. **Disque-Kochem, C., and B. Dreiseikelmann.** 1997. The cytoplasmic DNA-binding protein TraM binds to the inner membrane protein TraD in vitro. J Bacteriol **179:**6133-7.
36. **Ducote, M. J., and G. S. Pettis.** 2006. An in vivo assay for conjugation-mediated recombination yields novel results for *Streptomyces* plasmid pIJ101. Plasmid **55:** 242-8.
37. **Ducote, M. J., S. Prakash, and G. S. Pettis.** 2000. Minimal and contributing sequence determinants of the *cis*-acting locus of transfer (*clt*) of streptomycete plasmid pIJ101 occur within an intrinsically curved plasmid region. J Bacteriol **182:**6834-41.
38. **Dunny, G. M., B. A. Leonard, and P. J. Hedberg.** 1995. Pheromone-inducible conjugation in *Enterococcus faecalis*: interbacterial and host-parasite chemical communication. J Bacteriol **177:**871-6.
39. **Duran, R.** 1998. New shuttle vectors for *Rhodococcus* sp. R312 (formerly *Brevibacterium* sp. R312), a nitrile hydratase producing strain. J Basic Microbiol **38:**101-6.
40. **Durrenberger, M. B., W. Villiger, and T. Bachi.** 1991. Conjugational junctions: morphology of specific contacts in conjugating *Escherichia coli* bacteria. J Struct Biol **107:**146-56.
41. **Durvasula, R. V., A. Gumbs, A. Panackal, O. Kruglov, S. Aksoy, R. B. Merrifield, F. F. Richards, and C. B. Beard.** 1997. Prevention of insect-borne disease: an approach using transgenic symbiotic bacteria. Proc Natl Acad Sci U S A **94:**3274-8.
42. **Errington, J., J. Bath, and L. J. Wu.** 2001. DNA transport in bacteria. Nat Rev

Mol Cell Biol **2:**538-45.

43. **Fernandes, P. J., J. A. Powell, and J. A. Archer.** 2001. Construction of *Rhodococcus* random mutagenesis libraries using Tn*5* transposition complexes. Microbiology **147:**2529-36.
44. **Finnerty, W. R.** 1992. The biology and genetics of the genus *Rhodococcus*. Annu Rev Microbiol **46:**193-218.
45. **Firth, N., K. Ippen-Ihler, and R. A. Skurray.** 1996. Structure and function of the F factor and mechanism of conjugation., p. 2377-2401. *In* F. C. Neidhard, R. Curtiss III., J. L. Ingraham, E. C. C. Lin, K. S. Low, B. Magasanik, W. S. Reznikoff, M. Schaechter, and H. E. Umbarger (ed.), Escherichia coli and Salmonella, vol. 2. ASM Press, Washington, D.C.
46. **Folsom, B. R., D. R. Schieche, P. M. DiGrazia, J. Werner, and S. Palmer.** 1999. Microbial desulfurization of alkylated dibenzothiophenes from a hydrodesulfurized middle distillate by *Rhodococcus erythropolis* I-19. Appl Environ Microbiol **65:**4967-72.
47. **Frost, L. S., K. Ippen-Ihler, and R. A. Skurray.** 1994. Analysis of the sequence and gene products of the transfer region of the F sex factor. Microbiol Rev **58:**162-210.
48. **Gallagher, J. R., E. S. Olson, and D. C. Stanley.** 1993. Microbial desulfurization of dibenzothiophene: a sulfur-specific pathway. FEMS Microbiol Lett **107:**31-5.
49. **Goodfellow, M., M. Mordarski, and S. T. Williams.** 1984. The Biology of the actinomycetes. Academic Press, London ; Orlando.
50. **Gowan, B., and E. R. Dabbs.** 1994. Identification of DNA involved in *Rhodococcus* chromosomal conjugation and self-incompatibility. FEMS Microbiol Lett **115:**45-50.
51. **Gray, K. A., O. S. Pogrebinsky, G. T. Mrachko, L. Xi, D. J. Monticello, and C. H. Squires.** 1996. Molecular mechanisms of biocatalytic desulfurization of fossil fuels. Nat Biotechnol **14:**1705-9.
52. **Grohmann, E., G. Muth, and M. Espinosa.** 2003. Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev **67:**277-301, table of contents.
53. **Grossman, M. J., M. K. Lee, R. C. Prince, K. K. Garrett, G. N. George, and I. J. Pickering.** 1999. Microbial desulfurization of a crude oil middle-distillate fraction: analysis of the extent of sulfur removal and the effect of removal on remaining sulfur. Appl Environ Microbiol **65:**181-8.
54. **Grossman, M. J., M. K. Lee, R. C. Prince, V. Minak-Bernero, G. N. George, and I. J. Pickering.** 2001. Deep desulfurization of extensively hydrodesulfurized middle distillate oil by *Rhodococcus* sp. strain ECRD-1. Appl Environ Microbiol **67:**1949-52.
55. **Gutierrez, T., R. P. Learmonth, P. D. Nichols, and I. Couperwhite.** 2003. Comparative benzene-induced fatty acid changes in a *Rhodococcus* species and its benzene-sensitive mutant: possible role of myristic and oleic acids in tolerance. J Chem Ecol **29:**2369-78.
56. **Hashimoto, Y., M. Nishiyama, F. Yu, I. Watanabe, S. Horinouchi, and T. Beppu.** 1992. Development of a host-vector system in a *Rhodococcus* strain and its use for expression of the cloned nitrile hydratase gene cluster. J Gen Microbiol

138:1003-10.

57. **Hayes, W.** 1952. Recombination in *Bact. coli* K 12; unidirectional transfer of genetic material. Nature **169:**118-9.
58. **Hirasawa, K., Y. Ishii, M. Kobayashi, K. Koizumi, and K. Maruhashi.** 2001. Improvement of desulfurization activity in *Rhodococcus erythropolis* KA2-5-1 by genetic engineering. Biosci Biotechnol Biochem **65:**239-46.
59. **Hopwood, D. A.** 1959. Linkage and the mechanism of recombination in *Streptomyces coelicolor*. Ann N Y Acad Sci **81:**887-98.
60. **Howard, M. T., W. C. Nelson, and S. W. Matson.** 1995. Stepwise assembly of a relaxosome at the F plasmid origin of transfer. J Biol Chem **270:**28381-6.
61. **Huang, W., J. Jia, J. Cummings, M. Nelson, G. Schneider, and Y. Lindqvist.** 1997. Crystal structure of nitrile hydratase reveals a novel iron centre in a novel fold. Structure **5:**691-9.
62. **Hughes, J., Y. C. Armitage, and K. C. Symes.** 1998. Application of whole cell rhodococcal biocatalysts in acrylic polymer manufacture. Antonie Van Leeuwenhoek **74:**107-18.
63. **Ihler, G., and W. D. Rupp.** 1969. Strand-specific transfer of donor DNA during conjugation in *E. coli*. Proc Natl Acad Sci U S A **63:**138-43.
64. **Izumi, Y., T. Ohshiro, H. Ogino, Y. Hine, and M. Shimao.** 1994. Selective Desulfurization of Dibenzothiophene by *Rhodococcus erythropolis* D-1. Appl Environ Microbiol **60:**223-226.
65. **Jager, W., A. Schafer, J. Kalinowski, and A. Puhler.** 1995. Isolation of insertion elements from gram-positive *Brevibacterium*, *Corynebacterium* and *Rhodococcus* strains using the *Bacillus subtilis sacB* gene as a positive selection marker. FEMS Microbiol Lett **126:**1-6.
66. **Jensen, G. B., A. Wilcks, S. S. Petersen, J. Damgaard, J. A. Baum, and L. Andrup.** 1995. The genetic basis of the aggregation system in *Bacillus thuringiensis* subsp. *israelensis* is located on the large conjugative plasmid pXO16. J Bacteriol **177:**2914-7.
67. **Kalkus, J., C. Dorrie, D. Fischer, M. Reh, and H. G. Schlegel.** 1993. The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini. J Gen Microbiol **139:**2055-65.
68. **Kalkus, J., M. Reh, and H. G. Schlegel.** 1990. Hydrogen autotrophy of *Nocardia opaca* strains is encoded by linear megaplasmids. J Gen Microbiol **136:**1145-51.
69. **Kalscheuer, R., M. Arenskotter, and A. Steinbuchel.** 1999. Establishment of a gene transfer system for *Rhodococcus opacus* PD630 based on electroporation and its application for recombinant biosynthesis of poly(3-hydroxyalkanoic acids). Appl Microbiol Biotechnol **52:**508-15.
70. **Kendall, K. J., and S. N. Cohen.** 1988. Complete nucleotide sequence of the *Streptomyces lividans* plasmid pIJ101 and correlation of the sequence with genetic properties. J Bacteriol **170:**4634-51.
71. **Kendall, K. J., and S. N. Cohen.** 1987. Plasmid transfer in *Streptomyces lividans*: identification of a *kil-kor* system associated with the transfer region of pIJ101. J Bacteriol **169:**4177-83.
72. **Kertesz, M. A., and C. Wietek.** 2001. Desulfurization and desulfonation:

applications of sulfur-controlled gene expression in bacteria. Appl Microbiol Biotechnol **57:**460-6.

73. **Kieser, T., D. A. Hopwood, H. M. Wright, and C. J. Thompson.** 1982. pIJ101, a multi-copy broad host-range *Streptomyces* plasmid: functional analysis and development of DNA cloning vectors. Mol Gen Genet **185:**223-8.
74. **Kim, D., Y. S. Kim, S. K. Kim, S. W. Kim, G. J. Zylstra, Y. M. Kim, and E. Kim.** 2002. Monocyclic aromatic hydrocarbon degradation by *Rhodococcus* sp. strain DK17. Appl Environ Microbiol **68:**3270-8.
75. **Klimke, W. A., C. D. Rypien, B. Klinger, R. A. Kennedy, J. M. Rodriguez-Maillard, and L. S. Frost.** 2005. The mating pair stabilization protein, TraN, of the F plasmid is an outer-membrane protein with two regions that are important for its function in conjugation. Microbiology **151:**3527-40.
76. **Kobayashi, M., T. Nagasawa, and H. Yamada.** 1992. Enzymatic synthesis of acrylamide: a success story not yet over. Trends Biotechnol **10:**402-8.
77. **Kobayashi, M., M. Nishiyama, T. Nagasawa, S. Horinouchi, T. Beppu, and H. Yamada.** 1991. Cloning, nucleotide sequence and expression in *Escherichia coli* of two cobalt-containing nitrile hydratase genes from *Rhodococcus rhodochrous* J1. Biochim Biophys Acta **1129:**23-33.
78. **Kobayashi, M., T. Onaka, Y. Ishii, J. Konishi, M. Takaki, H. Okada, Y. Ohta, K. Koizumi, and M. Suzuki.** 2000. Desulfurization of alkylated forms of both dibenzothiophene and benzothiophene by a single bacterial strain. FEMS Microbiol Lett **187:**123-6.
79. **Konig, C., D. Eulberg, J. Groning, S. Lakner, V. Seibert, S. R. Kaschabek, and M. Schlomann.** 2004. A linear megaplasmid, p1CP, carrying the genes for chlorocatechol catabolism of *Rhodococcus opacus* 1CP. Microbiology **150:**3075-87.
80. **Kopec, J., A. Bergmann, G. Fritz, E. Grohmann, and W. Keller.** 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:**401-9.
81. **Kosjek, B., W. Stampfer, M. Pogorevc, W. Goessler, K. Faber, and W. Kroutil.** 2004. Purification and characterization of a chemotolerant alcohol dehydrogenase applicable to coupled redox reactions. Biotechnol Bioeng **86:**55-62.
82. **Kosono, S., M. Maeda, F. Fuji, H. Arai, and T. Kudo.** 1997. Three of the seven *bphC* genes of *Rhodococcus erythropolis* TA421, isolated from a termite ecosystem, are located on an indigenous plasmid associated with biphenyl degradation. Appl Environ Microbiol **63:**3282-5.
83. **Kostichka, K., L. Tao, M. Bramucci, J. F. Tomb, V. Nagarajan, and Q. Cheng.** 2003. A small cryptic plasmid from *Rhodococcus erythropolis*: characterization and utility for gene expression. Appl Microbiol Biotechnol **62:**61-8.
84. **Kulakov, L. A., M. J. Larkin, and A. N. Kulakova.** 1997. Cryptic plasmid pKA22 isolated from the naphthalene degrading derivative of *Rhodococcus rhodochrous* NCIMB13064. Plasmid **38:**61-9.
85. **Kulakova, A. N., M. J. Larkin, and L. A. Kulakov.** 1997. The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064.

Microbiology **143 (Pt 1):**109-15.

86. **Kulakova, A. N., T. M. Stafford, M. J. Larkin, and L. A. Kulakov.** 1995. Plasmid pRTL1 controlling 1-chloroalkane degradation by *Rhodococcus rhodochrous* NCIMB13064. Plasmid **33:**208-17.
87. **Kurenbach, B., D. Grothe, M. E. Farias, U. Szewzyk, and E. Grohmann.** 2002. The *tra* region of the conjugative plasmid pIP501 is organized in an operon with the first gene encoding the relaxase. J Bacteriol **184:**1801-5.
88. **Lanka, E., and B. M. Wilkins.** 1995. DNA processing reactions in bacterial conjugation. Annu Rev Biochem **64:**141-69.
89. **Larkin, M. J., R. De Mot, L. A. Kulakov, and I. Nagy.** 1998. Applied aspects of *Rhodococcus* genetics. Antonie Van Leeuwenhoek **74:**133-53.
90. **Larkin, M. J., L. A. Kulakov, and C. C. Allen.** 2005. Biodegradation and *Rhodococcus*--masters of catabolic versatility. Curr Opin Biotechnol **16:**282-90.
91. **Lau, P. C., J. Garnon, D. Labbe, and Y. Wang.** 1996. Location and sequence analysis of a 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate hydrolase-encoding gene (*bpdF*) of the biphenyl/polychlorinated biphenyl degradation pathway in *Rhodococcus* sp. M5. Gene **171:**53-7.
92. **Lechevalier, M. P., A. C. Horan, and H. Lechevalier.** 1971. Lipid composition in the classification of *nocardiae* and *mycobacteria*. J Bacteriol **105:**313-8.
93. **Lechevalier, M. P., H. Lechevalier, and A. C. Horan.** 1973. Chemical characteristics and classification of *nocardiae*. Can J Microbiol **19:**965-72.
94. **Lederberg, J.** 1952. Cell genetics and hereditary symbiosis. Physiol Rev **32:**403-30.
95. **Lederberg, J., and E. L. Tatum.** 1946. Gene recombination in *E. coli*. Nature **158:**558-562.
96. **Lee, M., M. K. Kim, I. Singleton, M. Goodfellow, and S. T. Lee.** 2006. Enhanced biodegradation of diesel oil by a newly identified *Rhodococcus baikonurensis* EN3 in the presence of mycolic acid. J Appl Microbiol **100:**325-33.
97. **Lessard, P. A., X. M. O'Brien, N. A. Ahlgren, S. A. Ribich, and A. J. Sinskey.** 1999. Characterization of IS*1676* from *Rhodococcus erythropolis* SQ1. Appl Microbiol Biotechnol **52:**811-9.
98. **Lessard, P. A., X. M. O'Brien, D. H. Currie, and A. J. Sinskey.** 2004. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. BMC Microbiol **4:**15.
99. **Li, M. Z., C. H. Squires, D. J. Monticello, and J. D. Childs.** 1996. Genetic analysis of the *dsz* promoter and associated regulatory regions of *Rhodococcus erythropolis* IGTS8. J Bacteriol **178:**6409-18.
100. **MacLellan, S. R., R. Zaheer, A. L. Sartor, A. M. MacLean, and T. M. Finan.** 2006. Identification of a megaplasmid centromere reveals genetic structural diversity within the *repABC* family of basic replicons. Mol Microbiol **59:**1559-75.
101. **Maeda, M., S. Y. Chung, E. Song, and T. Kudo.** 1995. Multiple genes encoding 2,3-dihydroxybiphenyl 1,2-dioxygenase in the gram-positive polychlorinated biphenyl-degrading bacterium *Rhodococcus erythropolis* TA421, isolated from a termite ecosystem. Appl Environ Microbiol **61:**549-55.
102. **Magnusson, H.** 1923. Spezifische infektioese penumonie beim fohlen. Ein neuer eitererregar beim pferd. Arch Wiss Prakt Tierheilkd **50:**22-37.

103. **Mangan, M. W., and W. G. Meijer.** 2001. Random insertion mutagenesis of the intracellular pathogen *Rhodococcus equi* using transposomes. FEMS Microbiol Lett **205:**243-6.
104. **Manning, P. A., G. Morelli, and M. Achtman.** 1981. traG protein of the F sex factor of *Escherichia coli* K-12 and its role in conjugation. Proc Natl Acad Sci U S A **78:**7487-91.
105. **Marmur, J., R. Rownd, S. Falkow, L. S. Baron, C. Schildkraut, and P. Doty.** 1961. The nature of intergeneric episomal infection. Proc Natl Acad Sci U S A **47:**972-9.
106. **Masai, E., K. Sugiyama, N. Iwashita, S. Shimizu, J. E. Hauschild, T. Hatta, K. Kimbara, K. Yano, and M. Fukuda.** 1997. The *bphDEF* meta-cleavage pathway genes involved in biphenyl/polychlorinated biphenyl degradation are located on a linear plasmid and separated from the initial *bphACB* genes in *Rhodococcus* sp. strain RHA1. Gene **187:**141-9.
107. **Masai, E., A. Yamada, J. M. Healy, T. Hatta, K. Kimbara, M. Fukuda, and K. Yano.** 1995. Characterization of biphenyl catabolic genes of gram-positive polychlorinated biphenyl degrader *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol **61:**2079-85.
108. **Matson, S. W., and B. S. Morton.** 1991. *Escherichia coli* DNA helicase I catalyzes a site- and strand-specific nicking reaction at the F plasmid *oriT*. J Biol Chem **266:**16232-7.
109. **Matsubara, T., T. Ohshiro, Y. Nishina, and Y. Izumi.** 2001. Purification, characterization, and overexpression of flavin reductase involved in dibenzothiophene desulfurization by *Rhodococcus erythropolis* D-1. Appl Environ Microbiol **67:**1179-84.
110. **Mills, D. A., C. K. Choi, G. M. Dunny, and L. L. McKay.** 1994. Genetic analysis of regions of the *Lactococcus lactis* subsp. *lactis* plasmid pRS01 involved in conjugative transfer. Appl Environ Microbiol **60:**4413-20.
111. **Minkley, E. G., Jr., S. Polen, C. C. Brinton, Jr., and K. Ippen-Ihler.** 1976. Identification of the structural gene for F-pilin. J Mol Biol **108:**111-21.
112. **Minnikin, D. E.** 1991. Chemical principles in the organization of lipid components in the mycobacterial cell envelope. Res Microbiol **142:**423-7.
113. **Minnikin, D. E.** 1982. Lipids: Complex lipids, their chemistry, biosynthesis and roles., p. 95-184. *In* C. Ratledge and J. L. Stanford (ed.), The Biology of Mycobacteria. Academic Press, London.
114. **Monticello, D. J.** 2000. Biodesulfurization and the upgrading of petroleum distillates. Curr Opin Biotechnol **11:**540-6.
115. **Nagasawa, T., K. Takeuchi, and H. Yamada.** 1991. Characterization of a new cobalt-containing nitrile hydratase purified from urea-induced cells of *Rhodococcus rhodochrous* J1. Eur J Biochem **196:**581-9.
116. **Nagy, I., G. Schoofs, J. Vanderleyden, and R. De Mot.** 1997. Transposition of the IS*21*-related element IS*1415* in *Rhodococcus erythropolis*. J Bacteriol **179:**4635-8.
117. **Naito, M., T. Kawamoto, K. Fujino, M. Kobayashi, K. Maruhashi, and A. Tanaka.** 2001. Long-term repeated biodesulfurization by immobilized *Rhodococcus erythropolis* KA2-5-1 cells. Appl Microbiol Biotechnol **55:**374-8.

118. **Nakashima, N., and T. Tamura.** 2004. Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. Appl Environ Microbiol **70:**5557-68.
119. **Nakashima, N., and T. Tamura.** 2004. A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C. Biotechnol Bioeng **86:**136-48.
120. **Nakayama, N., T. Matsubara, T. Ohshiro, Y. Moroto, Y. Kawata, K. Koizumi, Y. Hirakawa, M. Suzuki, K. Maruhashi, Y. Izumi, and R. Kurane.** 2002. A novel enzyme, 2'-hydroxybiphenyl-2-sulfinate desulfinase (DszB), from a dibenzothiophene-desulfurizing bacterium *Rhodococcus erythropolis* KA2-5-1: gene overexpression and enzyme characterization. Biochim Biophys Acta **1598:**122-30.
121. **Nandhagopal, N., A. Yamada, T. Hatta, E. Masai, M. Fukuda, Y. Mitsui, and T. Senda.** 2001. Crystal structure of 2-hydroxyl-6-oxo-6-phenylhexa-2,4-dienoic acid (HPDA) hydrolase (BphD enzyme) from the *Rhodococcus* sp. strain RHA1 of the PCB degradation pathway. J Mol Biol **309:**1139-51.
122. **Nelson, W. C., M. T. Howard, J. A. Sherman, and S. W. Matson.** 1995. The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid *oriT*. J Biol Chem **270:**28374-80.
123. **Nikodinovic, J., J. M. Dinges, S. C. Bergmeier, M. C. McMills, D. L. Wright, and N. D. Priestley.** 2006. Resolution of methyl nonactate by *Rhodococcus erythropolis* under aerobic and anaerobic conditions. Org Lett **8:**443-5.
124. **O'Mahony, R., J. Doran, L. Coffey, O. J. Cahill, G. W. Black, and C. O'Reilly.** 2005. Characterisation of the nitrile hydratase gene clusters of *Rhodococcus erythropolis* strains AJ270 and AJ300 and *Microbacterium* sp. AJ115 indicates horizontal gene transfer and reveals an insertion of IS*1166*. Antonie Van Leeuwenhoek **87:**221-32.
125. **Ochi, K.** 1995. Phylogenetic analysis of mycolic acid-containing wall-chemotype IV actinomycetes and allied taxa by partial sequencing of ribosomal protein AT-L30. Int J Syst Bacteriol **45:**653-60.
126. **Ohki, M., and J. Tomizawa.** 1968. Asymmetric transfer of DNA strands in bacterial conjugation. Cold Spring Harb Symp Quant Biol **33:**651-8.
127. **Ohshiro, T., Y. Hine, and Y. Izumi.** 1994. Enzymatic desulfurization of dibenzothiophene by a cell-free system of *Rhodococcus erythropolis* D-1. FEMS Microbiol Lett **118:**341-344.
128. **Ohshiro, T., and Y. Izumi.** 2000. Purification, characterization and crystallization of enzymes for dibenzothiophene desulfurization. Bioseparation **9:**185-8.
129. **Ohshiro, T., Y. Kanbayashi, Y. Hine, and Y. Izumi.** 1995. Involvement of flavin coenzyme in dibenzothiophene degrading enzyme system from *Rhodococcus erythropolis* D-1. Biosci Biotechnol Biochem **59:**1349-1351.
130. **Oldfield, C., N. T. Wood, S. C. Gilbert, F. D. Murray, and F. R. Faure.** 1998. Desulphurisation of benzothiophene and dibenzothiophene by actinomycete organisms belonging to the genus *Rhodococcus*, and related taxa. Antonie Van Leeuwenhoek **74:**119-32.

131. **Panicker, M. M., and E. G. Minkley, Jr.** 1992. Purification and properties of the F sex factor TraD protein, an inner membrane conjugal transfer protein. J Biol Chem **267:**12761-6.
132. **Patrauchan, M. A., C. Florizone, M. Dosanjh, W. W. Mohn, J. Davies, and L. D. Eltis.** 2005. Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. J Bacteriol **187:**4050-63.
133. **Pettis, G. S., and S. N. Cohen.** 1996. Plasmid transfer and expression of the transfer (*tra*) gene product of plasmid pIJ101 are temporally regulated during the *Streptomyces lividans* life cycle. Mol Microbiol **19:**1127-35.
134. **Piddington, C. S., B. R. Kovacevich, and J. Rambosek.** 1995. Sequence and molecular characterization of a DNA region encoding the dibenzothiophene desulfurization operon of *Rhodococcus* sp. strain IGTS8. Appl Environ Microbiol **61:**468-75.
135. **Priefert, H., X. M. O'Brien, P. A. Lessard, A. F. Dexter, E. E. Choi, S. Tomic, G. Nagpal, J. J. Cho, M. Agosto, L. Yang, S. L. Treadway, L. Tamashiro, M. Wallace, and A. J. Sinskey.** 2004. Indene bioconversion by a toluene inducible dioxygenase of *Rhodococcus* sp. I24. Appl Microbiol Biotechnol **65:**168-76.
136. **Prince, R. C., and M. J. Grossman.** 2003. Substrate preferences in biodesulfurization of diesel range fuels by *Rhodococcus* sp. strain ECRD-1. Appl Environ Microbiol **69:**5833-8.
137. **Reygers, U., R. Wessel, H. Muller, and H. Hoffmann-Berling.** 1991. Endonuclease activity of *Escherichia coli* DNA helicase I directed against the transfer origin of the F factor. Embo J **10:**2689-94.
138. **Saeki, H., M. Akira, K. Furuhashi, B. Averhoff, and G. Gottschalk.** 1999. Degradation of trichloroethene by a linear-plasmid-encoded alkene monooxygenase in *Rhodococcus corallinus* (*Nocardia corallina*) B-276. Microbiology **145 (Pt 7):**1721-30.
139. **Sallam, K. I., Y. Mitani, and T. Tamura.** 2006. Construction of random transposition mutagenesis system in *Rhodococcus erythropolis* using IS*1415*. J Biotechnol **121:**13-22.
140. **Samies, J. H., B. N. Hathaway, R. M. Echols, J. M. Veazey, Jr., and V. A. Pilon.** 1986. Lung abscess due to *Corynebacterium equi*. Report of the first case in a patient with acquired immune deficiency syndrome. Am J Med **80:**685-8.
141. **Schindler, J. F., P. A. Naranjo, D. A. Honaberger, C. H. Chang, J. R. Brainard, L. A. Vanderberg, and C. J. Unkefer.** 1999. Haloalkane dehalogenases: steady-state kinetics and halide inhibition. Biochemistry **38:**5772-8.
142. **Schreiner, A., K. Fuchs, F. Lottspeich, H. Poth, and F. Lingens.** 1991. Degradation of 2-methylaniline in *Rhodococcus rhodochrous*: cloning and expression of two clustered catechol 2,3-dioxygenase genes from strain CTM. J Gen Microbiol **137:**2041-8.
143. **Schwarz, F. V., V. Perreten, and M. Teuber.** 2001. Sequence of the 50-kb conjugative multiresistance plasmid pRE25 from *Enterococcus faecalis* RE25. Plasmid **46:**170-87.
144. **Sekine, M., S. Tanikawa, S. Omata, M. Saito, T. Fujisawa, N. Tsukatani, T. Tajima, T. Sekigawa, H. Kosugi, Y. Matsuo, R. Nishiko, K. Imamura, M. Ito,**

H. Narita, S. Tago, N. Fujita, and S. Harayama. 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol **8:**334-46.

145. **Sensfuss, C., M. Reh, and H. G. Schlegel.** 1986. No correlation exists between the conjugative transfer of the autotrophic character and that of plasmids in *Nocardia opaca* strains. J Gen Microbiol **132:**997-1007.
146. **Seto, M., K. Kimbara, M. Shimura, T. Hatta, M. Fukuda, and K. Yano.** 1995. A Novel Transformation of Polychlorinated Biphenyls by *Rhodococcus* sp. Strain RHA1. Appl Environ Microbiol **61:**3353-3358.
147. **Shao, Z., W. A. Dick, and R. M. Behki.** 1995. An improved *Escherichia coli-Rhodococcus* shuttle vector and plasmid transformation in *Rhodococcus* spp. using electroporation. Lett Appl Microbiol **21:**261-6.
148. **Silverman, M. P., M. H. Rogoff, and I. Wender.** 1961. Bacterial Oxidation of Pyritic Materials in Coal. Appl Microbiol **9:**491-6.
149. **Singer, M. E., and W. R. Finnerty.** 1988. Construction of an *Escherichia coli-Rhodococcus* shuttle vector and plasmid transformation in *Rhodococcus* spp. J Bacteriol **170:**638-45.
150. **Stecker, C., A. Johann, C. Herzberg, B. Averhoff, and G. Gottschalk.** 2003. Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. J Bacteriol **185:**5269-74.
151. **Sutcliffe, I. C.** 1998. Cell envelope composition and organisation in the genus *Rhodococcus*. Antonie Van Leeuwenhoek **74:**49-58.
152. **Takai, S., T. Sekizaki, T. Ozawa, T. Sugawara, Y. Watanabe, and S. Tsubaki.** 1991. Association between a large plasmid and 15- to 17-kilodalton antigens in virulent *Rhodococcus equi*. Infect Immun **59:**4056-60.
153. **Takai, S., Y. Watanabe, T. Ikeda, T. Ozawa, S. Matsukura, Y. Tamada, S. Tsubaki, and T. Sekizaki.** 1993. Virulence-associated plasmids in *Rhodococcus equi*. J Clin Microbiol **31:**1726-9.
154. **Tam, A. C., R. M. Behki, and S. U. Khan.** 1987. Isolation and characterization of an s-ethyl-N,N-dipropylthiocarbamate-degrading *Arthrobacter* strain and evidence for plasmid-associated s-ethyl-N,N-dipropylthiocarbamate degradation. Appl Environ Microbiol **53:**1088-93.
155. **Tanaka, Y., T. Matsui, J. Konishi, K. Maruhashi, and R. Kurane.** 2002. Biodesulfurization of benzothiophene and dibenzothiophene by a newly isolated *Rhodococcus* strain. Appl Microbiol Biotechnol **59:**325-8.
156. **Tanaka, Y., O. Yoshikawa, K. Maruhashi, and R. Kurane.** 2002. The *cbs* mutant strain of *Rhodococcus erythropolis* KA2-5-1 expresses high levels of Dsz enzymes in the presence of sulfate. Arch Microbiol **178:**351-7.
157. **Thompson, C. J., J. M. Ward, and D. A. Hopwood.** 1982. Cloning of antibiotic resistance and nutritional genes in streptomycetes. J Bacteriol **151:**668-77.
158. **Thompson, R., L. Taylor, K. Kelly, R. Everett, and N. Willetts.** 1984. The F plasmid origin of transfer: DNA sequence of wild-type and mutant origins and location of origin-specific nicks. Embo J **3:**1175-80.
159. **Tilford, P. E.** 1936. Fasciation of sweet peas caused by *Phytomonas fascians* sp. J Agric Res **53:**383-394.
160. **Tomas-Gallardo, L., I. Canosa, E. Santero, E. Camafeita, E. Calvo, J. A.**

Lopez, and B. Floriano. 2006. Proteomic and transcriptional characterization of aromatic degradation pathways in *Rhodoccocus* sp. strain TFB. Proteomics **6:**S119-S132.

161. **Traxler, B. A., and E. G. Minkley, Jr.** 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. J Mol Biol **204:**205-9.
162. **Treadway, S. L., K. S. Yanagimachi, E. Lankenau, P. A. Lessard, G. Stephanopoulos, and A. J. Sinskey.** 1999. Isolation and characterization of indene bioconversion genes from *Rhodococcus* strain I24. Appl Microbiol Biotechnol **51:**786-93.
163. **Trott, S., S. Burger, C. Calaminus, and A. Stolz.** 2002. Cloning and heterologous expression of an enantioselective amidase from *Rhodococcus erythropolis* strain MP50. Appl Environ Microbiol **68:**3279-86.
164. **Tsai, M. M., Y. H. Fu, and R. C. Deonier.** 1990. Intrinsic bends and integration host factor binding at F plasmid oriT. J Bacteriol **172:**4603-9.
165. **Uhlin, B. E., and K. Nordstrom.** 1985. Preferential inhibition of plasmid replication in vivo by altered DNA gyrase activity in Escherichia coli. J Bacteriol **162:**855-7.
166. **Uhlin, B. E., V. Schweickart, and A. J. Clark.** 1983. New runaway-replication-plasmid cloning vectors and suppression of runaway replication by novobiocin. Gene **22:**255-65.
167. **Voeykova, T., L. Emelyanova, V. Tabakov, and N. Mkrtumyan.** 1998. Transfer of plasmid pTO1 from *Escherichia coli* to various representatives of the order *Actinomycetales* by intergeneric conjugation. FEMS Microbiol Lett **162:**47-52.
168. **Waltermann, M., H. Luftmann, D. Baumeister, R. Kalscheuer, and A. Steinbuchel.** 2000. *Rhodococcus opacus* strain PD630 as a new source of high-value single-cell oil? Isolation and characterization of triacylglycerols and other storage lipids. Microbiology **146 (Pt 5):**1143-9.
169. **Warhurst, A. M., and C. A. Fewson.** 1994. Biotransformations catalyzed by the genus *Rhodococcus*. Crit Rev Biotechnol **14:**29-73.
170. **Warren, R., W. W. Hsiao, H. Kudo, M. Myhre, M. Dosanjh, A. Petrescu, H. Kobayashi, S. Shimizu, K. Miyauchi, E. Masai, G. Yang, J. M. Stott, J. E. Schein, H. Shin, J. Khattra, D. Smailus, Y. S. Butterfield, A. Siddiqui, R. Holt, M. A. Marra, S. J. Jones, W. W. Mohn, F. S. Brinkman, M. Fukuda, J. Davies, and L. D. Eltis.** 2004. Functional characterization of a catabolic plasmid from polychlorinated- biphenyl-degrading *Rhodococcus* sp. strain RHA1. J Bacteriol **186:**7783-95.
171. **Zechner, E. L., F. de la Cruz, R. Eisenbrandt, A. M. Grahn, G. Koraimann, E. Lanka, G. Muth, W. Pansegrau, C. M. Thomas, B. M. Wilkins, and M. Zatyka.** 2000. Conjugative-DNA transfer processes., p. xxiii, 419 p. *In* C. M. Thomas (ed.), The horizontal gene pool: bacterial plasmids and gene spread. Harwood Academic, Amsterdam.
172. **Zopf.** 1891. Uber ausscheidung von fettfarbstoffen seiten gewisse spaltpilze. Berichte der Deutschen Botanischen Gesellschaft **9:**22-28.

CHAPTER 2.

The conjugative megaplasmids of *Rhodococcus erythropolis* AN12.*

* author's note:
portions of this chapter have been submitted
as a manuscript under the title,
"TraA is Required for Megaplasmid Transfer in *Rhodococcus erythropolis* AN12", 2006

2.1. Introduction.

Rhodococcus sp. bacteria are Gram-positive actinomycetes that possess a variety of biochemical and metabolic properties relevant to environmental and industrial microbiology (6, 33). Gene clusters that enable these unique enzymatic pathways often reside on large episomal elements called megaplasmids. *Rhodococcus* megaplasmids exhibit either circular or linear topology, and range from less than 50 kb to greater than one megabasepairs in length (22, 27, 30, 32). Though megaplasmids 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 (1, 21) and other types of Gram-positive bacteria (13, 25).

Both *cis* and *trans*-acting components of plasmid conjugation have been well characterized for many Gram-negative bacteria (18). For most of these plasmids, the DNA processing is initiated in *cis* at the origin of transfer (*ori*T) by a *trans*-acting protein complex called the relaxasome. The core enzyme of this complex is called a relaxase, and it cleaves one strand of the plasmid at the *nic* site within the *ori*T via a transesterification reaction. The nicked plasmid is then unidirectionally transferred to the recipient cell as a singled-stranded DNA intermediate. In plasmid F, the relaxase is encoded by a gene called *traI* (20, 31). 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* (15, 17).

Several different *Rhodococcus* megaplasmids are known to be conjugative (5, 8, 12, 22, 27). The completed sequence of the linear pBD2 megaplasmid from *Rhodococcus erythropolis* BD2 revealed an ORF encoding a putative TraA-like relaxase (30). More recently, sequencing analysis of a circular megaplasmid isolated from *Rhodococcus erythropolis* PR4 also revealed another putative relaxase (26). This evidence suggests that a single-stranded DNA transfer system similar to other bacterial conjugation systems may function in transfer of *Rhodococcus* megaplasmids. However, no functional analysis for any *Rhodococcus* TraA-like relaxase has been reported.

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 (16), 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 megaplasmids, and that at least two of the AN12 megaplasmids (pREA400 and pREA250) can be mobilized to a closely related *R. erythropolis* strain, SQ1. Limited sequence analysis of the AN12 megaplasmids pREA400 and pREA250 revealed ORFs whose gene products are predicted to be involved in plasmid stability, transfer, and gene regulation.

2.2. Materials and methods.

2.2.1. Bacterial strains and culturing conditions. Bacterial strains used in the present study are summarized in Table 2.1. Both strains of *Rhodococcus erythropolis* AN12 and SQ1, as well as strains of *Escherichia coli*, were grown in LB liquid media or on LB plates with 2% agar (24) 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.

TABLE 2.1. Bacterial strains, plasmids, and primers used in this study

Strain, plasmid, or primer name	Description or sequence	Source or reference
DH5α	*Escherichia coli*; *supE*44 Δ*lacU*169 (ϕ80 *lacZ*ΔM15) *recA*1 *endA*1 *hsdR*17 *thi*-1 *gyrA*96 *relA*1	Invitrogen
EC100D *pir*-116	*E. coli*; F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lacZ*ΔM15 Δ*lacX*74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ⁻ *rps*L *nup*G *pir*-116(DHFR)	Epicentre
TOP10	*E coli*; F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *deoR araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^r) *endA*1 *nupG*	Invitrogen
AN12	environmental isolate of *Rhodococcus erythropolis*	(16)
SQ1	environmental isolate of *R. erythropolis*; Rf^R Sm^R	ATCC4277-1
AN12PL	AN12 derivative	This study
AN12PL-1F6	*R. erythropolis* AN12PL derivative recovered from transformation with EZTn transposome; (pREA400::EZTn); Km^R	This study
AN12-5A6	*R. erythropolis* AN12 derivative recovered from transformation with EZTn transposome; (pREA250::EZTn); Km^R	This study
JY524	*R. erythropolis* SQ1 derived transconjugant recovered from mating with AN12-1F6; (pREA400::EZTn); Km^R Rf^R Sm^R	This study
JY640	*R. erythropolis* SQ1 derived transconjugant recovered from mating with AN12-5A6; (pREA250::EZTn); Km^R Rf^R Sm^R	This study

2.2.2. DNA manipulation and plasmid construction. Plasmids used in the present study, as well as a brief description of the cloning strategies used for the construction of each, are summarized below in Table 2.2, except for pEZTn1F6 and pEZTn5A6, which are described here. All DNA modifying enzymes and DNA size ladders were purchased from New England Biolabs (Beverly, MA) and used according to manufacturer's instructions. 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 *Eco*RI for 2 h, followed by incubation at 65°C for 30 min to abolish *Eco*RI activity. The DNA fragments were ligated with T4 DNA ligase, supplemented with 1mM 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.

Table 2.2. Summary of plasmids used in this study

Plasmids	Description	Source or reference
pEZTn1F6	Plasmid containing the EZTn transpososome 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 *mob* PCR product amplified from pSUP301 as a *Mfe*I-*Pst*I fragment, then ligated into pAL298 digested with *Eco*RI and *Pst*I	(19, 28); This study

2.2.3. Preparation and standard transformation of electrocompetent AN12 cells. AN12 electrocompetent cells were prepared essentially as previously described (16), except cells were grown in NBYE/0.05% Tween-80 media in a 1L baffled flask with shaking until $O.D._{600}$ 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 1X TE buffer with 100 to 150 μl competent AN12 cells for 5 minutes prior to electroporation. This suspension was placed in an ice-cold sterile 2mm gapped electroporation cuvette. Electroporation was carried out at 2.5 kV, 25 μF, 400 Ω, with a Bio-Rad Gene Pulser (Hercules, CA). Immediately following pulsing, 200 μl of LB was added to the cells. Cells were allowed to recover after electroporation at 30°C for 2 hours with gentle agitation prior to selection with appropriate antibiotics.

2.2.4. Transposon mutagenesis. Two modifications to the above standard transformation method were made. 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 hours prior to plating on selective media.

2.2.5. *Rhodococcus* conjugation and megaplasmid conjugation frequency. Matings were carried out on solid LB agar surfaces as previously described (19). Plasmid conjugation frequencies were determined by plating appropriate serial dilutions of the mating cell resuspensions (consisting of 1:1 mixtures of donor cells and recipient cells) onto selective media, and plating control recipient cell dilutions to obtain recipient viable counts on non-selective media. Mating efficiencies were calculated as transconjugant CFU per recipient viable cell counts.

2.2.6. 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, 10% Triton-X100). 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, 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 1X TBE, yielding a final cell concentration of 100 μg/μl. 80 μl 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, 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, 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, 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 1X TBE at 14°C, 5V/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.

2.2.7. 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 0.25M HCl for 40 minute, denaturation with 0.5N 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 20X 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 (Roche Diagnostics Corp.). Hybridizing species were detected using Kodak Biomax Light scientific imaging film.

2.2.8. *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 HotStar Taq 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.

2.3. Results.

2.3.1. Discovery of AN12 extrachromosomal elements. It was shown previously that *Rhodococcus erythropolis* AN12 possesses at least one extrachromosomal circular replicon of ca. 6.3 kb, called pAN12 (16). Preliminary experiments in our lab suggested that at least two different species of megaplasmids also exist in the AN12 genome, however the yield and resolution of these replicon were not sufficiently conclusive (Figure 2.1A). Pulsed-field gel (PFG) electrophoresis parameters (culturing conditions, buffer conditions, etc.) were further optimized to examine the megaplasmid species of AN12 and *R. erythropolis* SQ1. One such parameter altered was the percentage of the agarose used to embed the cells. Indeed, by lowering the concentration of the embedding agarose, the resolution of AN12 replicons improved such that two discreet bands in addition to the chromosome can be resolved (Figure 2.1B). Another parameter we focused on improving was the lysis step, during which the embedded cells were exposed to both enzymes (lysozyme and mutanolysin) and detergents (sarcosine). With the unmodified PFG protocol, we noticed that after lysis the agarose plugs remained white, indicating that cells within were largely intact. We reasoned that this may be due

to the detergent inactivation of the enzymes; thus the second improvement consisted of altering the lysis step, such that the enzymatic and detergent treatments were separated. A procedure combining both changes in concentration of the embedding agarose and in the lysis method dramatically improved both the yield and resolution of these large *Rhodococcus* AN12 replicons (Figure 2.1C). Three distinct megaplasmid species in AN12 can now be detected using these PFG conditions (Figure 2.2). These plasmids were named pREA400, pREA250, and pREA100, for *Rhodococcus erythropolis* AN12, based on their approximate sizes in kb. One distinct megaplasmid migrating at 400 kb was observed in SQ1 (Figure 2.2), consistent with what has been previously shown for this strain (22). We propose to call this SQ1 megaplasmid, pRES400.

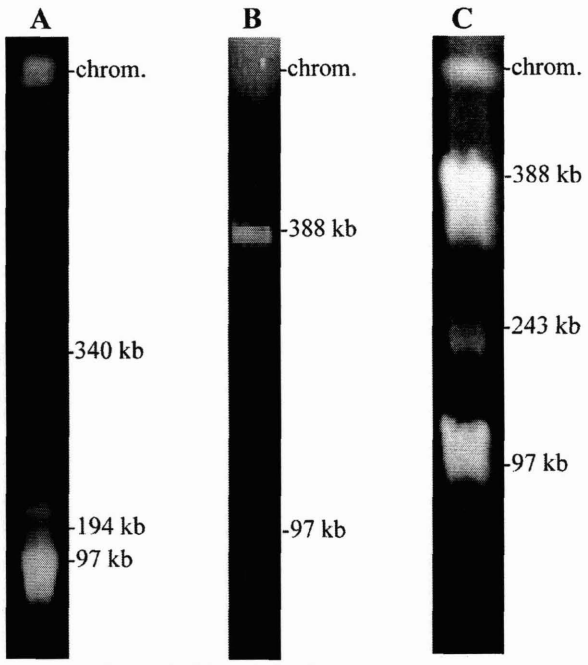


Figure 2.1. Improvements in pulsed-field gel (PFG) electrophoresis. *R. erythropolis* AN12 cells were treated either with (A) 1% embedding agarose, and lysozyme/mutanolysin/sarcosine, (B) 0.5% embedding agarose, and lysozyme/mutanolysin/sarcosine, or (C) 0.5% embedding agarose and separated lysozyme/mutanolysin and sarcosine. Replicons were then resolved using PFG using conditions stated in the materials and methods section. Sizes of the relevant lambda PFG ladder (NEB), as well as chromosomal band in each lane are indicated.

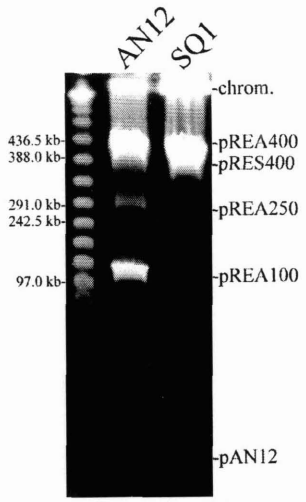


Figure 2.2. Pulsed-field gel (PFG) profiles of wild type AN12 and SQ1 replicons. Approximately 2 μg of the lambda ladder PFG marker were loaded to determine the migrating size of the megaplasmids. Chromosomal (chrom.) and distinguishable species of plasmids, as well as relevant ladder sizes (in kb) are labeled.

2.3.2. AN12 transposon library construction and genetic screen for transmissible elements. 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 (5, 27). We sought to determine whether, like pBD2 and pRHL2, the AN12 megaplasmids are conjugative. Earlier experiments in *Rhodococcus equi* (2) and *Rhodococcus rhodochrous* CW25 (10) demonstrated that transposons can be used effectively to generate random mutations *in vivo* . However, we sought to generate an AN12 transposon library not solely to induce mutations, but also to screen transposon associated sequences for transmissibility. To do so, a commercially available transposon-mediated system (EZ::TN <R6Kγ*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-wildtype (WT) and an AN12 derived strain, AN12PL, which lacks pREA100 and possibly pREA250, were used in the generation of this library. The AN12 WT strain exhibits glossy domed colony morphology and its cells in liquid cultures are evenly suspended, whereas the AN12PL strain exhibits dull flat

colony morphology and its cells in liquid cultures are flocculent. The majority of the AN12 colonies exhibit the wild-type glossy morphology, although the glossy appearance tends to diminish as the colonies age. The transposon cocktail was electroporated into both AN12 WT and AN12PL competent cells, and subsequent transposon establishment in the genome was selected based on kanamycin resistance (Km^R). In this way, a library consisting of a total of 432 mutants was created. Genomic DNA was prepared from a subset of mutants and screened using Southern blotting techniques to ensure that transposon insertion sites were relatively random (N. Sengupta and S. Windsor, data not shown). A subset of 119 mutants were 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) (19). SQ1-derived transconjugants can be selected on the basis of triple antibiotic resistance (Km^R Rf^R Sm^R).

The matings were conducted by pooling 4 or 5 individual donor AN12 or AN12PL mutant cultures and incubating with an equivalent volume of SQ1 recipient culture, first on LB agar plates for 24 hours, 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. Individual matings of all mutants in these two pools revealed that a single mutant donor in each case was responsible. These mutants were named AN12PL-1F6 and AN12-5A6. Conjugal transfer of the tagged AN12PL-1F6 element was determined to occur at a frequency of 7.1×10^{-4} (standard deviation of 9.9×10^{-5}, n = 5) per recipient SQ1 cell, and the tagged AN12-5A6 element at 5.2×10^{-4} (standard deviation of 1.3×10^{-4}, n = 4) per recipient cell, comparable to the frequencies determined for pBD2 and pRHL2 (5, 27).

2.3.3. 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. *Rhodococcus* DNA adjacent to each transposon insertion site was rescued from AN12PL-1F6 and AN12-5A6 in the form of plasmids that we named pEZTn1F6 and pEZTn5A6 (Fig. 2.3), respectively. Sequencing and restriction enzyme analysis showed that pEZTn1F6 and pEZTn5A6 are approximately 11 kb and 16 kb, respectively. To

demonstrate that DNA had indeed transferred from AN12 to SQ1, these plasmids were used as templates to synthesize probes that were hybridized to fragmented genomic DNA isolated from each donor, recipient and transconjugant (Figure 2.3). 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.

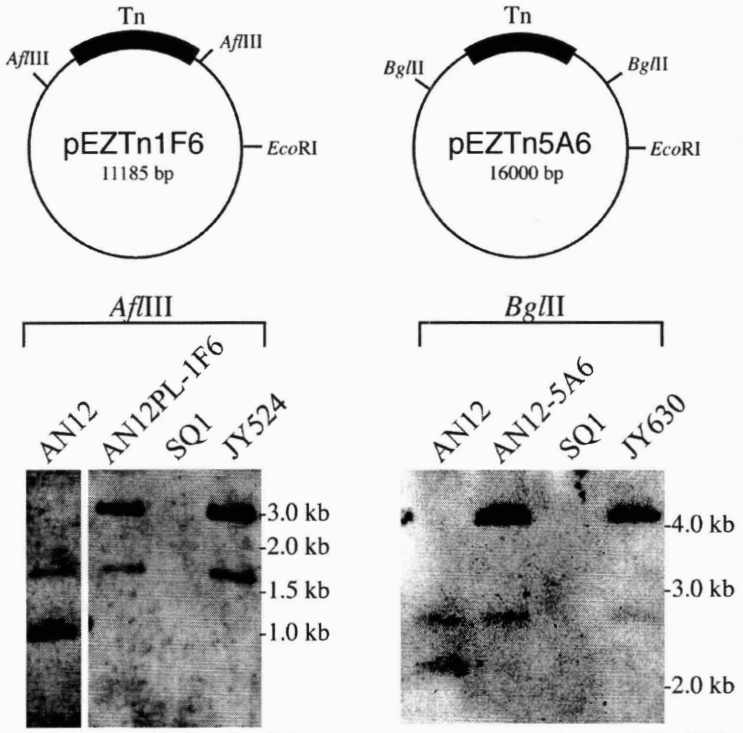


Figure 2.3. Southern blot analysis of wild type (AN12), donor (AN12PL-1F6 or AN12-5A6), recipient (SQ1), and transconjugant (JY524 or JY630) DNA. Genomic DNA (gDNA) was isolated from the indicated strains and treated with either *Afl*III or *Bgl*II enzymes to distinguish between wild type AN12 and AN12 strains which bear transposon insertions (AN12PL-1F6 and AN12-5A6). These DNA species were then hybridized to digoxygenin (DIG) probes generated with the plasmid rescue products, pEZTn1F6 and pEZTn5A6. Relevant 1kb DNA ladder sizes are indicated to the right. Expected hybridizing product size in each lane are as follows: for *Afl*III digests, AN12 (WT)- 1,017 bp, AN12PL-1F6 and JY524- 3,018 bp; for *Bgl*II digests, AN12 (WT)- 2,593 bp, and 2,155 bp, AN12-5A6 and JY630- 2,593 bp and 4,156 bp. No hybridizing DNA species were expected for in SQ1 gDNA treated with either enzyme.

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. 2.4A, the transconjugant from the AN12-5A6 x SQ1

mating clearly gained the pREA250 megaplasmid. PFG analysis showed that other transconjugants isolated from this mating also gained the pREA250 element (data not shown). This strongly suggests 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 if 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 (Figure 2.4B). 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.

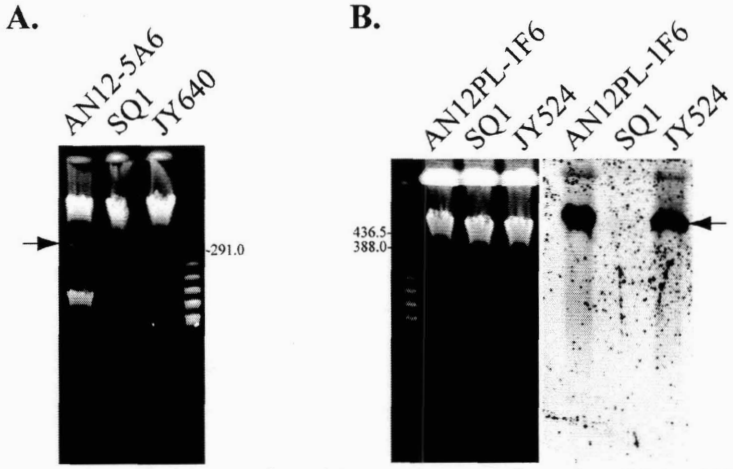


Figure 2.4. Transposons-tagged megaplasmids can be transferred to SQ1. (A) Pulsed-field gel (PFG) profiles of donor strain, AN12-5A6, recipient strain, SQ1, and SQ1-derived transconjugant, JY640. The AN12 replicon, pREA250, that has been mobilized to SQ1 is labeled with an arrow. (C) PFG (left) and subsequent 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 kb.

2.3.4. Sequence analysis of the AN12PL-1F6 transposon insertion region.

Plasmid "addiction" systems ensure the stable inheritance of a plasmid through rounds of cell division by killing off cells that have lost the plasmid (9). 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 for 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 (30).

Sequence analysis of pEZTn1F6 revealed three open reading frames (ORFs), *traA, pemI,* and *pemK*, at or adjacent to the transposon insertion in pREA400 (Figure 2.5). The rest of this chapter includes only analysis of the latter two ORF, as the functional analysis of the pREA400 encoded *traA* will be discussed in the next chapter of this thesis. Upon closer examination, the transposon inserted precisely one basepair downstream of the start codon of a *pemK*-like toxin gene, likely disrupting its expression. Like the operon organization of other toxin genes involved in post-segregational killing, the pREA400 *pemK* ORF (318 bp) is preceded by a 312 bp ORF, *pemI*, likely encoding antitoxin activity. PFAM analysis of pREA400 PemK readily identifies similar proteins (Figure 2.6). The *Escherichia coli* R100 plasmid encoded PemK (accession no. P13976) and a predicted protein from *Mycobacterium bovis* (accession no. P64860) are 13% and 50% identical to pREA400 PemK, respectively. However, neither PFAM nor BLASTP analysis against the non-redundant protein database using translated pREA400 *pemI* could identify comparably similar proteins. This was not unexpected as genes encoding antidotes are generally identified only by proximities to the toxin genes since antitoxin genes show little sequence similarity to each other.

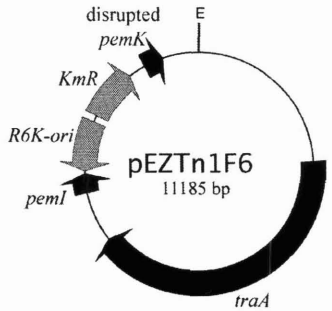


Figure 2.5. Detailed schematic diagram of the plasmid rescue product from recircularized genomic DNA isolated from AN12PL-1F6. Gray or black coloring is used to differentiate between ORFs and functions encoded by the transposon (*R6K-ori* and *KmR*) from *Rhodococcus* pREA400 (*traA*, *pemI,* and *pemK*), respectively. The *Eco*RI site (E) used to recircularize this plasmid is indicated.

```
R.eryth_PemK    1   --------------------MTAALRGQIYWADIG-------KGEKPWVVVSNNVRNRNL 33
M.bovis_P64860  1   --------------------MNAPLRGQVYRCDLG-------YGAKPWLIVSNNARNRHT 33
E.coli_PemK     1   MLKYQLKNENGWMHRRLVRRKSDMERGEIWLVSLDPTAGHEQQGTRPVLIVTPAAFNRVY 60

R.erth_PemK     34  NTVLAARVTTTPKTGVPTAVPLGAADP---LVGSILADDLIQLFDDEIAASRPAGALSPA 90
M.bovis_P64860  34  ADVVAVRLTTTRRT-IPTWVAMGPSDP---LTGYVNADNIETLGKDELGDY--LGEVTPA 87
E.coli_PemK     61  RLPVVVPVTSGGNFARTAGFAVSLDGVGIRTTGVVRCDQPRTIDMKARGGKR-LERVPET 119

R.eryth_PemK    91  TVVKLNKALAIALGLP-- 106
M.bovis_P64860  88  TMNKINTALATALGLPWP 105
E.coli PemK     120 IMNEVLGRLSTILT---- 133
```

Figure 2.6. Sequence alignment of PemK toxin proteins. The intact *Rhodococcus* pREA400 *pemK* gene product sequence was used to identify similar proteins by PFAM and BLASTP analyses. One hypothetical protein from *Mycobacterium bovis* (accession no. P64860), and PemK encoded by plasmid R100 in *Escherichia coli* (accession no, P13976) were identified. A ClustalW alignment was then performed. Identical residues in at least two of the three aligning sequences are indicated with gray boxes.

Anticipating plasmid instability due to the transposon disruption of the *pemK* ORF, we looked for the curing of the transposon tagged pREA400 megaplasmid in AN12PL-1F6. However, our preliminary evidence suggests that even without antibiotic selection for 10 consecutive days, the AN12PL-1F6 subcultures retained the ability to subsequently thrive on LB media supplemented with kanamycin, suggesting that the tagged megaplasmid is inherited stably. We do not interpret this result as evidence against the PemK-like toxin's potential function in pREA400 plasmid stability, but rather

we propose that other selective pressures, such as essential genes, prevent the loss of this plasmid.

2.3.5. Sequence analysis of the AN12PL-1F6 transposon insertion region. The sequence analysis of the pREA250 region flanking the transposon insertion did not immediately reveal genes involved in plasmid function. Using BLASTX and subsequent BLASTP analysis, 6 putative ORFs were defined, as well as two perfect inverted repeat elements that is likely to form a stem-loop/hairpin structure. Though most of the ORFs are most similar to conserved hypothetical ORFs in other actinomycetes or mycolata, the ORF2 likely encodes a TetR type repressor, and ORF6 likely encodes a outer membrane protein.

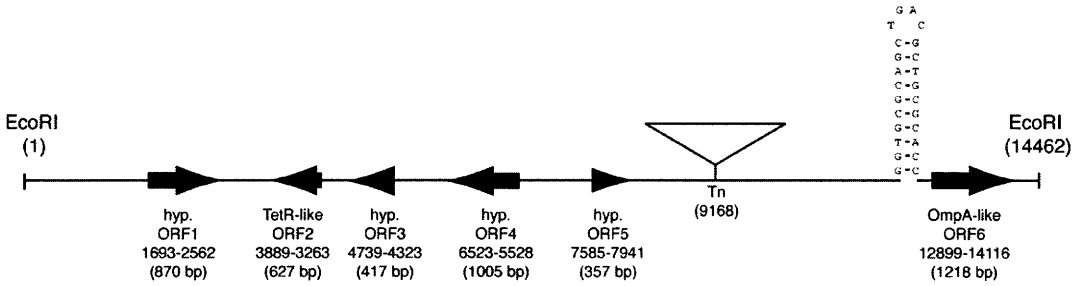


Figure 2.7. The megaplasmid pREA250 transposon insertion region is shown as a linearized 14,462 bp fragment from the *Eco*RI sites used to generate pEZTn5A6. The transposon is represented as a inverted triangle. A putative hairpin structure with 10 bp perfect inverted repeats and a loop consisting of 4 bp is drawn (not to scale).

2.4. Discussion.

Common modes by which bacteria exchange genetic information include phage-mediated transduction, transposition, and plasmid conjugation. Exploitation and further understanding of the last 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 these 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. Many plasmid-encoded mechanisms, including the regulation of its replication, partition, and addiction (post-segregational killing), are known to affect plasmid copy number in bacterial systems (4, 7). Should we find that AN12 megaplasmids pREA400 and pREA250 are maintained at different copy numbers, if would be of interest to determine which of these mechanisms are encoded by these *Rhodococcus* megaplasmids, and whether any of these mechanisms contribute to the observed difference in plasmid copy number.

By screening an AN12 transposon library, it was determined that two of the three AN12 megaplasmids are conjugative to and maintained by *R. erythropolis* SQ1. Consistent with this finding, a putative *traA* open reading frame encoding a conjugation initiating relaxase was found on the AN12 megaplasmid, pREA400. The functional analysis and significance of pREA400 *traA* will be the topic of discussion in the next chapter. Mobilization frequencies of pREA400 and pREA250 were determined to be more than 1 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, since pREA400 and pREA250 exhibit essentially the same mobilization frequency.

On pulsed-field gels, circular plasmids are known to change in apparent mobilities (with respect to the reference linear DNA marker) upon changing electrophoresis conditions (29). We suspect that *R. erythropolis* AN12 pREA400 and pREA100 are linear plasmids since changing PFG electrophoresis parameters do not appreciably affect their apparent migrating sizes (compare Figure 2.1C to Figure 2.2). Likewise, we propose that pREA250 is a circular plasmid, since it migrates as a 243 kb fragment in Figure 2.1C and as a 291 kb fragment in Figure 2.2.

Linear plasmids are not uncommon in the actinomycetes. In fact, the first linear plasmid discovered was isolated from an antibiotic producing strain of *Streptomyces* in 1979 (11). The ends of linear plasmids fall into two groups, 1) the covalently closed hairpin end types found in the spirochete *Borrelia* (23), and 2) the 5' protein-associated, invertron-containing linear plasmids. The nature of the latter class of linear plasmids has been used to distinguish megaplasmid topology. Since protein-associated (non-denaturing) plasmids do not migrate out of pulsed-field gel plugs, authors have argued that the absence of the megaplasmid band on PFG when plugs were not treated with proteinase K is evidence for the linearity of plasmid (14).

A conceptually simpler and unambiguous way to assess *Rhodococcus* megaplasmid topology has been proposed by J. Taylor and J. Archer (personal communication) in which the recognition sequence of a rare-cutting restriction enzyme, such as *Swa*I could be integrated into the plasmid of interest. The undigested and digested plasmids could then be resolved on a PFG. Should the single megaplasmid band become two species, then one could conclude that the plasmid in question is linear. Should the plasmid be circular, it should remain as one band, though the migration size will likely be different. AT-rich restriction enzyme recognition sites, such as *Swa*I (ATTTAAAT) and *Pac*I (TTAATTAA), have not yet been found in the incomplete genome sequence of *Rhodococcus aetherivorans* I24 (J. Archer, personal communication), and thus are not expected to be present on the megaplasmids examined nor complicate topological analyses. This strategy could be used to ascertain the topology of AN12 megaplasmids, pREA400, pREA250 and pREA100 in the near future.

PFG analysis shows that AN12PL-1F6 has been cured of megaplasmid pREA100, and possibly pREA250. One trivial explanation for this observation might have been that the AN12PL-1F6 mutant is an SQ1 contaminant, since SQ1 carries a megaplasmid of approximately 400 kb. However, four lines of evidence 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, pEZTn1F6, 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 megaplasmids 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 megaplasmids. 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 megaplasmids 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 (3). It would be very interesting to determine whether this plasmid fluidity is the case as well for rhodococci.

2.5. References.

1. **Argandona, M., F. Martinez-Checa, I. Llamas, E. Quesada, and A. del Moral.** 2003. Megaplasmids in Gram-negative, moderately halophilic bacteria. FEMS Microbiol Lett **227:**81-6.
2. **Ashour, J., and M. K. Hondalus.** 2003. Phenotypic mutants of the intracellular actinomycete *Rhodococcus equi* created by in vivo *Himar1* transposon mutagenesis. J Bacteriol **185:**2644-52.
3. **Bey, S. J., M. F. Tsou, C. H. Huang, C. C. Yang, and C. W. Chen.** 2000. The homologous terminal sequence of the *Streptomyces lividans* chromosome and SLP2 plasmid. Microbiology **146 (Pt 4):**911-22.
4. **Chattoraj, D. K.** 2000. Control of plasmid DNA replication by iterons: no longer paradoxical. Mol Microbiol **37:**467-76.
5. **Dabrock, B., M. Kesseler, B. Averhoff, and G. Gottschalk.** 1994. Identification and characterization of a transmissible linear plasmid from *Rhodococcus erythropolis* BD2 that encodes isopropylbenzene and trichloroethene catabolism. Appl Environ Microbiol **60:**853-60.
6. **de Carvalho, C. C., and M. M. da Fonseca.** 2005. The remarkable *Rhodococcus erythropolis*. Appl Microbiol Biotechnol **67:**715-26.
7. **del Solar, G., and M. Espinosa.** 2000. Plasmid copy number control: an ever-growing story. Mol Microbiol **37:**492-500.

8. **Desomer, J., P. Dhaese, and M. Van Montagu.** 1988. Conjugative transfer of cadmium resistance plasmids in *Rhodococcus fascians* strains. J Bacteriol **170:**2401-5.
9. **Engelberg-Kulka, H., and G. Glaser.** 1999. Addiction modules and programmed cell death and antideath in bacterial cultures. Annu Rev Microbiol **53:**43-70.
10. **Fernandes, P. J., J. A. Powell, and J. A. Archer.** 2001. Construction of *Rhodococcus* random mutagenesis libraries using Tn*5* transposition complexes. Microbiology **147:**2529-36.
11. **Hayakawa, T., T. Tanaka, K. Sakaguchi, N. Otake, and H. Yonehara.** 1979. A linear plasmid-like DNA in *Streptomyces* sp. producing lankacidin group antibiotic. J Gen Appl Microbiol **25:**255-260.
12. **Kalkus, J., C. Dorrie, D. Fischer, M. Reh, and H. G. Schlegel.** 1993. The giant linear plasmid pHG207 from *Rhodococcus* sp. encoding hydrogen autotrophy: characterization of the plasmid and its termini. J Gen Microbiol **139:**2055-65.
13. **Kinashi, H., M. Shimaji, and A. Sakai.** 1987. Giant linear plasmids in *Streptomyces* which code for antibiotic biosynthesis genes. Nature **328:**454-6.
14. **Konig, C., D. Eulberg, J. Groning, S. Lakner, V. Seibert, S. R. Kaschabek, and M. Schlomann.** 2004. A linear megaplasmid, p1CP, carrying the genes for chlorocatechol catabolism of *Rhodococcus opacus* 1CP. Microbiology **150:**3075-87.
15. **Kopec, J., A. Bergmann, G. Fritz, E. Grohmann, and W. Keller.** 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:**401-9.
16. **Kostichka, K., L. Tao, M. Bramucci, J. F. Tomb, V. Nagarajan, and Q. Cheng.** 2003. A small cryptic plasmid from *Rhodococcus erythropolis*: characterization and utility for gene expression. Appl Microbiol Biotechnol **62:**61-8.
17. **Kurenbach, B., D. Grothe, M. E. Farias, U. Szewzyk, and E. Grohmann.** 2002. The *tra* region of the conjugative plasmid pIP501 is organized in an operon with the first gene encoding the relaxase. J Bacteriol **184:**1801-5.
18. **Lanka, E., and B. M. Wilkins.** 1995. DNA processing reactions in bacterial conjugation. Annu Rev Biochem **64:**141-69.
19. **Lessard, P. A., X. M. O'Brien, D. H. Currie, and A. J. Sinskey.** 2004. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. BMC Microbiol **4:**15.
20. **Matson, S. W., and H. Ragonese.** 2005. The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. J Bacteriol **187:**697-706.
21. **Pedraza, R. O., and J. C. Diaz Ricci.** 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:**130-3.
22. **Priefert, H., X. M. O'Brien, P. A. Lessard, A. F. Dexter, E. E. Choi, S. Tomic, G. Nagpal, J. J. Cho, M. Agosto, L. Yang, S. L. Treadway, L. Tamashiro, M. Wallace, and A. J. Sinskey.** 2004. Indene bioconversion by a toluene inducible dioxygenase of *Rhodococcus* sp. I24. Appl Microbiol Biotechnol **65:**168-76.

23. **Saint Girons, I., I. G. Old, and B. E. Davidson.** 1994. Molecular biology of the *Borrelia*, bacteria with linear replicons. Microbiology **140 (Pt 8):**1803-16.
24. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
25. **Scholle, M. D., C. A. White, M. Kunnimalaiyaan, and P. S. Vary.** 2003. Sequencing and characterization of pBM400 from *Bacillus megaterium* QM B1551. Appl Environ Microbiol **69:**6888-98.
26. **Sekine, M., S. Tanikawa, S. Omata, M. Saito, T. Fujisawa, N. Tsukatani, T. Tajima, T. Sekigawa, H. Kosugi, Y. Matsuo, R. Nishiko, K. Imamura, M. Ito, H. Narita, S. Tago, N. Fujita, and S. Harayama.** 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol **8:**334-46.
27. **Shimizu, S., H. Kobayashi, E. Masai, and M. Fukuda.** 2001. Characterization of the 450-kb linear plasmid in a polychlorinated biphenyl degrader, *Rhodococcus* sp. strain RHA1. Appl Environ Microbiol **67:**2021-8.
28. **Simon, R., U. Priefer, and A. Puehler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagensis in gram negative bacteria. Bio/Technology **1:**784-791.
29. **Simske, J. S., and S. Scherer.** 1989. Pulsed-field gel electrophoresis of circular DNA. Nucleic Acids Res **17:**4359-65.
30. **Stecker, C., A. Johann, C. Herzberg, B. Averhoff, and G. Gottschalk.** 2003. Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. J Bacteriol **185:**5269-74.
31. **Traxler, B. A., and E. G. Minkley, Jr.** 1988. Evidence that DNA helicase I and *oriT* site-specific nicking are both functions of the F TraI protein. J Mol Biol **204:**205-9.
32. **Trott, S., S. Burger, C. Calaminus, and A. Stolz.** 2002. Cloning and heterologous expression of an enantioselective amidase from *Rhodococcus erythropolis* strain MP50. Appl Environ Microbiol **68:**3279-86.
33. **van der Geize, R., and L. Dijkhuizen.** 2004. Harnessing the catabolic diversity of rhodococci for environmental and biotechnological applications. Curr Opin Microbiol **7:**255-61.

CHAPTER 3.

Development of a targeted gene disruption strategy for genetic and molecular analysis of *traA* required for plasmid conjugation in *Rhodococcus erythropolis*, and an analogous method utilized in the fruit fly, *Drosophila melanogaster*

* author's note:
portions of this chapter have been submitted
as a manuscript under the title,
"TraA is Required for Megaplasmid Transfer in *Rhodococcus erythropolis* AN12", 2006

3.1. Introduction.

As of this writing, the gene sequences and annotations of 361 genomes, including well characterized genetic model organisms such as *Escherichia coli* (8), *Saccharomyces cerevisiae* (23), *Caenorhabditis elegans* (1), *Drosophila melanogaster* (3, 45), and *Mus musculus* (73), are available. The coupling of genomic sequence information with advances in gene-specific disruption strategies- including RNA interference (RNAi) for *C. elegans* (20, 25) and *D. melanogaster* (34), as well as targeted gene replacement, disruption, and deletion for yeast (56) to mice (67)- has revolutionized genetic research. Specifically, the reverse genetics (genotype to phenotype) approach to gene function analysis not only complements, but now arguably rivals the more classical forward genetics (phenotype to genotype) approach in utility and scale. Indeed, with the phenotypic analysis of 2,026 systematically generated gene-specific deletions in *Saccharomyces cerevisiae*, three years after the completion of the yeast genome, functional genomics was born (74).

The lack of classical genetic tools for rhodococci, particularly bacteriophages and efficient transformation methods, has greatly hampered their biological characterization in the past. Only one transducing bacteriophage, Q4, for *Rhodococcus erythropolis* has ever been described and is not widely used (15). The transformation efficiency of most strains of rhodococci is mediocre at best, with one microgram of DNA typically yielding less than 1,000 CFUs upon electroporation without optimization. Searches for other means of rhodococci transformation has led to the discovery of *E. coli-Rhodococcus* interspecies conjugation (60, 71). *Rhodococcus* transformations using these methods are often more efficient (10,000 exconjugants) than electroporation, but the rates are still not comparable to the competence of *E. coli*. Thus, while chemical mutagenesis using either ethyl methanesulfonate (EMS) or *N*-methyl-*N'*-nitro-*N*-nitrosoguanidine (NMG) has been carried out successfully in rhodococci (26, 72), the subsequent correlation of the mutant phenotype to a specific genetic locus or pathway is difficult.

To overcome these difficulties, four methods of targeted gene disruption have been developed to facilitate *Rhodococcus* genetic analysis. Both the λ-red method adapted for the analysis of benzoate and phthalate metabolism in *Rhodococcus* sp. RHA1 (48), and the Ω element method adapted for the analysis of vanillin production in

Rhodococcus opacus PD360 (49) rely on infrequent double crossover event and two rounds of drug resistance/sensitivity screening. A method described for generating gene knockouts in *R. erythropolis* SQ1 which takes advantage of single cross-over recombination and the *sacB* counter-selection, also relies on two rounds of drug resistance/sensitivity screening (69). Such elaborate methods have been developed for different strains of rhodococci because homologous recombination is rare, and illegitimate recombination is frequent (17). However, one recent study of carotenoid production in *R. erythropolis* ATCC47072 successfully employed a simple single cross-over recombination strategy with a targeting cassette of 800 bp homologous sequence to generate a gene knockout (66).

The serendipitous discovery of the putative *traA* ORF proximal to the transposon insertion in *Rhodococcus* AN12PL-1F6 begs the question of whether its gene product might be involved in pREA400 megaplasmid conjugation. We chose to use a reverse genetic approach, by using a novel targeted gene disruption method, to directly address *traA* gene function. Our method differs from the four previously reported methods in that the homologous recombination is stimulated by a doubled-stranded DNA break.

A landmark study in *Saccharomyces cerevisiae* directly examined the recombinogenic potential of DNA lesions *in vivo* (46). Nonreplicating plasmids which bear homologous sequences to various *S. cerevisiae* genetic loci (*LEU2* and *HIS3*) were introduced either as intact plasmids or prelinearized plasmids, with the restriction enzyme recognition/DNA lesion site either within the homology cassettes or elsewhere on the plasmid. The strain into which these DNA species were introduced was deficient for amino acid production; thus genetic transformation could be detected as survival on media without either leucine or histidine. While circular plasmids, and non-site-specifically linearized DNA exhibited low transformation efficiency, linearized plasmid containing DNA lesions within the homologous regions showed up to 2,000 fold increase in the transformation efficiency. This study also showed that the high frequency integration event was dependent on *RAD52*, whose gene product was known to be involved in doubled-stranded DNA repair and gene conversion (32, 50).

Not long after, elements (double-stranded break and homology-driven recombination) from this yeast study were adapted to generate chromosomal gene

disruptions of the *HPRT* locus in mouse embryonic stem (ES) cells (67), at once demonstrating the mutagenic potential of this strategy and its application in other eukaryotic systems. Use of a rare-cutting endonuclease called I-*Sce*I and its 18 bp recognition sequence (5' TAGGGATAACAGGGTAAT 3') to generate double-stranded breaks (DSB) in COS-1 and ES cells further improved the recombination frequency (12, 18, 52, 53, 61), and directly led to the development of a site-specific gene targeting method in *Drosophila melanogaster* (51).

In this chapter, I will first discuss using the gene targeting method of Rong and Golic (2000) to characterize *ptpmeg1* gene inactivation in *Drosophila melanogaster*, a project I pursued earlier in my graduate career. Next I will discuss the adaptation of DSB induced homologous recombination to study the phenotypic consequence of the *traA* genetic disruption in *Rhodococcus*.

3.2. Materials and methods.

3.2.1. Fly stocks, bacterial strains and culturing conditions. Fly stocks pertinent to *ptpmeg1* gene targeting were obtained from K. Golic. Other fly stocks important for ectopic expression UAS-*Ptpmeg1*^{WT} or UAS-*ptpmeg1*$^{\Delta FERM}$, such as *daughterless*-Gal4 (75) and *elav*-Gal4 (43) were obtained from the Bloomington Drosophila Stock Center. Bacterial strains used in the present study are summarized below in Table 3.1. Both strains of *Rhodococcus erythropolis* AN12 and SQ1, as well as strains of *Escherichia coli*, were grown in LB liquid media or on LB plates with 2% agar (55) supplemented with the following antibiotics purchased from Sigma-Aldrich (St. Louis, MO) as appropriate; ampicillin (Ap, 50 μg/ml), chloramphenicol (Cm, 12.5 μg/ml), 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.

Table 3.1. Bacterial strains used in work described in this chapter.

Strain names	Description	Source or reference
Drosophila[a]		
TOP10 F'	*Escherichia coli* strain used to propagate *Drosophila* cDNA library clones	Invitrogen
DH10B	*E. coli* strain used to propagate *Drosophila* BAC library clones	Invitrogen
Rhodococcus[a, b]		
JY825	*R. erythropolis* AN12 transformed with pJY37 to disrupt the pREA400 megaplasmid encoded *traA*; 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	*R. erythropolis* SQ1 derived transconjugant recovered from mating with JY926; Km^R Gm^R Rf^R Sm^R	This study

a. Work pertaining to either the *ptpmeg1* (*Drosophila*) knockout or the *traA* (*Rhodococcus*) has been indicated.
b. see Table 2.1 for descriptions of parental strains of wild type *R. erythropolis* AN12 and *R. erythropolis* SQ1, as well as *E. coli* DH5α.

3.2.2. DNA manipulation and plasmid construction. Plasmids used in the present study, as well as a brief description of the cloning strategies used for the construction of each, are summarized in Table 3.2, except for pEZTn1F6 and pEZTn5A6, which were described previously in Chapter 2, section 2. All DNA modifying enzymes and DNA size ladders were purchased from New England Biolabs (Beverly, MA), unless otherwise noted, and used according to manufacturer's instructions. All PCR primers used in this study were purchased from Integrated DNA Technologies (Coralville, IA) and are summarized below in Table 3.3. *Drosophila* genomic DNA was prepared as previously described (58). *Rhodococcus* genomic DNA was prepared as previously described (42). *Rhodococcus* plasmid DNA was isolated using the procedures previously described (38). A *Drosophila* BAC clone spanning the *Ptpmeg* region was obtained from I. Rebay. *Drosophila* cDNA clones carrying various splice isoforms of *Ptpmeg* were obtained from Research Genetics (Huntville, AL). All DNA sequencing reactions were performed by the MIT Biopolymers Laboratory.

Table 3.2. Summary of plasmids and cloning strategies used in this chapter.

Plasmids	Description	Source or reference
Drosophila[a]		
BACR13F06/ RP98-13F6	pBACe3.6 clone with a 167,968 bp insert corresponding to scaffold AE003458 and 61C1 region of the 3rd chromosome; Cm^R	(31)

continued on next page

LD22982	*Ptpmeg* full-length cDNA cloned into pOT2a vector; Cm^R	(4)
LD27491	$Ptpmeg^{\Delta FERM}$ cDNA cloned into the pOT2a vector; Cm^R	(4)
LD13416	*Ptpmeg* incomplete FERM domain cDNA cloned into the pOT2a vector; Cm^R	(4)
pUASt	P-element vector for ectopic gene expression; *white* (eye color) selectable marker; Ap^R	(9)
pTV2	P-element based gene targeting vector; 2 FRT sites, *white* selectable marker; Ap^R	(51)
pBS-MegBAC1	pBluescript(+) vector (Stratagene) containing a 2,892 bp *Eco*RI and *Spe*I (corresponding from nt. 114,730 to 117,626) fragment isolated from BACR13F06	This work
pBS-MegBAC2	pBS-MegBAC1 modified such that the unique *Acc*65I site (nt. 116,500 of BACR13F06) within the *Ptpmeg* cassette was abolished by Klenow following digestion	This work
pBS-MegISceI	pBS-MegBAC2 modified such that annealed oligonucleotides, JCY025/JCY026, containing the I-*Sce*I recognition site was subcloned into the unique *Bst*EII site (nt. 116,097 of BACR13F06) within the *Ptpmeg* cassette	This work
pTV2-MegISceI	pTV2 carrying a ca. 3.0kb *Not*I-*Kpn*I fragment subcloned from pBS-MegISceI corresponding to the modified *Ptpmeg* cassette	This work
pUASt-LD27491	A 2.3 kb *Eco*RI and *Bgl*II fragment from LD27491 (pOT2a) clone was ligated to pUASt; a *Kpn*I and *Hind*III from LD27491 (1.4 kb) was ligated to pBluescript II KS(+); both pUASt and pBluescript II KS(+) resulting intermediate constructs were then digested with *Xba*I and *Kpn*I, and the smaller digestion product from intermediate pBluescript II KS(+) was then ligated to the larger digested product of intermediate pUASt construct to yield the final construct	This work
pUASt-LD22982	A 2,169 bp *Eco*RI and *Acc65*I fragment from LD22982 (pOT2a), a 3,859 bp *Nco*I and *Acc*65I fragment from pUASt-LD27491, and a 6.589 bp *Nco*I and *Eco*RI fragment from pUASt-LD27491 were ligated together in a 3-way ligation to yield the final construct	This work
Rhodococcus[a]		
pEZTn1F6	Plasmid containing the EZTn transpososome and flanking AN12 genomic sequences isolated from AN12-1F6; Km^R	Ch. 2, (76)
pEZTn5A6	Plasmid containing the EZTn transposome and flanking AN12 genomic sequences isolated from AN12-5A6; Km^R	Ch. 2, (76)
pREA400	endogenous AN12 megaplasmid	Ch. 2, (76)
pREA250	endogenous AN12 megaplasmid	Ch. 2, (76)
pREA100	endogenous AN12 megaplasmid	Cha. 2, (76)
pAL349	RP4 *mob* PCR product amplified from pSUP301 as a *Mfe*I-*Pst*I fragment, then ligated into pAL298 digested with *Eco*RI and *Pst*I	(42, 60); This study
pJY37	plasmid used for *traA* targeted disruption; primers JYP560 and JYP561 were used to amplify a 706 bp fragment from AN12 genomic DNA corresponding to the interior 5' region of the *traA* gene, which was then cloned into pCR2.1Topo	This study
pJY48	*Fsp*I deletion of pAL349	This study
pJY49B	4900 bp *Xmn*I fragment isolated rom pEZTn1F6 containing the entire *traA* ORF cloned into pAL349 digested with *Fsp*I	This study

a. Work pertaining to either the *ptpmeg* (*Drosophila*) knockout or the *traA* (*Rhodococcus*) has been indicated.

Table 3.3. Summary of primers used in this chapter.

Primers	Sequence	Reference
Drosophila[a]		
JCY025	forward primer, when annealed to JCY026 and ligated to pBS-MegBAC2, generates an I-*Sce*I site in place of *Ptpmeg* endogenous *Bst*EII site; 5' GTCACTAGGGATAACAGGGTAAT 3'	This work
JCY026	reverse primer, when annealed to JCY026 and ligated to pBS-MegBAC2, generates an I-*Sce*I site in place of *Ptpmeg* endogenous *Bst*EII site; 5' GTGACATTACCCTGTTATCCCTA 3'	This work
MEGKO DIAG1	*ptpmeg* targeted disruption verification primer from nt. 114,550 to 114,572 of BACR13F06 5' ATCGAGTAGTGTGTGTGGTTGCC 3'	This work
WHITEKO DIAG1	*ptpmeg* targeted disruption verification primer 5' GAAGCGAGAGGAGTTTTGGCAC 3'	This work
MEGKO DIAG2	*ptpmeg* targeted disruption verification primer from nt. 117,820 to 117,798 of BACR13F06 5' GGTTCCGAGTGGTTCATTTCGTC 3'	This work
WHITEKO DIAG2	*ptpmeg* targeted disruption verification primer; 5' CGCACAATGTCCAACGGATACAC 3'	This work
Rhodococcus[a]		
JYP543	reverse primer to amplify pCR2.1Topo; 5' ACGCCAAGCTTGGTACCGAGC 3'	(76); This study
JYP557	forward primer to amplify AN12 genomic DNA; 5' GTGGCCATGTGATGACCCTCC 3'	(76); This study
JYP560	forward primer to amplify AN12 genomic DNA; 5' GGCGATGGCTACGAGTACCTG 3'	(76); This study
JYP561	reverse primer to amplify AN12 genomic DNA; 5' CGGGTGAAGCACGCTTCCTTC 3'	(76); This study
JYP576	reverse primer to amplify AN12 genomic DNA; 5' CGCCACTTGCCGTCCTTTTCC 3'	{Yang, submitted #107; This study

a. Work pertaining to either the *ptpmeg* (*Drosophila*) knockout or the *traA* (*Rhodococcus*) has been indicated.

3.2.3. Preparation and standard transformation of electrocompetent AN12 cells.

AN12 electrocompetent cells were prepared essentially as previously described (38), except cells were grown in NBYE/0.05% Tween-80 media in a 1L baffled flask with shaking until $O.D._{600}$ 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 1X TE buffer with 100 to 150 μl competent AN12 cells for 5 minutes prior to electroporation. This suspension was placed in an ice-cold sterile 2mm gapped electroporation cuvette. Electroporation was carried out at 2.5 kV, 25 μF, 400 Ω, with a Bio-Rad Gene Pulser (Hercules, CA). Immediately following pulsing, 200 μl of LB was added to the cells. Cells were allowed to recover after electroporation at 30°C for 2 hours with gentle agitation prior to selection with appropriate antibiotics.

3.2.4. Transformations of *D. melanogaster* with pUASt-*Ptpmeg* constructs and pTV2-MegISceI, and *R. erythropolis* AN12 with pJY37. The P-element constructs (pUASt-LD22982, pUASt-LD22982, or pTV2-MegISceI) were mixed (1:1 conc./conc.) with the Δ2-3 helper plasmid (63), then injected into embryos collected from isogenic fly strain, y^1 w^1 Canton-S. The chromosomal locations of these P-elements were mapped with the y^1 w^1 *CyO/Tft* balancer strain. Subsequent manipulations with P[FRT; $ptpmeg^{I\text{-}SceI}$; w^+; FRT] carrier flies to yield *ptpmeg* germline disruption strains was done as described (51), and will be discussed in further detail below. The *Rhodococcus* integration plasmid, pJY37, bears four features required to generate a *traA* site-specific gene disruption. First, the plasmid does not self-replicate in *Rhodococcus*. Second, the plasmid carries a selectable drug resistance marker. Third, the plasmid carries ≥ 600 bp of targeting sequence homologous and internal to the *traA* ORF, such that this cassette contains as little of the ORF as possible to avoid potential functions associated with the truncated *traA* gene product following recombination. Fourth, within this targeting cassette, there should be an unique restriction site (*Bsp*EI). To generate the *traA* mutant, 2 μg of this pJY37 was prelinearized at the *Bsp*EI site in a 20 μl digestion reaction, then 1 μl of this reaction was used for each transformation. The transformation was then carried out as described previously, except after electroporation, cells were allowed to recover for 24 hours prior to plating on selective media.

3.2.5. *Rhodococcus* conjugation and megaplasmid mobilization frequency. Matings were carried out as described earlier in Chapter 2, Section 2.

3.2.6. Pulsed-field gel electrophoresis (PFGE). PFG electrophoresis was performed as described earlier in Chapter 2, Section 2.

3.2.7. Southern blot analysis. Hybridizations were performed as described earlier in Chapter 2, Section 2.

3.2.8. *Rhodococcus erythropolis* colony PCR. Colony PCR reactions were performed as described earlier in Chapter 2, Section 2.

3.3. Results.

3.3.1. Targeted gene disruption of the protein tyrosine phosphatase, Ptpmeg, in *Drosophila melanogaster*. Protein tyrosine kinases (PTKs) and phosphatases (PTPs) are antagonistic modulators of substrate protein function, and mediate cell signaling in many developmental contexts. Examples of cytoplasmic PTK (*Abl*) and PTP (*Csw*) as well as transmembrane receptor type PTK, or RPTKs (*Sev*) and RPTPs (*Ptp69d*) have been identified as central regulators of *Drosophila* nervous system development (5, 7, 21, 22, 29). A previously uncharacterized protein tyrosine phosphatase, Ptpmeg, was identified which exhibits similarities to the mammalian phosphatase, PTPN4/PTPMEG, in that both proteins encode PDZ domains flanked by N-terminal FERM domains, and C-terminal PTP domains (Figure 3.1). PDZ domains often coordinate protein-protein and protein-scaffolding interactions at the cell surface (70). The first protein identified with a PDZ domain, PSD-95/SAP90, was found to bind to the carboxy terminus of the Shaker K+ channel (35). FERM domains- so named for the founding members of this protein family which bear this module, Band 4.1, Ezrin, Radixin, and Moesin- also localize proteins to the plasma membrane, and serve as links between the plasma membrane and the cytoskeleton (11). Because mammalian PTPN4/PTPMEG transcripts can be detected in the mouse cerebellum, and PTPMEG protein binds to the glutamate receptor subunits, GluRδ2 and GluRε1 (30), we reasoned that *Drosophila* Ptpmeg may also function in nervous system development.

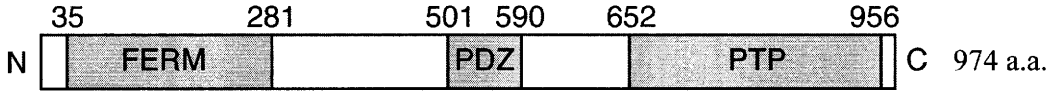


Figure 3.1. A schematic diagram of *Drosophila* Ptpmeg domains. Numbers above each indicated domain denotes amino acid positions in context of the full length (974 a.a.) protein encoded by the LD22982 cDNA. CD-Search (44) with the Ptpmeg sequence indicate that its FERM, PDZ, and PTP domains are 98%, 100%, and 99% similar to consensus sequences of domains with accession nos. CDD 24244, CDD 29049, and CDD 28929, respectively.

Two specific aims were designed to address Ptpmeg function. First, a stable *ptpmeg* disruption would be generated using the strategy outlined by Rong et al. (2000). Genetic complementation using either ubiquitous or tissue-specific Gal4 drivers and

pUASt constructs bearing the wild-type *Ptpmeg* would then allow us to address whether any putative phenotypes were specifically due to *ptpmeg* disruption. Second, the effects of *Ptpmeg* misexpression would be determined by driving these same pUASt constructs in an otherwise normal *Ptpmeg* genetic background. The *Drosophila Ptpmeg* locus is located on the left arm of the 3rd chromosome at cytological map position, 61C1.

To begin, the complete sequence of seven individual *Ptpmeg* cDNA clones generated by the cDNA and EST projects of Berkeley Drosophila Genome Database (BDGP) was determined. The analysis of these clones revealed that at least three different splice isoforms of Ptpmeg are expressed during development (Figure 3.2) (4, 36). The majority of the clones, LD22982, LD16634. LD24325, LD38703, and LP01515, encode a full length Ptpmeg protein of 975 amino acids. One clone, LD13416 encodes a N-terminally truncated protein predicted to contain only a partial FERM domain and to be 780 amino acids in length. The last clone, LD27491 encodes a N-terminally truncated, 580 residue protein with no FERM domain. P-element based UAS (pUASt) constructs bearing the LD22982, or *Ptpmeg*^{WT}, and LD27491, or *ptpmeg*$^{\Delta FERM}$ were generated and injected into 910 and 870 embryos, respectively. Eight independent transgenic fly lines bearing the former (UAS-*Ptpmeg*^{WT}) and thirteen lines bearing the latter (UAS- *ptpmeg*$^{\Delta FERM}$) were created. My preliminary experiments indicate that animals were phenotypically normal when UAS-*Ptpmeg*^{WT} expression was induced ubiquitously with *daughterless*-Gal4, or in the developing CNS with *elav*-Gal4. Interestingly, the ectopic expression of the UAS-*ptpmeg*$^{\Delta FERM}$ with *elav*-Gal4 resulted in completely penetrant late pupae death (possibly due to a failure of the flies to eclose) with flies exhibiting glazed rough eyes.

For gene targeting, a bacterial artificial chromosome (BAC) clone, BACR13F06, containing the entire *Ptpmeg* locus and neighboring genes was identified. DNA from BACR13F06 corresponding to an incomplete *Ptpmeg* from a *Spe*I site to an *Eco*RI site (internal fragment minus regulatory, start, and stop codons) was then subcloned into an intermediate vector, pBluescript KS(+). Prior to introducing this *Ptpmeg* fragment to the targeting vector, pTV2 (51), modifications were made such that 1) an *I-SceI* recognition sequence was inserted into the FERM encoding sequence to generate the double-stranded break, and 2) an *Acc*65I site within the PDZ encoding sequence was digested, treated

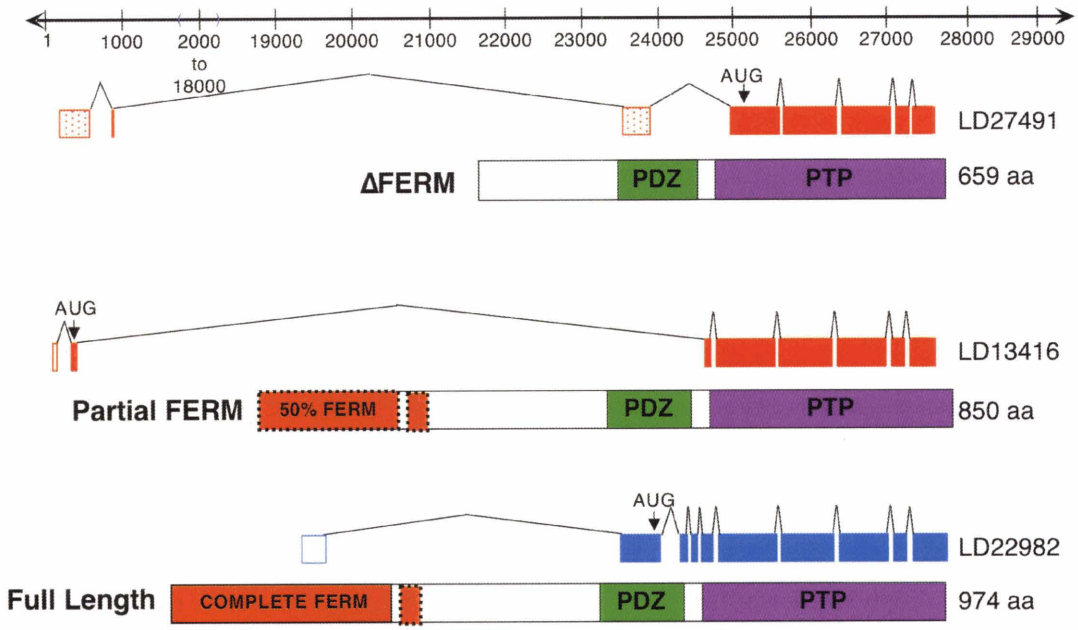


Figure 3.2. The splicing variants and end products of *Ptpmeg*. Numbering on the ruler reflect lengths in basepairs at the *Ptpmeg* genomic region. Exons (coding sequencing in solid, non-coding in dotted) are depicted as boxes and introns as gaps. Junctions were determined by comparing cDNA (LD27491, LD13416, and LD22982) sequences against the genomic DNA. The predicted start sites for each isoform predicted is indicated (AUG and arrow). Smaller transcripts (red) are differentiated from the full length transcript (blue). Products of these transcripts are shown below and labeled with domains.

with Klenow polymerase, and then religated to induce premature stop codons in the final targeting product, such that only a protein with a FERM domain and a partial PDZ would be expressed. Since all of the Ptpmeg splice variants appear to encode the PDZ and PTP domains, this targeted disruption should affect Ptpmeg function regardless of any potential transcriptional regulation (Figure 3.3).

A diagram of the fly crosses used to generate the final homozygous *ptpmeg* mutant flies is outlined in Figure 3.4. Roughly 1,800 y^1 w^1 homozygous embryos were injected with the pTV2-MegISceI construct. Two independent transformants out of 279 survivors were analyzed further. The location of the P-elements (P[FRT-MegISceI-w^+-FRT]) in both of these lines were mapped to the 2nd chromosome, and two "donor" lines, KA6-1 and KA6-2, that were isogenic for the P[FRT-MegISceI-w^+-FRT] insertion were established. These donor flies were mated to flies carrying inducible (heat-shock)

sources of the FLP recombinase and I-*Sce*I enzyme on the 2nd chromosome. Embryos, 0-3 day old, collected from this mating were then heat-shocked. Over 900 virgins with mosaic eyes, which were heterozygous at the 2nd chromosome, were then mated to males carrying sources of eye-specific FLP recombinase (*ey*-FLP). This cross is expected to yield w^- eyes in animals wild type for the *Ptpmeg* locus, because the intact P-element (on the second chromosome) contains two FRT sites on the second chromosome which, upon exposure to FLP, should be excised in the eye. Conversely, the *w*+ marker should be stable in animals in which the P-element was mobilized in the germline, since only one FRT site remains following homologous recombination during the initial treatment with heat-shock induced FLP recombinase. Of the 30,000 progenies screened from this cross, 11 animals (A through K) retained the w^+ phenotype. The *w*+ marker mapped to the 2nd chromosome in three of these animals (A, C, and D), and hence reflected non-specific integration events of the excised P[FRT-MegISceI-w+-FRT]. These lines were not examined further. The *ptpmeg* targeting cassette-associated w+ marker all mapped to the third chromosome in lines B, E, F, G, H, I, J, and K. Line F perished without progeny. These lines were then balanced with *TM3/TM6b* flies. My preliminary data suggested that the putative *ptpmeg* disruption does not affect the viability of these flies, as homozygous flies could be generated. However, few progeny result when these *ptpmeg/ptpmeg* animals were mated to each other.

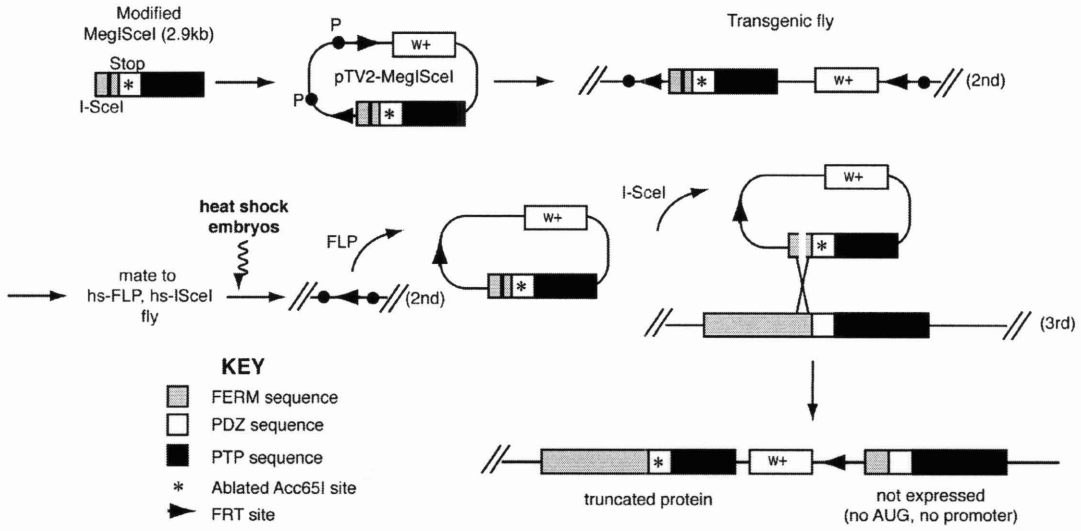


Figure 3.3. *In vivo* targeting of the *Drosophila Ptpmeg* locus.

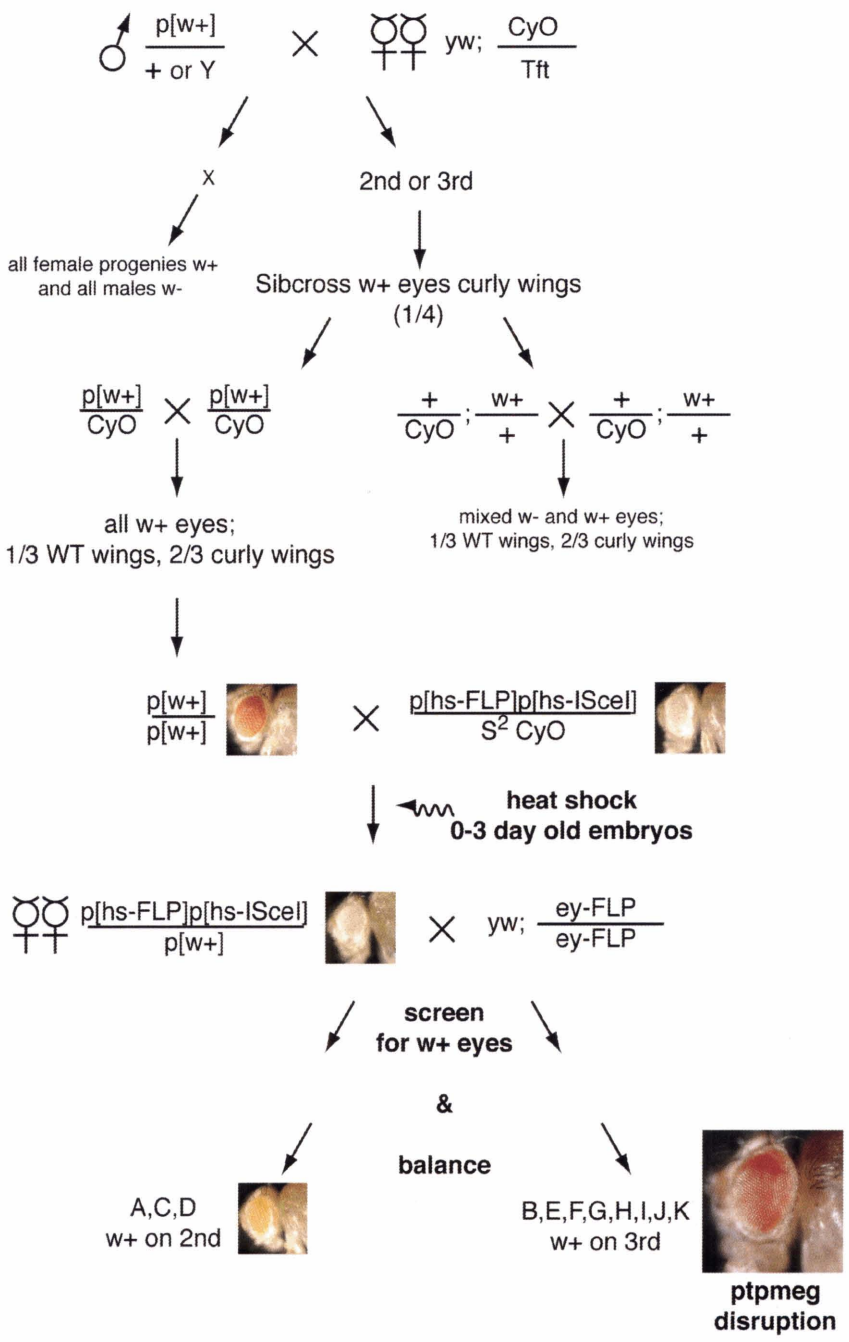


Figure 3.4. Genetics of generating a targeted *ptpmeg* disruption. All relevant crosses after the generation of pTV2-MegISceI transformants are shown. P[FRT-MegISceI-w+-FRT] are abbreviated (p[w+]). Desirable outcomes are indicated with larger font size than undesirable outcomes.

Genomic DNA from homozygous flies of remaining lines (B, E, G, H and I) was prepared for further analysis. Two pairs of primers, MEGKODIAG1/WHITEKODIAG1 and WHITEKODIAG2/MEGKODIAG2, were designed to specifically amplify *ptpmeg* sequences at targeted loci (Figure 3.5A), since one primer in each pair is specific to genomic sequence not found on targeting vector, pTV2, and the other primer in each pair is specific to pTV2 sequence. Indeed, genomic DNA from B, E, G, H, and I, gave rise to the expected PCR products (Figure 3.5B). The first set of PCR products from each of these lines were also digested with *Acc*65I, since homologous recombination could lead to the gene replacement of this site, and hence the removal of premature stop codons introduced with the ablated *Acc*65I site. Results indicated that this mutation was not replaced with wild type sequences in lines B, E, H and I, and should yield a carboxyl-terminal truncated product as expected (Figure 3.5C).

An antibody against the Ptpmeg protein, which could differentiate between mutant from wild type animals, would have been helpful to confirm the nature of this mutant allele, however, I was not able to generate this reagent before leaving the Garrity Lab, in which this part of thesis was performed. I was able to briefly note that the mutant animals exhibited duplicated bristles at the scutellum (Figure 3.6), on the head, and on the wings. Also, the homozygous animals tended to exhibit uncoordinated and twitching behaviors.

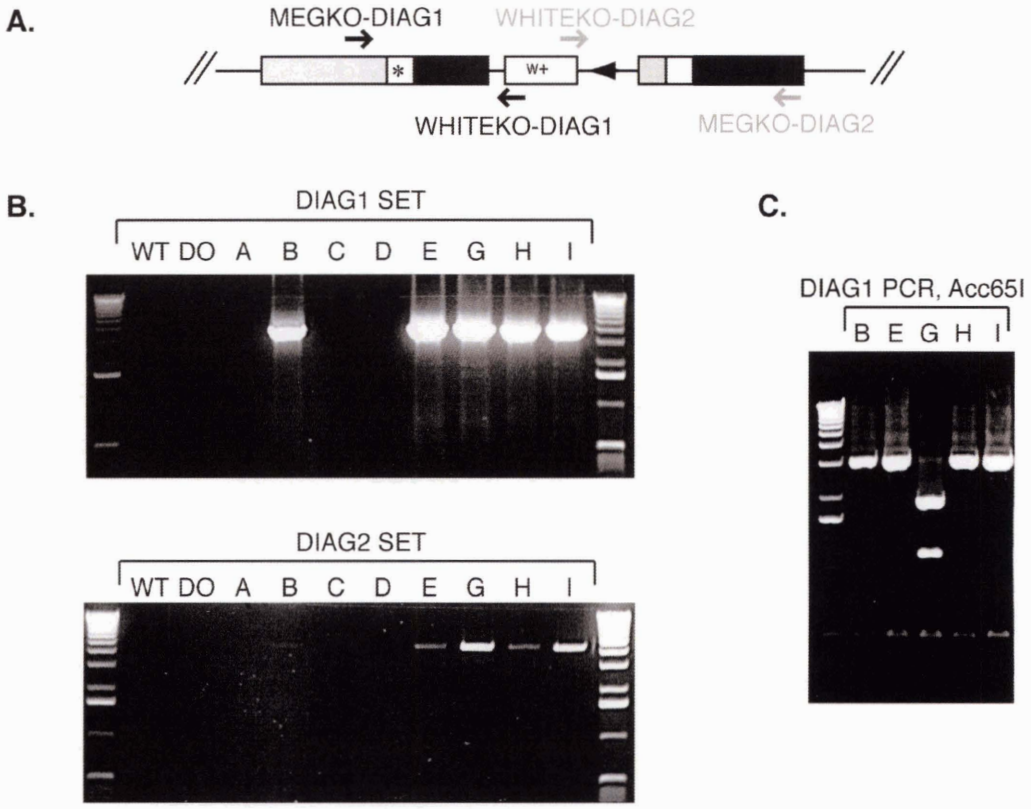


Figure 3.5. Verification of the targeted *ptpmeg* locus. (A and B) Primers, MEGKO-DIAG1/WHITEKO-DIAG1 (black) and WHITEKO-DIAG2/MEGKO-DIAG2 (gray) were used to amplify genomic DNA from the indicated *ptpmeg* mutants. WT (wild type) and DO (donor) genomic DNA samples (P[w+]/P[w+]) were used as controls. (C) PCR products from the DIAG1 reactions were purified and digested with *Acc*65I to detect the retention of the nonsense mutation. Expected sizes from *ptpmeg* mutant flies are 3.0 kb and 0.5 kb; is the *Acc*65I site is fixed during recombination, the expected bands sizes are 1.8 kb, 1.2 kb and 0.5 kb. DNA ladder from X were used in these panels to estimate fragment sizes.

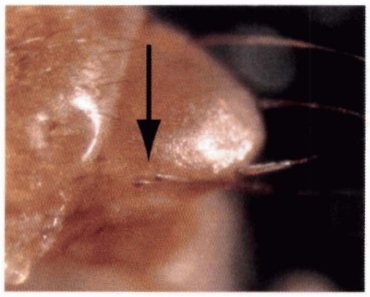


Figure 3.6. Adult *ptpmeg* homozygous mutants exhibit duplicated scutellar bristles. One or more scutellar bristles can be found in 4/5 homozygous flies in confirmed lines, B, E, H and I.

3.3.2. Sequence analysis of pREA400 encoded *traA*. One aim of the *Rhodococcus* megaplasmid 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 this pREA400 megaplasmid *traA* encodes a protein of 1,506 amino acids (a.a.). Proteins putatively involved in plasmid transfer isolated from Gram-positive actinomycetes, *Gordonia westfalica* (10) and *Arthrobacter aurescens* (54), were among those most similar to the pREA400 TraA. The TraA (Genbank no. CAE09129) from *G. westfalica* is encoded by a 101 kb megaplasmid, and is 31% identical to pREA400 TraA. It's unclear whether the *traA* gene product (Genbank no. AAS20144) from *A. aurescens* also resides on a megaplasmid, however, this protein is 29% identical to pREA400 TraA.

A conserved domain (CD) search revealed a pREA400 TraA region (a.a. 544 to 963) that is similar to ATP-dependent RecD type helicases (44). In addition, this region is 27% identical to TraA relaxase/helicase encoded by the *Agrobacterium tumerifaciens* pTiC58 plasmid (19), and 22% identical to the TraI relaxase/helicase of the *Escherichia coli* F factor (2, 41, 65).

It was not clear based on BLAST and CD search analysis whether pREA400 *traA* encodes relaxase function. Recently, a crystal structure of the relaxase domain of F factor TraI protein (Genbank no. AAC44186) was reported, and key residues (Y16, D81, H146, H157, and H159) of active site were determined (41). Based on the previously described relaxase/helicases, we anticipated that the pREA400 TraA relaxase domain may be located at the N-terminus. Thus, a partial alignment of pREA400 TraA a.a. 1-369 to sequences from F factor TraI, as well as the *Gordonia* and *Arthrobacter* TraA's using ClustalW analysis was completed (Fig. 3.7). We found that pREA400 TraA is 26%, 43% and 42% identical to F factor TraI, pKB1 TraA, and *Arthrobacter* TraA, respectively, over this N-terminal region. Importantly, all of the key residues required for catalysis and divalent cation binding found in the crystal study of F factor TraI were conserved. Taken with the alignment to helicases, we propose that pREA400 TraA, like the F factor TraI, is a consolidated relaxase/helicase conjugation initiating protein.

```
AAC44186          1   MMSIAQVRSAGSAGNYYTDKDNY.............  23
pREA400_traA      1   --MTLHVLHAGDGYEYLTSQVAT.............  21
CAE09129          1   MTATIHKLTAGDGYEYLTKSVAA.............  23
AAS20144          1   -MMSLHVLSAGTGYLYYTQETAS.............  22

AAC44186         79  GYDLTFSAPKSVSMMAMLGED---KRLIDAMNQAVD 111
pREA400_traA    195  GFDLVFTPPKSVSTMWALADDDLRRQIERIHHETVK 230
CAE09129        206  GYDMTFSPVKSVSALWAVAPLPMAEKIEAAHRQAVA 241
AAS20144        186  GYDLTFTPVKSISVLWALGDADTRRLVEDAQQAALN 221

AAC44186        112  FAVRQVEALAS-TRVMTDGQSETVLTGNLVMALFNH 146
pREA400_traA    231  DTLGWIEKEACFTRTGATSQEHDQTTG-VVATLYDH 265
CAE09129        242  DALAFIEREACLSRLGTDGIAQVDTDG-LIAAAFTH 276
AAS20144        222  DSIEYLETHALATRLGTNGIAQSSVKGGLTATAFRH 257

AAC44186        147  DTSRDQEPQLHTHAV--VANVTQHNG---EWKTLSS 177
pREA400_traA    266  YDSRAGDPNLHTHAVLSVKVCTEKDG---KWRALFS 298
CAE09129        277  RDSRAGDPDLHTHVAISNKVRVRDAAGITRWMALDG 312
AAS20144        258  HDSRLGDPNLHTHVVVSNKVQDLAGN----WKSIDG 289

AAC44186        178  DKVGKTGFIENVYANQIAFGRLYREKLKEQVEALGY 213
pREA400_traA    299  STLHRYGVPASQRYNAAIMSKLHTELGFGLTERSTG 334
CAE09129        313  TPLFKATVSASEVYNSRIELYLQRDLGVSFSERFSA 348
AAS20144        290  KLLHRSAVAVSEHYNTRIQAHLEDHG--VRFEARTV 323

AAC44186        214  ETEVVGKHGMWEMPGVPVEAFSGRSQTIREAVGEDA 249
pREA400_traA    335  RGRQN-VVEIAEVPQQLCEMFSSRRTQIETRRDQLV 369
CAE09129        349  DVRKRPVREIDGVSAELMARWSSRRAAIEERIDELS 384
AAS20144        324  NGSKQPVMEIASVPRELISLFSKRSEGIRTSLTELR 359
```

Figure 3.7. Alignment of four DNA relaxase domains. The TraI and TraA gene products isolated from *Escherichia* coli F plasmid, *Gordonia westfalica* pKB1 plasmid, and *Arthrobacter aurescens* are indicated with accession nos. AAC44186, CAE09129, and AAS20144, respectively, and have been aligned to pREA400 TraA residues 1 to 369. The alignment was performed using the ClustalW program (http://ebi.ac.uk/clustalw). Identical amino acids residues present in at least three of the four sequences are highlighted (gray boxes). Crystallographically determined (41) active site residues of the relaxase domain of F plasmid (Y16, D81, H146, H157, and H159), and aligning residues are indicated (bold). Also bolded are residues that align with R150 of the F plasmid, known to be important DNA substrate binding (28). Single a.a gaps in the alignment are indicated with dashes, and larger gaps (>10 a.a.) in the alignment are indicated with periods. The numberings 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.

3.3.3. Development and deployment of a new targeted gene disruption method to study *traA* gene function. A new 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 (14, 68). This phenomenon was first exploited to achieve genetic transformations of *Saccharomyces cerevisiae* (46). In this context, the targeting cassette is prelinearized *in vitro* prior to transformation of yeast cells. For *Drosophila melanogaster* (51) and other organisms (33), the targeting cassette is linearized *in vivo* by the expression of rare-cutting endonucleases.

The approach we adopted to generate a *traA* disruption is most similar to the yeast model (Figure 3.8). Briefly, a 707 bp fragment corresponding to nt. 51 to 758 of the predicted *traA* ORF was amplified with PCR for use as the targeting cassette. This PCR product was used to construct the disruption plasmid, pJY37, which was prelinearized at the unique *Bsp*EI site within the targeting cassette, then introduced to AN12 WT cells via electroporation. Cells were allowed to recover for 24 hours, then transformants that survived on selective media were further analyzed.

A colony PCR assay was developed to quickly screen whether the drug resistant transformants recombined at the *traA* locus or illegitimately elsewhere in the genome. Primers JYP577, which annealed to sequences upstream of the *traA* ORF (and therefore not present in the disruption vector), and JYP543, which annealed to sequences within the vector backbone, were designed for this assay (Figure 3.8). Only recombinants that integrated homologously at the *traA* locus should give the expected 800 bp product. Of the one hundred transformants analyzed with colony PCR, one gave a positive result (Figure 3.9). We named this candidate mutant strain, JY825.

We expected this 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 an amino-terminal truncated protein, the length of which is predicted to be less than 250 a.a.. The second *traA* includes most of the ORF,

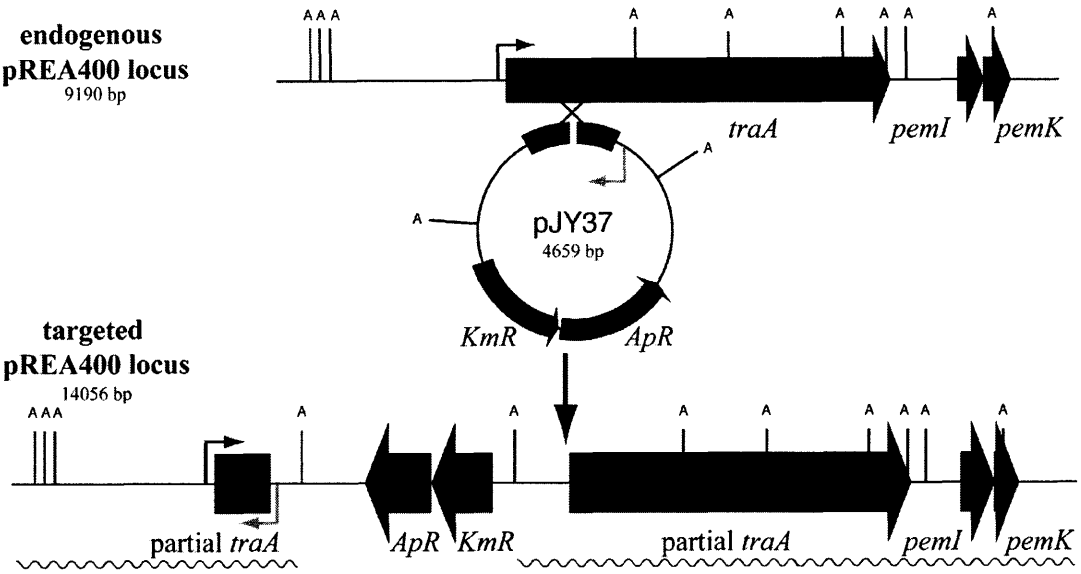


Figure 3.8. Schematic diagram of the *Rhodococcus* targeted gene-disruption strategy. The 9,190 bp of *Eco*RI fragment from pREA400 is shown with the *traA* ORF (nt. 2,691 to 7,214), *pemI* ORF (nt. 8,001 to 8,309) and *pemK* ORF (nt. 8,305 to 8,625). The integration plasmid, pJY37, is a derivative of pCR2.1Topo, and carries a *traA* specific DNA (nt. 2,714 to 3,421 of pEZTn1F6). pJY37 was prelinearized at the unique *Bsp*EI site to induce homologous recombination. The targeted pREA400 locus (14,056 bp) is depicted following the integration of pJY37. *Afl*III recognition sequences are labeled. Primer JYP557 designed for colony PCR screening anneals to sequences (nt. 2,680 to 2,700) upstream of the *traA* ORF, and sites where JYP557 anneal are depicted as a bent black arrows. Primer JYP543 designed for the same purpose anneals to pCR2.1Topo/pJY37 sequences 69 bp downstream of the region of *traA* homology. Sites where JYP543 anneal are depicted as bent gray arrows. DNAs predicted to anneal to a Southern probe (pEZTn1F6 template) are underlined (wavy lines).

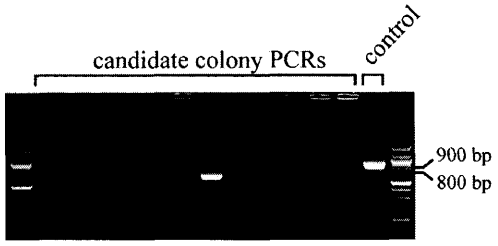


Figure 3.9. Identification and verification of the candidate *traA* disruption mutant strain, JY825. Colony PCR-based screening using primers JYP557 and JYP543 should yield an approximately 800 bp product if the *traA* locus is disrupted as expected. Each lane represents a kanamycin resistant AN12(pJY37) 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 ug 100 bp DNA ladder.

but no start codon nor promoter, and should not be expressed. Southern blot analysis was used to verify the targeted locus, using the digoxygenin probe generated from pEZTn1F6. Genomic DNA samples of AN12 WT, JY825, AN12PL-1F6 and SQ1 digested with *Afl*III enzyme were examined (Figure 3.10). 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.

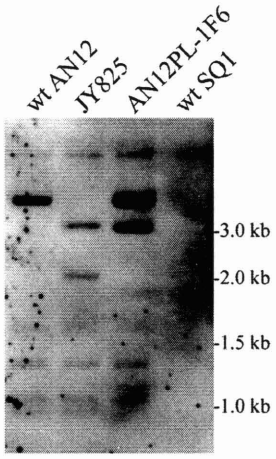


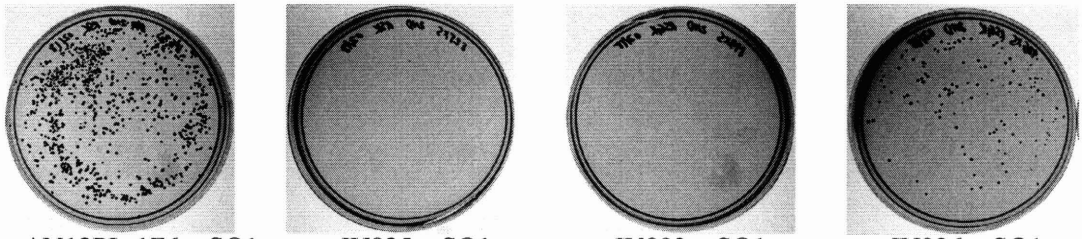
Figure 3.10. Verification of the candidate *traA* disruption mutant strain, JY825, with southern blot analysis. Genomic DNA fragments digested with *Afl*III from the indicated strains were hybridized to DIG-labeled probe made from pEZTn1F6 plasmid template. Expected hybridizing fragment sizes for DNA from each of the following strains are as follows: wild-type AN12- 3,588 bp, 1,342 bp, 1,100 bp, and 1,021 bp; JY825- 3,210 bp, 2,227 bp, 1,342 bp, 1,100 bp and 1,021 bp; AN12PL-1F6- 3,588 bp, 3,022 bp, 1,342 bp, and 1,100 bp. No hybridizing bands were expected in wild-type SQ1.

3.3.4. Phenotypic analysis of the *traA* mutant and genetic complementation of the *traA* defect. JY825 was mated to SQ1, along with AN12PL-1F6 as a positive control 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 pREA400 insertions of Km^R cassettes. While the AN12PL-1F6 donor strain frequently gives rise to greater than 200,000 transconjugant per mating, JY825 gives rise to less than 30 transconjugants per mating (Table 3.4, Fig. 3.11). This strongly suggests that *traA* gene function is required for pREA400 megaplasmid conjugation.

TABLE 3.4. Summary of mobilization frequency of AN12 megaplasmids to SQ1.

Donor Strain	No. Trials	Weighted Average Frequency[a]	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}

a. Frequency expressed as the number of transconjugant CFUs on selective plates by the number of recipient CFUs observed on permissive plates.



AN12PL-1F6 x SQ1 JY825 x SQ1 JY893 x SQ1 JY926 x SQ1

Figure 3.11. Genetic and phenotypic complementation of the *traA* disruption mutant. Donor strains AN12PL-1F6, JY825, JY893, or JY926 were mated to SQ1 recipients (marked as donor X 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.

To make certain that the defect in pREA400 transfer was specifically due to *traA* disruption and not 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 gentamycin resistance (Gm^R). The entire *traA* ORF and flanking sequences was subcloned into pAL349 digested with the *Fsp*I enzyme, resulting in the complementation plasmid, pJY49B. A control plasmid, pJY48, was also generated at the same time by religating pAL349 vector after the *Fsp*I digestion.

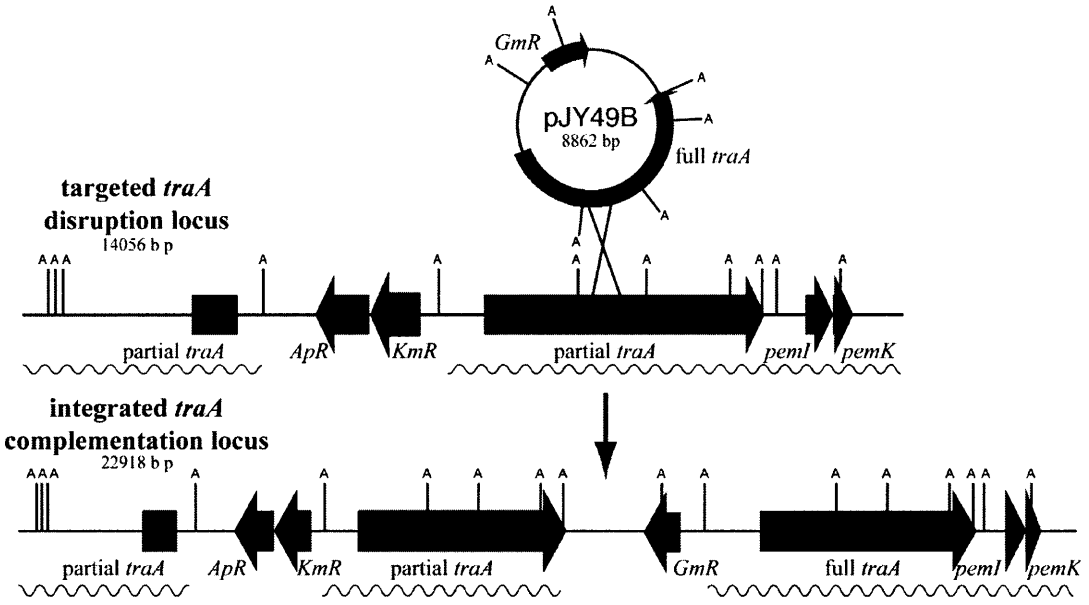


Figure 3.12. The integration of the complementation plasmid, pJY49B at the disrupted *traA* locus is depicted. pJY49B contains 4,900 bp homologous sequence corresponding to the entire *traA* ORF, 221 bp of upstream sequence, as well as 156 bp downstream sequence. The exact site of pJY49B integration is unknown. Sites corresponding to *Afl*III recognition sequences are labeled (A), as are the regions which anneal to Southern probes generated from pEZTn1F6 template (wavy lines).

Both pJY48 and pJY49B were introduced to JY825 *traA* mutant strain via electroporation, and transformants were selected following a 24 hour recovery period. Since neither pJY48 or pJY49B can replicate in *Rhodococcus*, the only way the cells could express the drug resistance is if the plasmids had recombined into the genome. Greater than 500 CFUs that were Gm^R were observed when pJY49B was used as the transforming DNA, whereas only 3 CFUs were observed when pJY48 was used. We named the 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 (Figure 3.12). This was indeed the case for the transformant, JY926, as confirmed by Southern blot analysis (Figure 3.13).

JY926 and JY893 were tested for the ability to give rise to triply resistant (Km^R Rf^R Sm^R) transconjugants 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 (Figure 3.11). In

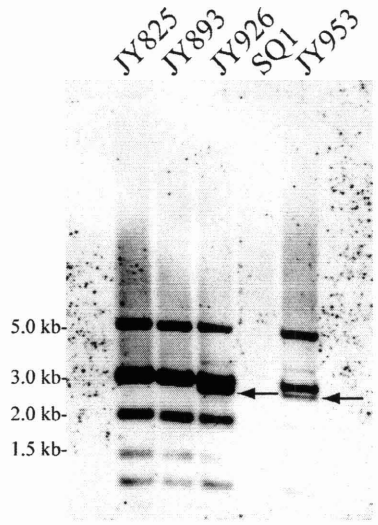


Figure 3.13. Southern blot confirmation. Genomic DNA from each sample was linearized with *Afl*III, then probed with digoxygenin labeled probe synthesized with pEZTn1F6 template DNA. Expected annealing DNA fragments are as follows: JY825 and JY893- 3,210 bp, 2,227 bp, 1,342 bp, and 1,100 bp; JY926 and JY953(JY926 x SQ1 transconjugant)- 3,210 bp, 2,855 bp, 2,227 bp, 1,342 bp and 1,100 bp. No DNA from SQ1 was expected to hybridize to the probe. The 2,855 bp band unique to the JY926 and JY953 are indicated (arrows).

contrast, the full copy of *traA* in the JY926 strain does indeed complement the original *traA* disruption in JY825, although JY926 does not transfer pREA400 to SQ1 as efficiently as AN12PL-1F6 (Table 3.4). From these data, we conclude that the pREA400 encoded *traA* gene function is required for megaplasmid transfer in *R. erythropolis* AN12.

3.4. Discussion.

We have devised a new targeted gene disruption strategy to study *Rhodococcus* gene function. Like similar strategies employed in eukaryotic systems for targeted insertions discussed previously, our method also relies on double-stranded DNA breaks to stimulate homologous recombination. In other bacterial systems, a number of DNA processing factors, including helicases, nucleases, and synapsis proteins, are required for DNA repair (6, 14). Two independent pathways, RecBCD and RecF, have been proposed to function in *Escherichia coli* (13), and both are reliant on the synapsis function of the RecA protein (62). DNA repair is an integral part of replication; as would

be expected, RecA is well conserved from bacteria to humans. Though DNA repair has not been characterized in rhodococci, it is expected that RecA plays a critical role in homologous recombination since sequence analysis of other actinomycetes suggests the presence of *recA* homologues in *Nocardia farcinica*, *Streptomyces coelicolor*, as well as *Mycobacterium tuberculosis* (http://www.expasy.org/unirules/MF_00268).

Although we have used this targeted gene disruption strategy to study the effects of *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. Evidence suggests that larger targeting cassettes facilitate the recombination in *Rhodococcus*, as the integration event of the pJY39 rescue construct was readily generated, whereas the *traA* gene disruption event was rare. Our preliminary evidence indicates that at least 600 bp of homologous sequence is necessary for site-specific recombination (data not shown) in *R. erythropolis* AN12. Thus, a modification to our approach will be necessary to study gene functions of smaller ORFs; in particular, the targeting cassette will need to be modified prior to recombination to encode nonsense mutations, both upstream and downstream of the unique restriction site to ensure gene disruption.

Though many putative relaxase functions have been annotated and predicted by various *Rhodococcus* sequencing efforts (59, 64), the present study is the first genetic analysis of a determinant of megaplasmid conjugation 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 two Gram-positive plasmid encoded relaxases have been examined; the *traA* gene product from the 30 kb *Streptococcus* plasmid, pIP501, and the *mobM* gene product from 5 kb *Streptococcus* plasmid, pMV158 (16, 24, 27, 37, 39).

For these relaxases, it is known that the transesterification reaction 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 (41, 47, 57, 65). It is expected that the relaxase encoded by pREA400 *traA* will function similarly because these residues appear to be conserved (Fig. 3.7). 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*.

Interestingly, pREA400 TraA, as well as the TraA's from *Gordonia westfalica* and *Arthrobacter aurescens* encode a gap of approximately 250 residues between the tyrosine and the first histidine in the active site, whereas F factor TraI and other Gram-negative bacteria relaxases encode gaps that are significantly smaller (150 residues or less). We do not anticipate that this gap length difference is solely due to plasmid origin from either Gram-negative or Gram-positive bacteria, since the pIP501 relaxase from the Gram-positive *Streptococcus* also exhibits a 100 a.a. gap between Y19 and H119 (NCBI accession no. AAA99466). Nor does the gap length appear to be dependent on the size of the plasmid, as the *Gordonia* pKB1 and *E.* coli F factor are plasmids of comparable sizes (101,000 bp and 99,159 bp, respectively). We speculate that this discrepancy in gap lengths may instead reflect the requirements to recognize, accommodate, and/or act upon differing DNA substrates or *nic* regions (40). Identification of the *nic* site of *R. erythropolis* pREA400 should help address this issue.

Our study suggests that the basic machinery required to initiate plasmid transfer appear to be evolutionarily conserved from previously well characterized plasmid, such as F, pTiC58, and RP4 to a newly characterized actinomycete megaplasmid, pREA400. Further characterization of the *traA* function, regulation, gene product localization and protein-protein interactions will elucidate conserved and novel aspects of actinomycete megaplasmid conjugation.

3.5. References.

1. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. Science **282:**2012-8.
2. **Abdel-Monem, M., G. Taucher-Scholz, and M. Q. Klinkert.** 1983. Identification of *Escherichia coli* DNA helicase I as the *traI* gene product of the F sex factor. Proc Natl Acad Sci U S A **80:**4659-63.
3. **Adams, M. D., S. E. Celniker, R. A. Holt, C. A. Evans, J. D. Gocayne, P. G.**

Amanatides, S. E. Scherer, P. W. Li, R. A. Hoskins, R. F. Galle, R. A. George, S. E. Lewis, S. Richards, M. Ashburner, S. N. Henderson, G. G. Sutton, J. R. Wortman, M. D. Yandell, Q. Zhang, L. X. Chen, R. C. Brandon, Y. H. Rogers, R. G. Blazej, M. Champe, B. D. Pfeiffer, K. H. Wan, C. Doyle, E. G. Baxter, G. Helt, C. R. Nelson, G. L. Gabor, J. F. Abril, A. Agbayani, H. J. An, C. Andrews-Pfannkoch, D. Baldwin, R. M. Ballew, A. Basu, J. Baxendale, L. Bayraktaroglu, E. M. Beasley, K. Y. Beeson, P. V. Benos, B. P. Berman, D. Bhandari, S. Bolshakov, D. Borkova, M. R. Botchan, J. Bouck, P. Brokstein, P. Brottier, K. C. Burtis, D. A. Busam, H. Butler, E. Cadieu, A. Center, I. Chandra, J. M. Cherry, S. Cawley, C. Dahlke, L. B. Davenport, P. Davies, B. de Pablos, A. Delcher, Z. Deng, A. D. Mays, I. Dew, S. M. Dietz, K. Dodson, L. E. Doup, M. Downes, S. Dugan-Rocha, B. C. Dunkov, P. Dunn, K. J. Durbin, C. C. Evangelista, C. Ferraz, S. Ferriera, W. Fleischmann, C. Fosler, A. E. Gabrielian, N. S. Garg, W. M. Gelbart, K. Glasser, A. Glodek, F. Gong, J. H. Gorrell, Z. Gu, P. Guan, M. Harris, N. L. Harris, D. Harvey, T. J. Heiman, J. R. Hernandez, J. Houck, D. Hostin, K. A. Houston, T. J. Howland, M. H. Wei, C. Ibegwam, et al. 2000. The genome sequence of *Drosophila melanogaster*. Science **287:**2185-95.

4. **Adams, M. D., J. M. Kelley, J. D. Gocayne, M. Dubnick, M. H. Polymeropoulos, H. Xiao, C. R. Merril, A. Wu, B. Olde, R. F. Moreno, and et al.** 1991. Complementary DNA sequencing: expressed sequence tags and human genome project. Science **252:**1651-6.
5. **Allard, J. D., H. C. Chang, R. Herbst, H. McNeill, and M. A. Simon.** 1996. The SH2-containing tyrosine phosphatase corkscrew is required during signaling by sevenless, Ras1 and Raf. Development **122:**1137-46.
6. **Amundsen, S. K., and G. R. Smith.** 2003. Interchangeable parts of the *Escherichia coli* recombination machinery. Cell **112:**741-4.
7. **Basler, K., and E. Hafen.** 1988. Control of photoreceptor cell fate by the sevenless protein requires a functional tyrosine kinase domain. Cell **54:**299-311.
8. **Blattner, F. R., G. Plunkett, 3rd, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. Collado-Vides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau, and Y. Shao.** 1997. The complete genome sequence of *Escherichia coli* K-12. Science **277:**1453-74.
9. **Brand, A. H., and N. Perrimon.** 1993. Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development **118:**401-15.
10. **Broker, D., M. Arenskotter, A. Legatzki, D. H. Nies, and A. Steinbuchel.** 2004. Characterization of the 101-kilobase-pair megaplasmid pKB1, isolated from the rubber-degrading bacterium *Gordonia westfalica* Kb1. J Bacteriol **186:**212-25.
11. **Chishti, A. H., A. C. Kim, S. M. Marfatia, M. Lutchman, M. Hanspal, H. Jindal, S. C. Liu, P. S. Low, G. A. Rouleau, N. Mohandas, J. A. Chasis, J. G. Conboy, P. Gascard, Y. Takakuwa, S. C. Huang, E. J. Benz, Jr., A. Bretscher, R. G. Fehon, J. F. Gusella, V. Ramesh, F. Solomon, V. T. Marchesi, S. Tsukita, K. B. Hoover, and et al.** 1998. The FERM domain: a unique module involved in the linkage of cytoplasmic proteins to the membrane.

Trends Biochem Sci **23:**281-2.

12. **Choulika, A., A. Perrin, B. Dujon, and J. F. Nicolas.** 1995. Induction of homologous recombination in mammalian chromosomes by using the I-*Sce*I system of *Saccharomyces cerevisiae*. Mol Cell Biol **15:**1968-73.
13. **Clark, A. J.** 1991. *rec* genes and homologous recombination proteins in *Escherichia coli*. Biochimie **73:**523-32.
14. **Cox, M. M.** 1999. Recombinational DNA repair in bacteria and the RecA protein. Prog Nucleic Acid Res Mol Biol **63:**311-66.
15. **Dabbs, E. R.** 1987. A generalised transducing bacteriophage for *Rhodococcus erythropolis*. Mol Gen Genet **206:**116-120.
16. **de Antonio, C., M. E. Farias, M. G. de Lacoba, and M. Espinosa.** 2004. Features of the plasmid pMV158-encoded MobM, a protein involved in its mobilization. J Mol Biol **335:**733-43.
17. **Desomer, J., M. Crespi, and M. Van Montagu.** 1991. Illegitimate integration of non-replicative vectors in the genome of *Rhodococcus fascians* upon electrotransformation as an insertional mutagenesis system. Mol Microbiol **5:**2115-24.
18. **Donoho, G., M. Jasin, and P. Berg.** 1998. Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol **18:**4070-8.
19. **Farrand, S. K., I. Hwang, and D. M. Cook.** 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:**4233-47.
20. **Fire, A., S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello.** 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature **391:**806-11.
21. **Garrity, P. A., C. H. Lee, I. Salecker, H. C. Robertson, C. J. Desai, K. Zinn, and S. L. Zipursky.** 1999. Retinal axon target selection in *Drosophila* is regulated by a receptor protein tyrosine phosphatase. Neuron **22:**707-17.
22. **Gertler, F. B., R. L. Bennett, M. J. Clark, and F. M. Hoffmann.** 1989. Drosophila *abl* tyrosine kinase in embryonic CNS axons: a role in axonogenesis is revealed through dosage-sensitive interactions with *disabled*. Cell **58:**103-13.
23. **Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin, and S. G. Oliver.** 1996. Life with 6000 genes. Science **274:**546, 563-7.
24. **Grohmann, E., L. M. Guzman, and M. Espinosa.** 1999. Mobilisation of the streptococcal plasmid pMV158: interactions of MobM protein with its cognate *oriT* DNA region. Mol Gen Genet **261:**707-15.
25. **Guo, S., and K. J. Kemphues.** 1995. *par-1*, a gene required for establishing polarity in *C. elegans* embryos, encodes a putative Ser/Thr kinase that is asymmetrically distributed. Cell **81:**611-20.
26. **Gutierrez, T., R. P. Learmonth, P. D. Nichols, and I. Couperwhite.** 2003. Comparative benzene-induced fatty acid changes in a *Rhodococcus* species and its benzene-sensitive mutant: possible role of myristic and oleic acids in tolerance. J Chem Ecol **29:**2369-78.

27. **Guzman, L. M., and M. Espinosa.** 1997. The mobilization protein, MobM, of the streptococcal plasmid pMV158 specifically cleaves supercoiled DNA at the plasmid *oriT*. J Mol Biol **266:**688-702.
28. **Harley, M. J., D. Toptygin, T. Troxler, and J. F. Schildbach.** 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:**6460-8.
29. **Herbst, R., P. M. Carroll, J. D. Allard, J. Schilling, T. Raabe, and M. A. Simon.** 1996. Daughter of sevenless is a substrate of the phosphotyrosine phosphatase Corkscrew and functions during sevenless signaling. Cell **85:**899-909.
30. **Hironaka, K., H. Umemori, T. Tezuka, M. Mishina, and T. Yamamoto.** 2000. The protein-tyrosine phosphatase PTPMEG interacts with glutamate receptor delta 2 and epsilon subunits. J Biol Chem **275:**16167-73.
31. **Hoskins, R. A., C. R. Nelson, B. P. Berman, T. R. Laverty, R. A. George, L. Ciesiolka, M. Naeemuddin, A. D. Arenson, J. Durbin, R. G. David, P. E. Tabor, M. R. Bailey, D. R. DeShazo, J. Catanese, A. Mammoser, K. Osoegawa, P. J. de Jong, S. E. Celniker, R. A. Gibbs, G. M. Rubin, and S. E. Scherer.** 2000. A BAC-based physical map of the major autosomes of *Drosophila melanogaster*. Science **287:**2271-4.
32. **Jackson, J. A., and G. R. Fink.** 1981. Gene conversion between duplicated genetic elements in yeast. Nature **292:**306-11.
33. **Jasin, M.** 1996. Genetic manipulation of genomes with rare-cutting endonucleases. Trends Genet **12:**224-8.
34. **Kennerdell, J. R., and R. W. Carthew.** 1998. Use of dsRNA-mediated genetic interference to demonstrate that *frizzled* and *frizzled 2* act in the *wingless* pathway. Cell **95:**1017-26.
35. **Kim, E., M. Niethammer, A. Rothschild, Y. N. Jan, and M. Sheng.** 1995. Clustering of Shaker-type K+ channels by interaction with a family of membrane-associated guanylate kinases. Nature **378:**85-8.
36. **Kopczynski, C. C., J. N. Noordermeer, T. L. Serano, W. Y. Chen, J. D. Pendleton, S. Lewis, C. S. Goodman, and G. M. Rubin.** 1998. A high throughput screen to identify secreted and transmembrane proteins involved in *Drosophila* embryogenesis. Proc Natl Acad Sci U S A **95:**9973-8.
37. **Kopec, J., A. Bergmann, G. Fritz, E. Grohmann, and W. Keller.** 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:**401-9.
38. **Kostichka, K., L. Tao, M. Bramucci, J. F. Tomb, V. Nagarajan, and Q. Cheng.** 2003. A small cryptic plasmid from *Rhodococcus erythropolis*: characterization and utility for gene expression. Appl Microbiol Biotechnol **62:**61-8.
39. **Kurenbach, B., D. Grothe, M. E. Farias, U. Szewzyk, and E. Grohmann.** 2002. The *tra* region of the conjugative plasmid pIP501 is organized in an operon with the first gene encoding the relaxase. J Bacteriol **184:**1801-5.
40. **Lanka, E., and B. M. Wilkins.** 1995. DNA processing reactions in bacterial conjugation. Annu Rev Biochem **64:**141-69.

41. **Larkin, C., S. Datta, M. J. Harley, B. J. Anderson, A. Ebie, V. Hargreaves, and J. F. Schildbach.** 2005. Inter- and intramolecular determinants of the specificity of single-stranded DNA binding and cleavage by the F factor relaxase. Structure **13:**1533-44.
42. **Lessard, P. A., X. M. O'Brien, D. H. Currie, and A. J. Sinskey.** 2004. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. BMC Microbiol **4:**15.
43. **Luo, L., Y. J. Liao, L. Y. Jan, and Y. N. Jan.** 1994. Distinct morphogenetic functions of similar small GTPases: *Drosophila* Drac1 is involved in axonal outgrowth and myoblast fusion. Genes Dev **8:**1787-802.
44. **Marchler-Bauer, A., and S. H. Bryant.** 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res **32:**W327-31.
45. **Myers, E. W., G. G. Sutton, A. L. Delcher, I. M. Dew, D. P. Fasulo, M. J. Flanigan, S. A. Kravitz, C. M. Mobarry, K. H. Reinert, K. A. Remington, E. L. Anson, R. A. Bolanos, H. H. Chou, C. M. Jordan, A. L. Halpern, S. Lonardi, E. M. Beasley, R. C. Brandon, L. Chen, P. J. Dunn, Z. Lai, Y. Liang, D. R. Nusskern, M. Zhan, Q. Zhang, X. Zheng, G. M. Rubin, M. D. Adams, and J. C. Venter.** 2000. A whole-genome assembly of *Drosophila*. Science **287:**2196-204.
46. **Orr-Weaver, T. L., J. W. Szostak, and R. J. Rothstein.** 1981. Yeast transformation: a model system for the study of recombination. Proc Natl Acad Sci U S A **78:**6354-8.
47. **Pansegrau, W., W. Schroder, and E. Lanka.** 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:**2782-9.
48. **Patrauchan, M. A., C. Florizone, M. Dosanjh, W. W. Mohn, J. Davies, and L. D. Eltis.** 2005. Catabolism of benzoate and phthalate in *Rhodococcus* sp. strain RHA1: redundancies and convergence. J Bacteriol **187:**4050-63.
49. **Plaggenborg, R., J. Overhage, A. Loos, J. A. Archer, P. Lessard, A. J. Sinskey, A. Steinbuchel, and H. Priefert.** 2006. Potential of *Rhodococcus* strains for biotechnological vanillin production from ferulic acid and eugenol. Appl Microbiol Biotechnol:1-11.
50. **Resnick, M. A., and P. Martin.** 1976. The repair of double-strand breaks in the nuclear DNA of *Saccharomyces cerevisiae* and its genetic control. Mol Gen Genet **143:**119-29.
51. **Rong, Y. S., and K. G. Golic.** 2000. Gene targeting by homologous recombination in *Drosophila*. Science **288:**2013-8.
52. **Rouet, P., F. Smih, and M. Jasin.** 1994. Expression of a site-specific endonuclease stimulates homologous recombination in mammalian cells. Proc Natl Acad Sci U S A **91:**6064-8.
53. **Rouet, P., F. Smih, and M. Jasin.** 1994. Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. Mol Cell Biol **14:**8096-106.
54. **Sajjaphan, K., N. Shapir, L. P. Wackett, M. Palmer, B. Blackmon, J. Tomkins, and M. J. Sadowsky.** 2004. Arthrobacter aurescens 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:**4402-7.
55. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
56. **Scherer, S., and R. W. Davis.** 1979. Replacement of chromosome segments with altered DNA sequences constructed in vitro. Proc Natl Acad Sci U S A **76:**4951-5.
57. **Scherzinger, E., V. Kruft, and S. Otto.** 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:**929-38.
58. **Sears, H. C., C. J. Kennedy, and P. A. Garrity.** 2003. Macrophage-mediated corpse engulfment is required for normal *Drosophila* CNS morphogenesis. Development **130:**3557-65.
59. **Sekine, M., S. Tanikawa, S. Omata, M. Saito, T. Fujisawa, N. Tsukatani, T. Tajima, T. Sekigawa, H. Kosugi, Y. Matsuo, R. Nishiko, K. Imamura, M. Ito, H. Narita, S. Tago, N. Fujita, and S. Harayama.** 2006. Sequence analysis of three plasmids harboured in *Rhodococcus erythropolis* strain PR4. Environ Microbiol **8:**334-46.
60. **Simon, R., U. Priefer, and A. Puehler.** 1983. A broad host range mobilization system for *in vivo* genetic engineering: transposon mutagensis in gram negative bacteria. Bio/Technology **1:**784-791.
61. **Smih, F., P. Rouet, P. J. Romanienko, and M. Jasin.** 1995. Double-strand breaks at the target locus stimulate gene targeting in embryonic stem cells. Nucleic Acids Res **23:**5012-9.
62. **Smith, G. R.** 1989. Homologous recombination in prokaryotes: enzymes and controlling sites. Genome **31:**520-7.
63. **Spradling, A. C., and G. M. Rubin.** 1982. Transposition of cloned P elements into Drosophila germ line chromosomes. Science **218:**341-7.
64. **Stecker, C., A. Johann, C. Herzberg, B. Averhoff, and G. Gottschalk.** 2003. Complete nucleotide sequence and genetic organization of the 210-kilobase linear plasmid of *Rhodococcus erythropolis* BD2. J Bacteriol **185:**5269-74.
65. **Street, L. M., M. J. Harley, J. C. Stern, C. Larkin, S. L. Williams, D. L. Miller, J. A. Dohm, M. E. Rodgers, and J. F. Schildbach.** 2003. Subdomain organization and catalytic residues of the F factor TraI relaxase domain. Biochim Biophys Acta **1646:**86-99.
66. **Tao, L., and Q. Cheng.** 2004. Novel beta-carotene ketolases from non-photosynthetic bacteria for canthaxanthin synthesis. Mol Genet Genomics **272:**530-7.
67. **Thomas, K. R., and M. R. Capecchi.** 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell **51:**503-12.
68. **van den Bosch, M., P. H. Lohman, and A. Pastink.** 2002. DNA double-strand break repair by homologous recombination. Biol Chem **383:**873-92.
69. **van der Geize, R., G. I. Hessels, R. van Gerwen, P. van der Meijden, and L. Dijkhuizen.** 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.
70. **van Ham, M., and W. Hendriks.** 2003. PDZ domains-glue and guide. Mol Biol

Rep **30:**69-82.

71. **Voeykova, T., L. Emelyanova, V. Tabakov, and N. Mkrtumyan.** 1998. Transfer of plasmid pTO1 from *Escherichia coli* to various representatives of the order *Actinomycetales* by intergeneric conjugation. FEMS Microbiol Lett **162:**47-52.
72. **Waltermann, M., H. Luftmann, D. Baumeister, R. Kalscheuer, and A. Steinbuchel.** 2000. *Rhodococcus opacus* strain PD630 as a new source of high-value single-cell oil? Isolation and characterization of triacylglycerols and other storage lipids. Microbiology **146 (Pt 5):**1143-9.
73. **Waterston, R. H., K. Lindblad-Toh, E. Birney, J. Rogers, J. F. Abril, P. Agarwal, R. Agarwala, R. Ainscough, M. Alexandersson, P. An, S. E. Antonarakis, J. Attwood, R. Baertsch, J. Bailey, K. Barlow, S. Beck, E. Berry, B. Birren, T. Bloom, P. Bork, M. Botcherby, N. Bray, M. R. Brent, D. G. Brown, S. D. Brown, C. Bult, J. Burton, J. Butler, R. D. Campbell, P. Carninci, S. Cawley, F. Chiaromonte, A. T. Chinwalla, D. M. Church, M. Clamp, C. Clee, F. S. Collins, L. L. Cook, R. R. Copley, A. Coulson, O. Couronne, J. Cuff, V. Curwen, T. Cutts, M. Daly, R. David, J. Davies, K. D. Delehaunty, J. Deri, E. T. Dermitzakis, C. Dewey, N. J. Dickens, M. Diekhans, S. Dodge, I. Dubchak, D. M. Dunn, S. R. Eddy, L. Elnitski, R. D. Emes, P. Eswara, E. Eyras, A. Felsenfeld, G. A. Fewell, P. Flicek, K. Foley, W. N. Frankel, L. A. Fulton, R. S. Fulton, T. S. Furey, D. Gage, R. A. Gibbs, G. Glusman, S. Gnerre, N. Goldman, L. Goodstadt, D. Grafham, T. A. Graves, E. D. Green, S. Gregory, R. Guigo, M. Guyer, R. C. Hardison, D. Haussler, Y. Hayashizaki, L. W. Hillier, A. Hinrichs, W. Hlavina, T. Holzer, F. Hsu, A. Hua, T. Hubbard, A. Hunt, I. Jackson, D. B. Jaffe, L. S. Johnson, M. Jones, T. A. Jones, A. Joy, M. Kamal, E. K. Karlsson, et al.** 2002. Initial sequencing and comparative analysis of the mouse genome. Nature **420:**520-62.
74. **Winzeler, E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J. D. Boeke, H. Bussey, A. M. Chu, C. Connelly, K. Davis, F. Dietrich, S. W. Dow, M. El Bakkoury, F. Foury, S. H. Friend, E. Gentalen, G. Giaever, J. H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D. J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, C. Rebischung, J. L. Revuelta, L. Riles, C. J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R. K. Storms, S. Veronneau, M. Voet, G. Volckaert, T. R. Ward, R. Wysocki, G. S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston, and R. W. Davis.** 1999. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. Science **285:**901-6.
75. **Wodarz, A., U. Hinz, M. Engelbert, and E. Knust.** 1995. Expression of *crumbs* confers apical character on plasma membrane domains of ectodermal epithelia of *Drosophila*. Cell **82:**67-76.
76. **Yang, J. C., P. A. Lessard, N. Sengupta, S. D. Windsor, X. M. O'Brien, M. Bramucci, V. Nagarajan, and A. J. Sinskey.** submitted. TraA is required for megaplasmid conjugation in *Rhodococcus erythropolis* AN12.

CHAPTER 4.

Characterization of the conjugation determinants of pAN12, a small replicon from *Rhodococcus erythropolis* AN12.*

*author's note:
portions of this chapter has been submitted
as a manuscript under the same title, 2006

4.1. Introduction.

Rhodococcus sp. bacteria are Gram-positive actinomycetes found ubiquitously in many terrestrial and marine environments. This versatility can partially be attributed to plasmid-encoded operons. In particular, megaplasmids found in these rhodococci often encode advantageous properties, such as the degradation of aromatic compounds (15), and virulence factors required for plant pathogenesis (4). Besides megaplasmids, which can range from 50 kb to greater than 1 megabasepairs in size, and can exhibit either linear and circular topologies, rhodococci may also harbor circular plasmids of smaller sizes. Circular plasmids are classified according to type of replication they undergo. Three general mechanisms of plasmid replication have been characterized; theta, strand displacement, and rolling circle (6). A number of smaller plasmids proposed to replicate via either the theta type (5, 11, 17, 19) or the rolling circle type (16, 21) of replication have been discovered in rhodococci. These *Rhodococcus* plasmids are cryptic, and have been studied mainly for use as *Escherichia coli-Rhodococcus* shuttle vectors. Several of these plasmids have proven to be useful for metabolic engineering purposes (11, 16).

Many species of *Rhodococcus* can exchange plasmids via conjugation. It has been proposed that mechanisms similar to those demonstrated for Gram-negative bacterial plasmids function in Gram-positive plasmid conjugation (10). These processes are initiated with proteins called relaxases or nickases that act *in trans* upon plasmid sequences required *in cis* called origins of transfer (*ori*T) , within which reside specific cleavage sequences called *nic* sites (18). It is assumed for these plasmids that a single-stranded DNA intermediate is transferred from donor to recipient cells. Interestingly, this is apparently not the case for a subset of plasmids found in mycelium-forming actinomycetes. For the pIJ101 conjugative plasmid of *Streptomyces lividans*, the transfer of double-stranded DNA (28) is mediated by a single plasmid-borne and membrane-bound gene product called Tra (25-27). Tra shares protein similarity with the Gram-negative FtsK ATPase, which functions in DNA segregation during cell division.

Though the exchange of genetic information has been documented for rhodococci as early as the 1960's (3), few reports exist that describe determinants governing plasmid conjugation for these actinomycetes. It is unknown whether the mechanisms in rhodococci are more analogous to single-stranded DNA transfer typified by Gram-

negative systems, or more closely resemble those observed for the *Streptomyces* pIJ101. Further understanding of the mechanisms of *Rhodococcus* plasmid conjugation will also facilitate its genetic analysis and manipulation. Since direct rhodococci transformations can prove difficult due to low transformation efficiency and/or DNA rearrangements, conjugation can be a powerful tool to introduce DNA in the context of a plasmid that is stably maintained by both the donor and recipient strains. Recently, the first conjugative small (<10 kb) *Rhodococcus* plasmid, pB264, was described in *Rhodococcus sp.* B264-1 (19). A 700 bp region required *in cis* for its transfer has been defined. The presence of seven inverted repeats and one direct repeat within this region suggests that one or more of these elements are substrates for the plasmid transfer mechanism. It is unknown for pB264 what protein factors are required *in trans* for its transfer, although it has been proposed that those factors may be encoded by genes which reside on B264-1 megaplasmids (19).

The goals for this study are to characterize the conjugal determinants of a small circular plasmid, pAN12, found in *R. erythropolis* strain, AN12. AN12 was first isolated and so named for its ability to use the aromatic compound, aniline, as a carbon source (2). The isolation of its cryptic plasmid, pAN12, and characterization of two plasmid-borne genes, *rep* and *div* (involved in plasmid replication and segregation during cell division, respectively), was recently reported (16). pAN12 was assigned to the pIJ101/pJV1 family of rolling-circle family of plasmid based on conservations of Rep protein motifs and sequence similarities of the plasmid replication origins. Though streptomycete plasmids pIJ101 and pJV1 have both been found to be conjugative (14, 33), it is not known whether pAN12 is conjugative, and if so, what are the determinants for its transfer.

In this report, we provide the first evidence of pAN12 conjugation and describe the identification of a minimal 61 bp *cis*-acting region on pAN12 necessary for its mobilization. Furthermore, we show that pAN12 does not encode all necessary factors for its own transmission. In fact, data suggests that the presence of the recently identified AN12 megaplasmid, pREA400, in the donor strain is required for conjugation of pAN12 and derivative plasmids.

4.2. Materials and methods.

4.2.1. Bacterial strains and culturing conditions. Bacterial strains used in the present study are summarized below in Table 4.1. Both *Rhodococcus erythropolis* AN12 and SQ1, as well as strains of *Escherichia coli*, were grown in LB liquid media or on LB plates with 2% agar (29) 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* and *R. ruber* DDO319 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.

4.2.2. Electrocompetent AN12 cells. AN12 electrocompetent cells were prepared essentially as previously described (16), except cells were grown in NBYE/0.05% Tween-80 media in a 1L baffled flask with shaking until $O.D._{600}$ of about 0.5 was reached.

4.2.3. *Rhodococcus* conjugation and plasmid conjugation efficiency. Matings using 100 μl were carried out on solid LB agar surfaces as previously described (19). The plasmid conjugation efficiencies were determined by plating appropriate serial dilutions of the mating cell resuspensions (consisting of 1:1 mixtures of donor cells and recipient cells) onto selective media, and plating control recipient cell dilutions to obtain recipient viable counts on non-selective media. Mating efficiencies were calculated as transconjugant CFU per recipient viable cell counts.

Table 4.1. Summary of bacterial strains used in this study.

Strain name	Description	Source or reference
DH5α	*Escherichia coli*; *supE*44 Δ*lacU*169 (ϕ80 *lacZ*ΔM15) *recA*1 *endA*1 *hsdR*17 *thi*-1 *gyrA*96 *relA*1	Invitrogen
EC100D *pir*-116	*E. coli*; F⁻ *mcr*A Δ(*mrr-hsd*RMS-*mcr*BC) φ80d*lacZ*ΔM15 Δ*lacX*74 *rec*A1 *end*A1 *ara*D139 Δ(*ara, leu*)7697 *gal*U *gal*K λ⁻ *rps*L *nup*G *pir*-116(DHFR)	Epicentre
TOP10	*E coli*; F- *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Φ80*lacZ*ΔM15 Δ*lacX*74 *recA*1 *deoR araD*139 Δ(*ara-leu*)7697 *galU galK rpsL* (Str^r) *endA*1 *nupG*	Invitrogen
AN12	environmental isolate of *Rhodococcus erythropolis*	(16)
SQ1	environmental isolate of *R. erythropolis*; Rf^R Sm^R	ATCC4277-1
DDO319	environmental isolate of *Rhodococcus ruber*	Gift from DuPont de Nemours
AN12PL	flaky derivative strain of AN12 missing pREA100	(34)
AN12PL-1F6	*R. erythropolis* AN12PL mutant recovered from transformation with EZTn transposome; (pREA400::EZTn); Km^R	(34)
AN12-5A6	*R. erythropolis* AN12 derivative recovered from transformation with EZTn transposome; (pREA250::EZTn); Km^R	(34)
JY524	*R. erythropolis* SQ1 derived transconjugant recovered from mating with AN12PL-1F6; (pREA400::EZTn); Km^R Rf^R Sm^R	(34)
JY650	*R. erythropolis* AN12 (pAL319); Gm^R	This study
JY660	*R. erythropolis* AN12 (pAL281); Gm^R	This study
JY671	*R. erythropolis* SQ1 (pAL319); Gm^R	This study
JY700	*R. erythropolis* AN12 (pAL321); Gm^R	This study
JY709	*R. erythropolis* SQ1 (pAL321); Gm^R	This study
JY720	*R. erythropolis* AN12 (pJY29); Gm^R	This study
JY735	*R. erythropolis* AN12 (pJY30); Gm^R	This study
JY737	*R. erythropolis* AN12 (pJY31); Gm^R	This study
JY750	*R. erythropolis* AN12 (pJY33); Gm^R	This study
JY813	*R. erythropolis* AN12(pJY35); Gm^R	This study
JY825	*R. erythropolis* AN12 *traA*:pJY37 Km^R	(34)
JY855	*R. erythropolis* AN12 *traA*::pJY37 transformed with pAL321; Gm^R Km^R	(34); This study

4.2.4. DNA manipulation and plasmid construction. Plasmids used in the present study, as well as a brief description of the cloning strategies used for the construction of each, are summarized below in Table 4.2. 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), and are summarized below in Table 4.3. *Rhodococcus* genomic DNA was prepared as previously described (19). *Rhodococcus* plasmid DNA was isolated using the procedures previously described (16).

Table 4.2. Summary of plasmids and cloning strategies used in this study.

Plasmids	Description	Source or reference
pREA400	endogenous AN12 megaplasmid	(34)
pREA250	endogenous AN12 megaplasmid	(34)
pREA100	endogenous AN12 megaplasmid	(34)
pAL281	*E. coli*-rhodococci shuttle vector; Gm^R	(19)
pAL319	pAL281 derivative carrying the intact pAN12 plasmid ligated at the *Pst*I site	This study
pAL321	pAL319 derivative with the pAN12 *rep* region deleted from 3413 to 4634 bp	This study
pJY29	Derivative of pAL321 carrying a deletion from 758 to 1819 bp between *Xmn*I sites	This study
pJY30	Derivative of pAL321 carrying a deletion from 1253 to 3413 bp between *Cla*I sites	This study
pJY31	Derivative of pAL321 carrying a deletion from 2559 to 4435 bp between *Sal*I sites	This study
pJY33	Derivative of pAL281 constructed by ligating the 740 bp from *XmnI* and *SalI* sites of pAL321 to unique *Eco*RV and *Xho*I sites of pAL281, respectively	This study
pJY35	Derivative of pAL281 constructed by ligating annealed complementary primers JYP573 and JYP574 into *Eco*RI and *Xho*I	This study

4.2.5. Pulsed-field gel electrophoresis (PFGE). PFG electrophoresis was carried out as previously described (34).

4.2.6. 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 0.25M HCl for 40 minute, denaturation with 0.5N NaOH for 30 min, and neutralization with Tris-Cl for 30 min. DNA was then transferred to positively-charged nylon membranes (Roche Diagnostics Corp., Indianapolis, IN) using 20X SSC as the transfer buffer for 24 h. DIG-11 dUTP labeled probes were generated and hybridizations were carried out using reagents in the DIG-High Prime DNA Labeling and Detection Starter Kit per manufacturer's instruction (Roche Diagnostics Corp.). Hybridizing species were detected using Kodak Biomax X-AR or Biomax Light scientific imaging films.

4.2.7. *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 HotStar Taq Polymerase Kit (Qiagen Sciences, Valencia, CA). DNA was amplified in a PTC-200 Cycler (Bio-Rad, Waltham, MA) using cycling parameters of an initial heating step 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 , ending with an extension cycle of 10 min at 72°C. Annealing temperatures and extension times were adjusted to optimize performance of each primer pair.

Table 4.3. Summary of primers used in this study.

Primers	Description	Source or reference
JYP516	forward primer to amplify AN12 chromosomal sequences; 5' GCTACTCATGCCTGCATTCTC 3'	This study
JYP517	reverse primer to amplify AN12 chromosomal sequences; 5' AGCGAAAGCGAGTCCGAATAG 3'	This study
JYP558	forward primer to amplify pAN12 sequences; 5' CTGATCTGCTGGTCAGTGCGG 3'	This study
JYP559	forward primer to amplify pAN12 sequences; 5' CGACCTTGGTGCACTAGTCGG 3'	This study
JYP573	forward primer to clone the pAN12 *clt*-like region; includes *Eco*RI linker; PAGE purified 5'Phos/AATTCCGTGGGCAGGTTTCGGCGTGAGGCGAGTTTTTCTCC TGCCTCATGTGCAACCTCCTCAAA 3'	This study
JYP574	reverse primer to clone the pAN12 *clt*-like region; includes *Xho*I linker; PAGE purified 5'Phos/TCGATTTGAGAAGGTTGCACATGAGGCAGGAGAAAAACTC GCCTCACGCCGAAACCTGCCCACGG 3'	This study

4.3. Results.

4.3.1. Discovery of pAN12 transmission, and definition of the *cis*-acting region required for conjugation. It was shown previously that *Rhodococcus erythropolis* AN12 harbors a 6.3 kb cryptic plasmid called pAN12 (16). Recently, three novel *R. erythropolis* AN12 megaplasmids, pREA400, pREA250, and pREA100 were also discovered using refined pulsed-field gel (PFG) separation techniques. Using a transposon mediated strategy to tag AN12 extrachromosomal replicons, we discovered that two of the three megaplasmid, pREA400 and pREA250, can be transferred to and maintained by a recipient strain of *R. erythropolis* called SQ1. In these experiments, SQ1 derived transconjugants could be distinguished from AN12 donor cells since SQ1 is naturally resistant to rifampicin (Rf) and streptomycin (Sm) antibiotics (19). In the course of characterizing pREA400 conjugation by PFG electrophoresis, we noticed the co-transfer of a smaller replicon, which we suspected was pAN12 (Fig. 4.1). Two additional lines of evidence confirmed that indeed, pAN12 can be mobilized. First, the plasmid can be isolated from wild type AN12 and SQ1(pREA400:Tn) transconjugants,

but not from the wild type SQ1 strain. Second, the extracted plasmids from both wild type AN12 and SQ1(pREA400:Tn) yielded the same restriction enzyme digests patterns, the band sizes of which could be predicted using the published pAN12 sequence.

A series of pAN12-derived plasmids were then constructed to further define the *cis*-acting region required for its mobilization. Because of its cryptic nature, pAN12 was fused with a plasmid (pAL281) encoding gentamicin resistance (Gm^R) to monitor its mobilization. pAL281 is capable of replication in both *E. coli* and *Rhodococcus sp.* SQ1 (19) and AN12 (this work). To verify that pAL281 is not conjugative by itself, AN12(pAL281), or JY660, was mated to SQ1, and transconjugants selected for using triple (Rf Sm Gm) drug selection. We failed to detect any transconjugants (Table 2), and interpret this data to mean that pAL281 cannot be mobilized detectably from AN12 to SQ1. Thus, pAL281 should be useful to analyze what sequences contributed by pAN12 in the context of a fusion plasmid are important for pAN12 conjugation.

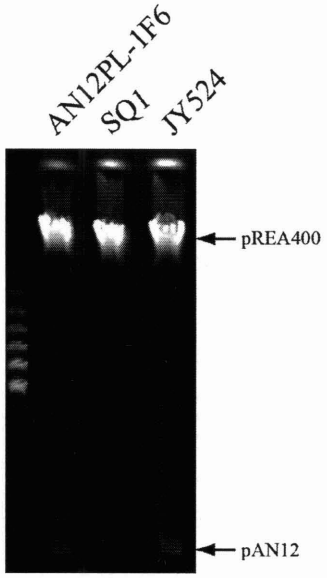


Figure 4.1. PFG of *Rhodococcus* replicons following conjugation. Lanes consist of DNA replicons from donor (AN12PL-1F6), recipient (SQ1), and the SQ1 derived transconjugant (JY524). The lambda PFG ladder was used to estimate sizes of replicons. Arrows point to replicons (pREA400 and pAN12) gained by *R. erythropolis* SQ1 following conjugation. Although only the transfer of megaplasmid pREA400:Tn by virtue of the Km^R marker it carries was selected, pAN12 was transferred as well.

The mobilization of the resulting pAN12-pAL281 fusion plasmid, pAL319, could be monitored based on the gain of Gm^R in recipient cells Because pAL319 might be unstable due to the presence of two functional plasmid replication systems (both NG2 and pAN12 *rep* regions), a second plasmid, pAL321, was created that deletes the pAN12 *rep* ORF. Representative plasmid maps of pAN12, pAL281, pAL319 and pAL321 are shown in Fig. 4.2. Intact pAL319 plasmid can be isolated from transformed AN12 cells, suggesting that it is maintained stably, however the endogenous pAN12 plasmid is frequently lost (data not shown), suggesting some level of incompatibility between pAL319 and pAN12. In contrast, pAL321 is maintained stably by AN12 along with the endogenous pAN12.

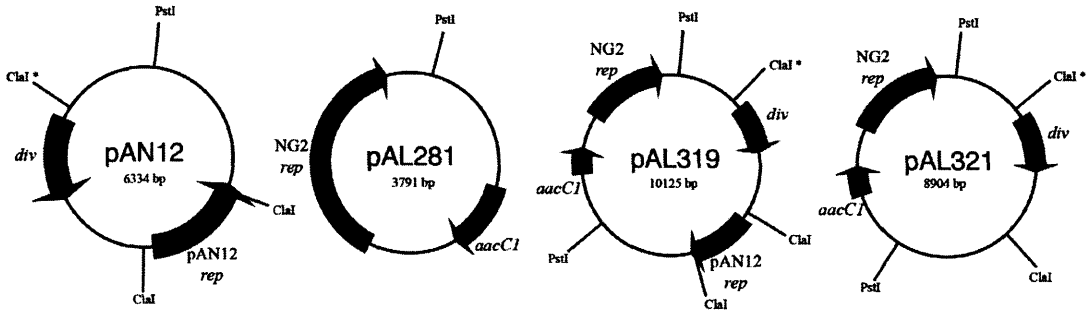


Figure 4.2. Maps of relevant plasmid constructs. Plasmids are not drawn to scale. The endogenous pAN12 plasmid was fused to pAL281 at an unique *Pst*I site, resulting in the plasmid pAL319. The pAL321 plasmid was constructed by digesting pAL319 with *Cla*I to remove the pAN12 *rep* ORF then re-ligating the major product. A *dam* methylated *Cla*I site (*) is labeled.

AN12 carrying either pAL319 or pAL321 were mated to SQ1. The appearance of roughly equal numbers of triply (Rf Sm Gm) resistant colonies strongly suggested that AN12 can mobilize both pAL319 and pAL321 at comparable frequencies, and that the pAN12 *rep* region is dispensable *in cis* for plasmid transfer. The stable maintenance of pAL321 in SQ1 was confirmed with a Southern blot using the entire pAL319 plasmid as a probe template and total genomic DNA samples from AN12, donor JY700, recipient SQ1, and transconjugant JY709 (Fig. 4.3). The endogenous pAN12 should be linearized with either the *Eco*RI and *Bgl*II double enzyme digest, or the *Pst*I single enzyme digest, and should also be detected using the pAL319 probe. Hybridizing bands of the expected sizes from each sample were observed, strongly suggesting that the pAL321 plasmid is

indeed conjugative, and that the SQ1 derived transconjugant JY709 stably maintains pAL321 . Interestingly, the endogenous pAN12 plasmid co-mobilized with pAL321 in all cases examined. Furthermore, pAL321 mobilization frequency, 1.5×10^{-5} (Table 4.4), is comparable to what was observed for pREA400 and pREA250 (34).

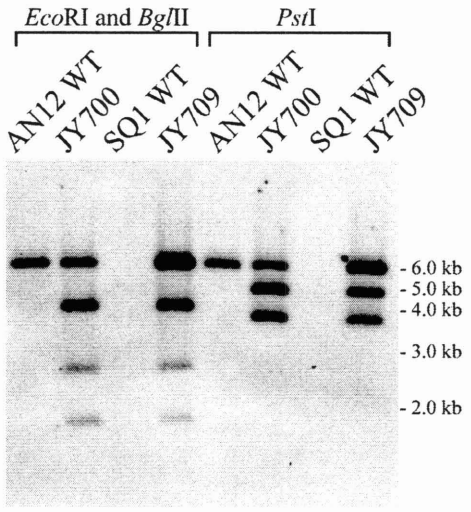


Figure 4.3. Southern blot analysis of pAL321 conjugation and maintenance. Genomic DNA prepared from AN12, pAL321 donor strain (JY700), SQ1, and transconjugant strain JY709, were digested with the indicated enzymes and transferred onto a positively charged nylon membrane. Treatment of pAL321 with *Eco*RI and *Bgl*II should produce 4,427 bp, 2,575 bp, and 1,890 bp products. Treatment of pAL321 with *Pst*I should produce 5,109 bp and 3,787 bp products. Linearized endogenous pAN12 is expected to migrate at 6,334 bp. A digoxygenin (DIG) labeled Southern probe was prepared using the entire pAL319 plasmid as template. Relevant 1kb ladder sizes (NEB) are indicated.

Using AN12(pAL321) as a positive control, and AN12(pAL281) as a negative control, AN12 carrying a series of pAL321 deletion plasmids were assayed for mobilization to SQ1 (Fig. 4.4). In each of these cases, plasmids were re-isolated from transconjugants to verify the transfer to and their subsequent maintenance by SQ1. The mobilization frequency for each pAL321-derivative plasmid was also determined (Table 2). The mobilization frequency for pJY30 (1.4×10^{-9} per recipient cell) is much lower than pAL321, strongly suggesting that sequences deleted between these *Cla*I sites are important for pAN12 mobility. Unlike pJY30, other deletion plasmids (pJY29 and pJY31) transferred to SQ1 at similar frequencies as pAL321. Taken together, data

strongly suggests that the 746 bp region between *Xmn*I and *Sal*I sites of pAN12 (from nt. 4034 to 4770, Genbank accession no. AY178757) is necessary *in cis* for pAN12 mobilization. Indeed, when this sequence was fused with the pAL281 backbone, it is sufficient to confer mobility to the resulting plasmid, pJY33, confirming this hypothesis (Table 4.4).

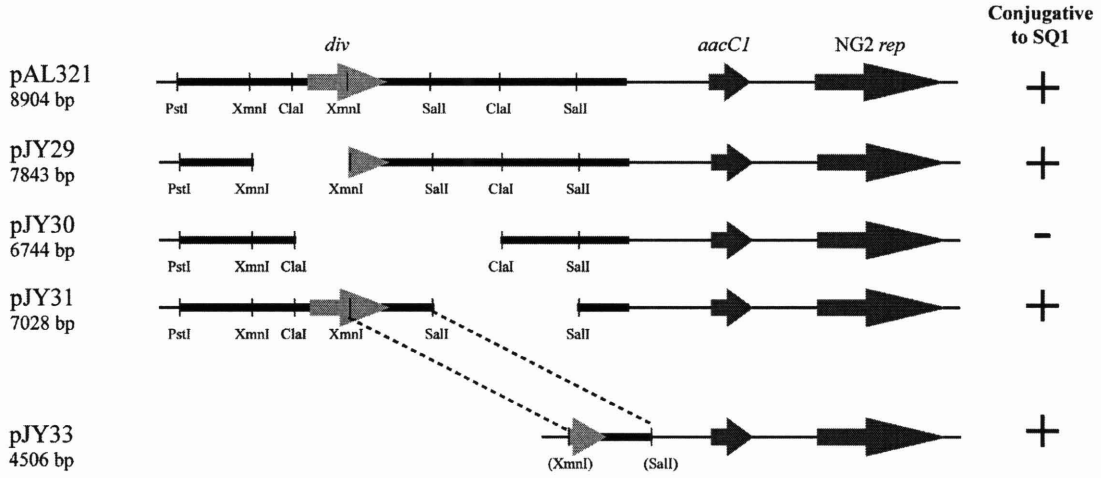


Figure 4.4. Diagram of the pAN12 constructs, and summary of transmissibility to SQ1. pAL281 plasmid sequences are represented as thin black horizontal lines, and pAN12 sequences are represented as thick dark horizontal lines. Relevant restriction enzyme recognition sites are noted with thin black vertical lines. pAL281 ORFs, *aacC1* (encoding Gm^R) and *rep* (encoding plasmid replication function) are indicated with dark gray arrows. The pAN12 ORF *div* is indicated with a light gray arrow. pJY29 through pJY31 are direct deletions of pAL321, and sequences deleted are indicated with gaps. pJY33 was constructed by fusing the indicated *Xmn*I-*Sal*I fragment to pAL281 digested with *Eco*RV and *Xho*I. Whether a plasmid can be mobilized (+) or not (-) to SQ1 is indicated on the right.

Table 4.4. Summary of mobilization efficiencies of pAN12 derived plasmids to SQ1.

Strain (Plasmid)	No. Trials	Weighted Average Mobilization Efficiency [a]	Standard Deviation
JY660 (pAL281)	4	0[b]	0[b]
JY700 (pAL321)	7	1.5×10^{-5}	4.0×10^{-7}
JY720 (pJY29)	6	2.0×10^{-6}	9.8×10^{-7}
JY735 (pJY30)	6	1.4×10^{-9}	3.5×10^{-10}
JY737 (pJY31)	5	1.4×10^{-6}	3.3×10^{-7}
JY750 (pJY33)	4	4.6×10^{-7}	2.3×10^{-8}
JY813 (pJY35)	4	2.4×10^{-7}	8.5×10^{-8}

a. Mobilization efficiencies were calculated by dividing the number of transconjugant CFUs observed on selective plates by the number of recipient CFUs observed on permissive plates.
b. The designation of "0" is defined as less than 1 event per 2.5×10^9 recipient CFU, the lower limit of detection.

4.3.2. Identification and characterization of the pAN12 *clt*-like region.

pAN12 belongs to the pIJ101/pJV1 family of rolling-circle plasmids (16). Protein sequence analysis suggests that not only are the replication proteins and origins of replications similar, the Div and Tra proteins of pAN12 and pIJ101 share core motifs with the FtsK/SpoIIIE family of DNA segregation proteins (9). Recently, a minimal 54 bp sequence (*clt*) in pIJ101 was found to be essential for its conjugation (8). We looked within the pAN12 *Xmn*I-*Sal*I 746bp fragment and asked whether any sequence is similar to the pIJ101 *clt*. Indeed, a 61bp region on pAN12 was found that shares approximately 50% nucleotide identity to the pJI101 *clt* (Fig. 4.5).

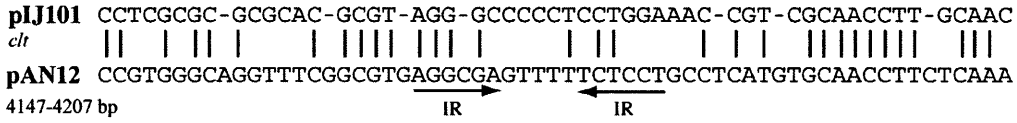
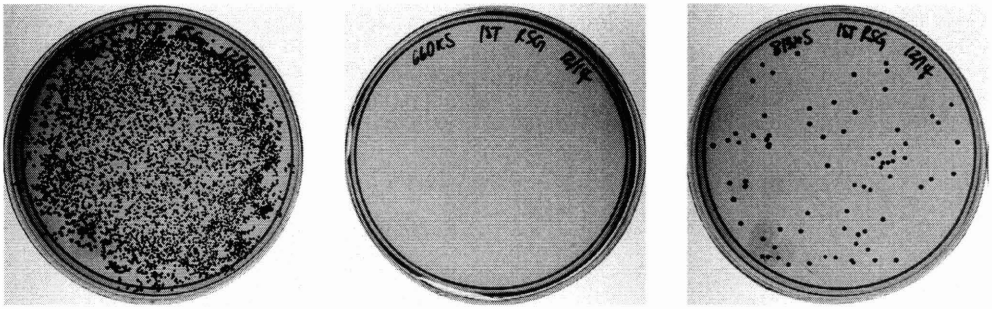


Figure 4.5. Identification of the minimal pAN12 *clt*-like region. The 54 bp *clt* region of pIJ101 (Genbank no. M21778) was aligned with a 61 bp region from nt. 4,147 to 4,207 of pAN12 (Genbank no. AY178757) using a computer-assisted program based on the BLOSUM62 alignment protocol with a medium penalty assessed for gaps (DNA Strider 1.4f6, CEA, France). Identical (32/61) residues are indicated with vertical lines. Seven gaps in pIJ101 are indicated with horizontal dashes. Imperfect inverted repeats (IR) in the pAN12 *clt*-like region are indicated with arrows.

To test the functionality of this sequence, oligonucleotides JYP573 and JYP574 were annealed and ligated into pAL281. The resulting plasmid, pJY35, was used to transform AN12. Gentamicin resistant transformants were analyzed for the presence of the intact plasmid by rescuing the plasmid and sequencing. The confirmed AN12(pJY35) strain, JY813, was then mated to SQ1, as were the AN12(pAL321) strain, JY700, and the AN12(pAL281) strain, JY660, as positive and negative controls, respectively (Fig. 4.6). Clearly, the addition of the pAN12 *clt*-like region confers mobility to the otherwise non-mobile pAL281. Although pJY35 is not transferred with the same efficiency as pAL321, this experiment demonstrated that the pAN12 *clt*-like region supplied the necessary *cis*-acting function to direct pAL281 transfer from AN12 to SQ1.



JY700 x SQ1 JY660 x SQ1 JY813 x SQ1

Figure 4.6. Addition of the pAN12 *clt*-like region confers transmissibility to pAL281. AN12 donor strains carrying either pAL321 (JY700), pAL281 (JY660), or pJY35 (JY813) were mated to SQ1 recipients. A 1:4 dilution of each mating cell suspension (donor X recipient) were plated onto LB agar supplemented with Rf, Sm, and Km. Plates were incubated for 96 hours at 30°C prior to being photographed.

4.3.3. Co-Mobilization of AN12 megaplasmids and pAN12 during conjugation.

Co-mobilization of pAN12 with pREA400:Tn prompted us to begin this investigation of the basis of pAN12 conjugation. Interestingly, the co-mobilization of pAN12 appeared to be less frequent when AN12-5A6 (which carries the megaplasmid pREA250:Tn) was mated to SQ1. We sought to determine an approximate frequency of pAN12 co-mobilization to explore its potential dependence on other plasmids for mobilization. Because PFG electrophoresis is a time-consuming process, a *Rhodococcus erythropolis* colony PCR assay was developed to examine larger sample numbers. The primers JYP558 and JYP559 were made to amplify pAN12-specific sequences. These primers generated a PCR product when AN12 cells were used in the reaction, and did not generate a product when wild-type SQ1 cells were used. In addition, positive control primers, JYP516 and JYP517, which amplify chromosomal sequences common to both AN12 and SQ1, gave expected products in both reactions (Fig. 4.7). This demonstrated that colony PCR is a viable technique to assay for presence of specific elements in both of these *R. erythropolis* strains. A total of 72 SQ1(pREA400:Tn) and 75 SQ1(pREA250:Tn) transconjugants were then tested for the presence of pAN12 using JYP558 and JYP559. pAN12 was co-mobilized in 34% of SQ1(pREA400) transconjugants examined, and only 7% of SQ1(pREA250) transconjugants examined (Fig. 4.8), suggesting that pAN12 co-transfer is common and that its frequency may depend on which megaplasmid is transferred.

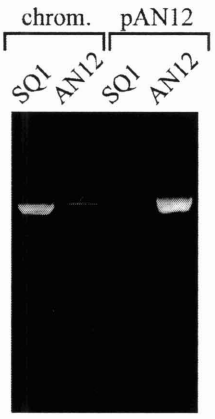


Figure 4.7. Colony PCR and co-mobilization of pAN12 with pREA400 megaplasmids. (A) Amplified sequence from colony PCR using primer pairs JYP516/JYP517, against common chromosomal (chrom.) sequences, and JYP558/JYP559, against pAN12 sequences, were resolved on a 2% agarose gel using the indicated strains as templates.

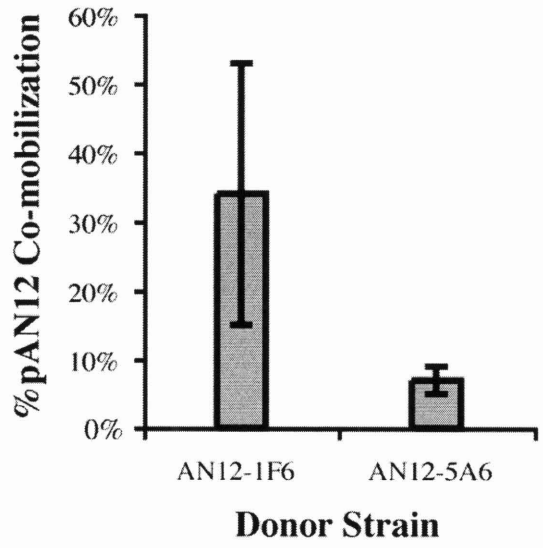


Figure 4.8. Megaplasmid-specific co-mobilization of pAN12. AN12-1F6 and AN12-5A6 were mated to SQ1 cells in three independent trials. Transconjugants that were triply (Rf Sm Km) resistant were tested for the presence of pAN12 using primer JYP558 and JYP559 in colony PCR reactions. The number of pAN12 positive transconjugants was divided by the number of transconjugants to yield a percentage of co-mobilization. The average percentage from all trials is plotted on the Y-axis for each donor (X-axis). Error bars representing the mean plus or minus one standard deviation are shown.

4.3.4. Involvement of pREA400 megaplasmid in pAL321 transfer. One possible explanation for the more frequent co-transfer of pAN12 with pREA400 than pREA250 could be that the larger AN12 megaplasmid is somehow involved in pAN12 transfer. This involvement was not expected for pAN12, as it is quite well documented that pIJ101 encodes all necessary *cis*- and *trans*-acting factors required for its conjugation, in the forms of the *clt* and Tra proteins, respectively. It was not possible to directly assay AN12 strains for the *trans*-acting requirements residing on megaplasmid(s) as no method currently exists to cure specific megaplasmids. Instead, SQ1 strains were generated that stably maintained pAL321 plasmid along with either the tagged pREA400 or the tagged pREA250 megaplasmid. These SQ1 derivative strains are useful as donors as they made it possible to test individual AN12 megaplasmid separately. A potential issue to overcome was that AN12 could not be used as the recipient in the following experiment when SQ1 strains are used as donors, as no selection remains to distinguish AN12-derived transconjugants from SQ1 donors. For this reason, the *Rhodococcus ruber* strain, DDO319, was used as the ultimate recipient since its growth at 37°C could be exploited as a selection criterion. Maintenance of pAL321 by DDO319 was not a concern since the NG2 replicon (supplied in the pAL281 backbone of pAL321) functions at 37°C (19). To eliminate the possibility that pAL281 backbone could confer non-pAN12 related transfer functions, an experiment was done in which AN12(pAL281) was mated to DDO319, and transfer of the gentamicin resistance marker used to monitor pAL281 transfer. No transconjugants were ever detected in these experiments, indicating that the pAL281 backbone would not be involved in the transfer of the pAL321 fusion plasmid between AN12 and DDO319.

The following experiment, depicted in Fig. 4.9, indicates that pREA400 is indeed involved in the transfer of pAN12 and its derivative, pAL321. Either AN12PL-1F6 or AN12-5A6 donors were mated to SQ1(pAL321) recipient cells. Twelve separate transconjugants that were quadruply (Rf, Sm, Gm, Km) resistant from each of these matings were then isolated and used as donors in individual matings with DDO319. Twelve SQ1(pAL321) strains that did not gain either tagged megaplasmid served as negative control. AN12(pAL321) strain, JY700, was always used as a positive control donor in these experiments. Only the transfers of pAL321, and not the tagged

megaplasmids, were selected for in these subsequent matings to DDO319. We found that eleven out of twelve SQ1(pREA400:Tn + pAL321) donors gave rise to DDO319(pAL321) transconjugants. In contrast, none of the twelve SQ1(pREA250:Tn + pAL321) strains and none of the twelve SQ1(pAL321) negative control donor strains tested gave rise to DDO319(pAL321) transconjugants. These results implicate a role for pREA400 encoded factor(s) in pAN12 conjugation. Interestingly, this factor(s) is not the *traA* encoded relaxase/helicase that we found to be indispensable for pREA400 conjugation (34). The strain, JY825, or AN12(pREA400Δ*traA*::Km + pAL321), efficiently transferred pAL321 to both *R. erythropolis* SQ1 and *R. ruber* DDO319 (Fig. 4.10).

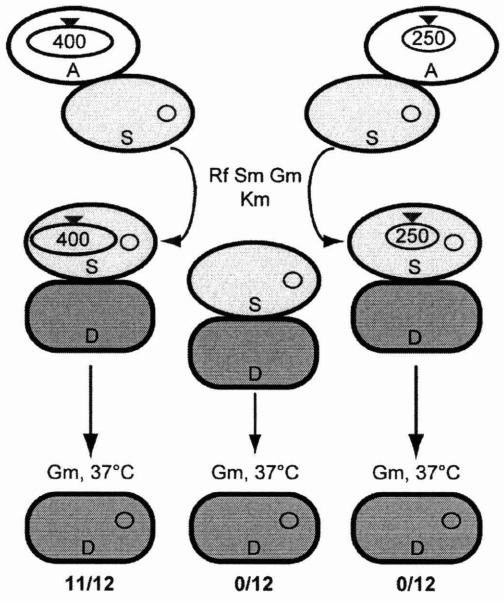
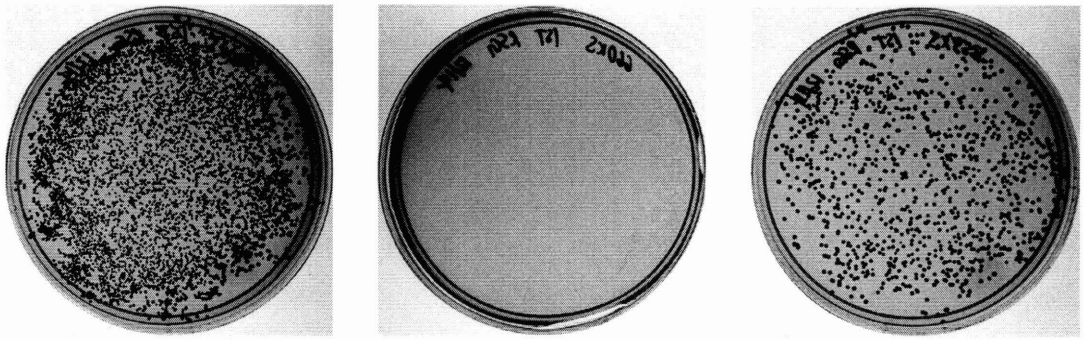


Figure 4.9. Involvement of the *R. erythropolis* AN12 megaplasmid, pREA400, in pAL321 conjugation. *R. erythropolis* strains are represented as large ovals whereas the *Rhodococcus ruber* strains, DDO319, are represented as dark rounded rectangles (D). AN12 derivatives (A, white oval) have been distinguished from SQ1 derivatives (S, light gray oval) by the label and shading. Transposon-tagged megaplasmids are indicated as smaller ovals with inverted triangles. The numbers inside the ovals distinguish between pREA400:Tn (400) and pREA250:Tn (250). The pAN12-derived plasmids, pAL321 (Gm^R), have been indicated as small, unlabelled circles Selections (antibiotics and/or temperature) used at each mating step have been indicated. The numbers of successful matings using individual SQ1 donors of the indicated plasmid composition that produced viable *R. ruber* DDO319(pAL321) transconjugants after selections, versus the total number of such matings, are expressed as fractions.



JY700 x SQ1 JY660 x SQ1 JY855 x SQ1

Figure 4.10. Megaplasmid encoded *traA* function is not required for pAN12 transfer. Donors which are either wildtype (JY700), or mutant (JY855) for the *traA* locus on pREA400 were assayed for the ability to transfer pAL321 to recipient SQ1 strains. JY660, an AN12 strain carrying the non-transmissible pAL281 backbone plasmid was used as a negative control A 1:4 dilution of the indicated mating cell suspension (donor x recipient) was plated onto LB supplemented with Rf, Sm, and Gm antibiotics. Plates were photographed after 5 days of incubation at 30°C.

4.4. Discussion.

We have demonstrated pAN12 mobility, both as the unmodified endogenous plasmid and as fusion plasmids. It was shown previously that pAN12 is related to the *Streptomyces* conjugative plasmid, pIJ101 (16). Recently, a minimal 54 bp region (*clt*) required for pIJ101 conjugation was identified (8). Based on the plasmids' similarities, we looked for a *clt*-like sequence in pAN12 using a computer alignment program. Indeed, we were able to find a 61 bp locus on pAN12 that exhibits up to 50% identity with the pIJ101 *clt.* This level of sequence alignment is surprising and perhaps even more significant given that *clt* regions identified on other related *Streptomyces* plasmids, such as pJV1 and pSN22, exhibit little sequence identity with pIJ101 *clt* (13, 32). These plasmid elements appear to serve analogous *cis*-acting roles, as pAL281 gained pAN12 *clt*-like region and transferability concomitantly. The work with *Streptomyces* pIJ101 has shown that its *clt* region consists of 3 direct repeats of G(A/C)AAC, as well as one imperfect inverted repeat (8). We were unable to detect the same direct repeats within the pAN12 *clt*-like region. However, we were able to locate an imperfect inverted repeat that partially overlapped with the *Streptomyces* pIJ101 *clt* inverted repeat (Fig. 3). The

significance of these repeat regions for both plasmids, specifically whether they serve as binding sites or substrates for DNA transfer factors, remains unknown. Evidence does suggest that these *clt* and *clt*-like regions do not behave as the canonical *oriT* and *nic* sites, as an attempt to detect cleavage at the pIJ101 *clt* site has failed (8).

pAN12 and pIJ101 also encode Div and Tra, respectively. The amino acid sequences of these proteins are similar to the FtsK/SpoIIIE ATP-dependent translocases, which function as DNA pumps in chromosome segregation during cell division (9). Similarities between the two plasmids thus far lead us to speculate that, like pIJ101, pAN12 mobilization may be dependent on its Div protein. The mobilization of the pAL321 deletion plasmid, pJY29, in which the Div protein function was removed, might argue against this proposal. However, it is important to note the endogenous pAN12 plasmid is stably maintained with this plasmid in the donor cell, and its intact Div protein may act *in trans* for the mobilization of pJY29. As we have not found conditions in which pAN12 could be cured or separated from pAL321, we cannot rule out the potential function of pAN12 encoded Div protein to direct plasmid transfer. Intriguingly, data suggests that during sporulation, *Bacillus subtilis* SpoIIIE acts upon a double-stranded (ds) DNA substrate (1, 20, 30). Relatedly, a dsDNA transfer mechanism has also been proposed for *Streptomyces* plasmids pIJ101 (7) and pSAM2 (28) which encode SpoIIIE-like proteins. As pAN12 also encode a SpoIIIE-like protein, it is possible that pAN12 transfers as dsDNA during *Rhodococcus* conjugation.

pAN12 is frequently co-mobilized with the tagged megaplasmid, pREA400 We believe this is the first documentation of plasmid co-mobilization in *Rhodococcus*. Our experiments indicate an important role for pREA400 megaplasmid encoded factor(s) for pAN12 transmission, as only SQ1 cells that harbor the tagged pREA400 megaplasmid are capable of pAN12 transmission to *R. ruber* DDO319. We have ruled out the involvement of megaplasmid-encoded TraA relaxase in pAN12 transfer, since the *traA* mutant is able to transfer pAN12 derivatives to SQ1. Although we do not yet know the identity of these megaplasmid encoded factor(s) that are involved in pAN12 transfer, they likely act at cell boundaries to establish physical contact between the donor and recipient cells. Perhaps these are factors similar to mating pair formation (mpf) components required for the conjugation of many bacterial plasmids (10, 31), or factors similar to the

aggregation substance (AS) required for the transfer of Gram-positive *Enterococcus faecalis* plasmid, pCF10 (12, 23). Sequencing of pREA400 will help address whether these mpf or AS genes are encoded by *R. erythropolis* AN12.

There are apparently at least two different modes of plasmid conjugation in *Rhodococcus erythropolis* AN12. Indeed, protein sequence analysis of the megaplasmid-encoded TraA relaxase (34) indicates that it is most similar to the relaxases required for the "classical" single-stranded DNA (ssDNA) plasmid transfer, whereas the pAN12 *cis*-acting determinant and encoded Div protein are most similar to features of *Streptomyces* plasmids that undergo dsDNA plasmid transfer (28). Interestingly, this is not the first actinomycete identified to date with the coding capacity for both the classical single-stranded and *Streptomyces*-like plasmid transfer systems. Plasmids have been isolated from *Bifidobacteria* with either ORFs encoding a putative relaxase (24) or a putative FtsK/SpoIIIE septal DNA translocator protein (22). Future functional studies of both AN12 *div* and *traA* gene products should shed light on mechanisms underlying actinomycete plasmid conjugation.

4.5. References.

1. **Adams, R. T., and R. G. Wake.** 1980. Highly specific labeling of the *Bacillus subtilis* chromosome terminus. J Bacteriol **143:**1036-8.
2. **Bramucci, M. G., C. M. McCutchen, M. Singh, S. M. Thomas, B. S. Larsen, J. Buckholz, and V. Nagarajan.** 2002. Pure bacterial isolates that convert p-xylene to terephthalic acid. Appl Microbiol Biotechnol **58:**255-9.
3. **Brownell, G. H., and K. L. Kelly.** 1969. Inheritance of mating factors in nocardial recombinants. J Bacteriol **99:**25-36.
4. **Crespi, M., E. Messens, A. B. Caplan, M. van Montagu, and J. Desomer.** 1992. Fasciation induction by the phytopathogen *Rhodococcus fascians* depends upon a linear plasmid encoding a cytokinin synthase gene. Embo J **11:**795-804.
5. **De Mot, R., I. Nagy, A. De Schrijver, P. Pattanapipitpaisal, G. Schoofs, and J. Vanderleyden.** 1997. Structural analysis of the 6 kb cryptic plasmid pFAJ2600 from *Rhodococcus erythropolis* NI86/21 and construction of *Escherichia coli-Rhodococcus* shuttle vectors. Microbiology **143 (Pt 10):**3137-47.
6. **del Solar, G., R. Giraldo, M. J. Ruiz-Echevarria, M. Espinosa, and R. Diaz-Orejas.** 1998. Replication and control of circular bacterial plasmids. Microbiol Mol Biol Rev **62:**434-64.
7. **Ducote, M. J., and G. S. Pettis.** 2005. An *in vivo* assay for conjugation-mediated recombination yields novel results for *Streptomyces* plasmid pIJ101. Plasmid.

8. **Ducote, M. J., S. Prakash, and G. S. Pettis.** 2000. Minimal and contributing sequence determinants of the *cis*-acting locus of transfer (*clt*) of streptomycete plasmid pIJ101 occur within an intrinsically curved plasmid region. J Bacteriol **182:**6834-41.
9. **Errington, J., J. Bath, and L. J. Wu.** 2001. DNA transport in bacteria. Nat Rev Mol Cell Biol **2:**538-45.
10. **Grohmann, E., G. Muth, and M. Espinosa.** 2003. Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev **67:**277-301, table of contents.
11. **Hirasawa, K., Y. Ishii, M. Kobayashi, K. Koizumi, and K. Maruhashi.** 2001. Improvement of desulfurization activity in *Rhodococcus erythropolis* KA2-5-1 by genetic engineering. Biosci Biotechnol Biochem **65:**239-46.
12. **Kao, S. M., S. B. Olmsted, A. S. Viksnins, J. C. Gallo, and G. M. Dunny.** 1991. Molecular and genetic analysis of a region of plasmid pCF10 containing positive control genes and structural genes encoding surface proteins involved in pheromone-inducible conjugation in *Enterococcus faecalis*. J Bacteriol **173:**7650-64.
13. **Kataoka, M., Y. M. Kiyose, Y. Michisuji, T. Horiguchi, T. Seki, and T. Yoshida.** 1994. Complete nucleotide sequence of the *Streptomyces nigrifaciens* plasmid, pSN22: genetic organization and correlation with genetic properties. Plasmid **32:**55-69.
14. **Kieser, T., D. A. Hopwood, H. M. Wright, and C. J. Thompson.** 1982. pIJ101, a multi-copy broad host-range *Streptomyces* plasmid: functional analysis and development of DNA cloning vectors. Mol Gen Genet **185:**223-8.
15. **Konig, C., D. Eulberg, J. Groning, S. Lakner, V. Seibert, S. R. Kaschabek, and M. Schlomann.** 2004. A linear megaplasmid, p1CP, carrying the genes for chlorocatechol catabolism of *Rhodococcus opacus* 1CP. Microbiology **150:**3075-87.
16. **Kostichka, K., L. Tao, M. Bramucci, J. F. Tomb, V. Nagarajan, and Q. Cheng.** 2003. A small cryptic plasmid from *Rhodococcus erythropolis*: characterization and utility for gene expression. Appl Microbiol Biotechnol **62:**61-8.
17. **Kulakov, L. A., M. J. Larkin, and A. N. Kulakova.** 1997. Cryptic plasmid pKA22 isolated from the naphthalene degrading derivative of *Rhodococcus rhodochrous* NCIMB13064. Plasmid **38:**61-9.
18. **Lanka, E., and B. M. Wilkins.** 1995. DNA processing reactions in bacterial conjugation. Annu Rev Biochem **64:**141-69.
19. **Lessard, P. A., X. M. O'Brien, D. H. Currie, and A. J. Sinskey.** 2004. pB264, a small, mobilizable, temperature sensitive plasmid from *Rhodococcus*. BMC Microbiol **4:**15.
20. **Lewis, P. J., S. R. Partridge, and J. Errington.** 1994. Sigma factors, asymmetry, and the determination of cell fate in *Bacillus subtilis*. Proc Natl Acad Sci U S A **91:**3849-53.
21. **Nakashima, N., and T. Tamura.** 2004. Isolation and characterization of a rolling-circle-type plasmid from *Rhodococcus erythropolis* and application of the plasmid to multiple-recombinant-protein expression. Appl Environ Microbiol **70:**5557-68.

22. **O'Riordan, K., and G. F. Fitzgerald.** 1999. Molecular characterisation of a 5.75-kb cryptic plasmid from *Bifidobacterium breve* NCFB 2258 and determination of mode of replication. FEMS Microbiol Lett **174:**285-94.
23. **Olmsted, S. B., S. M. Kao, L. J. van Putte, J. C. Gallo, and G. M. Dunny.** 1991. Role of the pheromone-inducible surface protein Asc10 in mating aggregate formation and conjugal transfer of the *Enterococcus faecalis* plasmid pCF10. J Bacteriol **173:**7665-72.
24. **Park, M. S., D. W. Shin, K. H. Lee, and G. E. Ji.** 1999. Sequence analysis of plasmid pKJ50 from *Bifidobacterium longum.* Microbiology **145 (Pt 3):**585-92.
25. **Pettis, G. S., and S. N. Cohen.** 2000. Mutational analysis of the *tra* locus of the broad-host-range *Streptomyces* plasmid pIJ101. J Bacteriol **182:**4500-4.
26. **Pettis, G. S., and S. N. Cohen.** 1996. Plasmid transfer and expression of the transfer (*tra*) gene product of plasmid pIJ101 are temporally regulated during the *Streptomyces lividans* life cycle. Mol Microbiol **19:**1127-35.
27. **Pettis, G. S., and S. N. Cohen.** 2001. Unraveling the essential role in conjugation of the Tra protein of *Streptomyces lividans* plasmid pIJ101. Antonie Van Leeuwenhoek **79:**247-50.
28. **Possoz, C., C. Ribard, J. Gagnat, J. L. Pernodet, and M. Guerineau.** 2001. The integrative element pSAM2 from *Streptomyces*: kinetics and mode of conjugal transfer. Mol Microbiol **42:**159-66.
29. **Sambrook, J., and D. W. Russell.** 2001. Molecular cloning : a laboratory manual, 3rd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
30. **Sargent, M. G.** 1980. Specific labeling of the *Bacillus subtilis* chromosome terminus. J Bacteriol **143:**1033-5.
31. **Schroder, G., and E. Lanka.** 2005. The mating pair formation system of conjugative plasmids-A versatile secretion machinery for transfer of proteins and DNA. Plasmid **54:**1-25.
32. **Servin-Gonzalez, L.** 1996. Identification and properties of a novel *clt* locus in the *Streptomyces* phaeochromogenes plasmid pJV1. J Bacteriol **178:**4323-6.
33. **Servin-Gonzalez, L., A. I. Sampieri, J. Cabello, L. Galvan, V. Juarez, and C. Castro.** 1995. Sequence and functional analysis of the *Streptomyces phaeochromogenes* plasmid pJV1 reveals a modular organization of *Streptomyces* plasmids that replicate by rolling circle. Microbiology **141 (Pt 10):**2499-510.
34. **Yang, J. C., P. A. Lessard, N. Sengupta, S. D. Windsor, X. M. O'Brien, M. Bramucci, V. Nagarajan, and A. J. Sinskey.** submitted. TraA is required for megaplasmid conjugation in *Rhodococcus erythropolis* AN12.

CHAPTER 5.

Conclusions and recommendations for future work.

5.1. Uncovering additional *Rhodococcus* megaplasmid conjugation determinants.

I have demonstrated that *Rhodococcus erythropolis* AN12 encodes three megaplasmids, two of which (pREA400 and pREA250) are conjugative. Given the large apparent sizes of pREA400 and pREA250 (400 kb and 250 kb, respectively), it is likely these plasmids bear most, if not all, factors needed for self-transmission. *R. erythropolis* pREA400 is likely to be a linear plasmid, and pREA250 is likely to be a circular plasmid. Further understanding of the transfer requirements for either linear or circular *Rhodococcus* megaplasmid will be necessary to achieve the goal of using bacterial conjugation to metabolically engineering *Rhodococcus* strains. For instance, it may be possible to integrate a defined origin or transfer (*ori*T) to transfer an otherwise non-mobilizable megaplasmid with desirable properties. Conversely, understanding of the *trans*-acting factors required for plasmid conjugation could be used to prevent unwanted environmental spreading of plasmid-associated function simply by inactivating these conjugation genes. The specific future work concerning the transmission of these megaplasmids will be outlined below for pREA400, which can be modified as necessary to further characterize pREA250 conjugation.

5.1.1. Defining the pREA400 origin of transfer (*ori*T) region. Sequence and subsequent mutant analysis revealed a megaplasmid-encoded gene product, TraA, essential for pREA400 transfer. TraA is a relaxase/helicase highly similar to the TraI relaxase/helicase of the conjugative *Escherichia coli* F plasmid. The relaxases identified and characterized thus far from conjugative Gram-positive plasmids have shown greater resemblance to the relaxases of the IncP family of plasmids (4, 12), and less similarity to the IncF relaxases, of which TraI is a founding member. *Rhodococcus* pREA400 TraA is unique in its resemblance to F plasmid TraI in that respect. However, since the enzymatic mechanisms of IncP and IncF classes of relaxases are conserved, it is expected that a *nic* site or relaxase substrate is present within a plasmid origin of transfer (*ori*T).

It will be of interest to identify and isolate the pREA400 *ori*T for two reasons. First, identification of the *ori*T sequence will aid in defining the pREA400 TraA enzymatic properties, such as its substrate binding specificity and nuclease kinetics. Artificial plasmids or labeled oligonucleotides bearing the F *ori*T sequence have been invaluable for the characterization of relaxosome formation and assessments of TraI

relaxase residues (5, 10). Second, the pREA400 *ori*T may be used as a bait when immobilized on a DNA affinity matrix to identify other conjugation factor/relaxosome components that may or may not physically associate with pREA400 TraA. For the F plasmid, it has been demonstrated that purified TraY and IHF can bind to *oriT* sequences *in vitro* (8, 11, Howard, 1995 #289). Therefore, protein:*ori*T association may allow for the identification of *Rhodococcus* conjugation factors following DNA affinity purification and subsequent mass spectrometry analysis of binding protein sequence.

The *ori*T's of conjugative plasmids have traditionally been identified by subcloning into a non-mobilizable plasmid, then assaying for the chimeric plasmid to be transferred to recipient cells. The vector pAL281 used previously to identify the *clt*-like region the mobilizable plasmid, pAN12, will be and ideal library backbone to identify the *ori*T of pREA400 as it can 1) replicate in both *E.coli* and *Rhodococcus*, and 2) is not mobilizable by itself. Megaplasmid-specific DNA can be isolated from PFG, by taking the DNA containing PFG slice and migrating the DNA into low melting point agarose. DNA will then be precipitated following β-agarase treatment. The megaplasmid DNA can then be cut with *Xho*I restriction enzyme, estimated to cleave once every 2,000 bp of sequence. This strategy should yield an isolate since *ori*Ts are generally less than 500 bp in size (9). Though this experiment has not been done yet for the purpose of isolating the *ori*T, I have successfully isolated megaplasmid-specific DNA and generated small pREA400- and pREA100-specific libraries (Figure 5.1).

Improvements to the efficiency of library construction, such as the modification to the multi-cloning site (MCS) region to disrupt a *lacZ* reporter gene, will be examined further. Once generated, library plasmids will be transformed into wild type *Rhodococcus erythropolis* AN12, and strains harboring the clones can be identified using the pAL281-encoded gentamycin drug resistance. AN12 library transformants will then be pooled and mated to SQ1. Transconjugants bearing the *ori*T containing library plasmid can then be directly assayed.

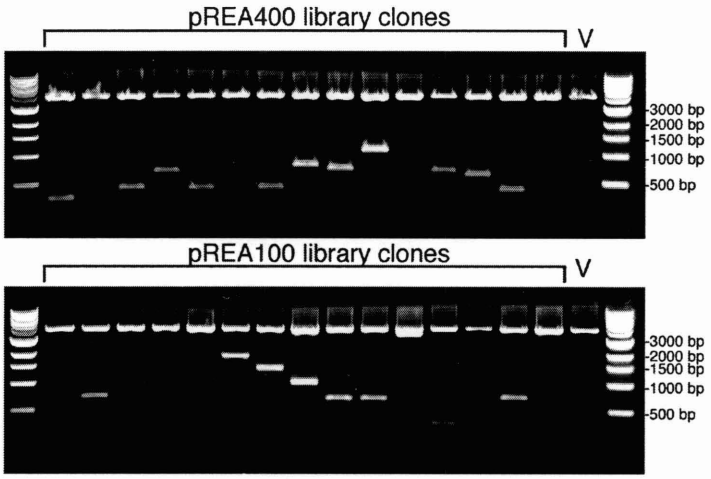


Figure 5.1. Megaplasmids-specific libraries construction. pREA400 and pREA100 DNA were isolated, then digested with *Eco*RV prior to ligating with pCR2.1Topo digested with *Eco*RV and treated with calf-intestinal phosphatase. White colonies were picked for plasmid preparations, and digest with *Eco*RV to estimate the insert size. Vector (V) digested with *Eco*RV were used as a control.

5.1.2. Site-directed mutagenesis of TraA and purification of TraA proteins for enzymatic assays. Conservation of key residues between pREA400 TraA and F plasmid TraI suggests that TraA is the enzymatic component of a nucleoprotein complex required to initiate DNA processing for *Rhodococcus* megaplasmid transfer. Those key residues need to be individually assessed at both genetic and biochemical levels to verify the functional conservation. The genetic approach is straightforward. Site-directed mutagenesis, especially targeting the conserved residues at Y14, H276, and H278, can be performed on the pJY49B containing the wild type copy of pREA400 *traA*. The tyrosine will be substituted with a phenylalanine (Y14F), and the histidine residues will be substituted alanine (H276A and H278A). These *traA* mutations can then be targeted to the JY825 *traA* disruptant to assay for genetic complementation of the pREA400 conjugation defect *in vivo*. Since these mutations lie to the 5' portion of the gene, no modifications to the current complementation strategy will be necessary, although these studies will require careful sequence analysis of the *traA* allele to ascertain these intended mutations are not corrected *in vivo* during homologous recombination.

The wild type and mutated *traA* constructs can also be fused to an epitope element, encoding either 6xHis or Myc, at the carboxy-terminus for protein affinity

purification. Unfortunately, *traA* subcloned in an independently replicating plasmid, such as pAL281, cannot be used *in vivo* in *Rhodococcus* for complementation studies or direct epitope tagging, as this construct is unstable and frequently rearranged (data not shown). This instability may be due to an as yet unidentified incompatibility locus that overlaps the *traA* gene, such as the pREA400 origin of replication or a partitioning gene. It could also be that the unregulated expression of *traA* due to plasmid copy number is toxic to the cells. Thus, I propose to generate the tagged proteins recombinantly in *E. coli*, although theoretically, the wild type TraA fusion protein could also be generated *in vivo* in *Rhodococcus* by using a modified gene targeting scheme.

Once the recombinant protein expression is verified with Western blot analysis, purified tagged traA proteins will be compared to tagged wild type TraA using a circular dichroism spectrophotometer (CD) to detect potential protein misfolding. Since these traA proteins carry point-mutations, it is very unlikely that the overall secondary protein structures will be affected. Purified proteins will then be assessed for binding and nuclease activity on the *ori*T substrate *in vitro*, in the presence or absence of the divalent cation, Mg^{2+}, known to be required cofactors for TraI function (2). All proteins, wild type TraA and mutant traA's, should bind to the *ori*T with equal affinity on gel mobility shift assays using [^{32}P] end-labeled *ori*T substrate. Wild type TraA should cleave the *ori*T in a Mg2+ dependent manner. This nuclease activity can be detected by precipitating the substrate DNA after treatment with TraA, denaturing the DNA with NaOH to separate the complementary strands, then resolving the products on an agarose gel. Site- and strand-specific cleavage should result in an unmodified ssDNA and two smaller ssDNA species, whereas when the DNA is pretreated with mutated traA that cannot cleave its substrate, only co-migrating products should result (3). It is expected that the Y14F mutant will not be able to cleave at the *nic* site since the hydroxyl group necessary for the nucleophilic attack is missing. Likewise, the H276A and H278A mutants should be defective for substrate cleavage, as these histidine residues coordinate Mg^{2+} binding, which aid in nucleophilic attack.

5.1.3. Isolation of additional pREA400 conjugation factors. Sequence up to 30 kb upstream of the *traA* locus will be determined. It is known that the *traY* locus is distal to the *traI* locus in the transfer operon of the F sex factor, so this may not be the

most efficient way to isolate relaxosome components. One method to do so involving DNA affinity matrices has been briefly outlined above. However, direct examination of the ORFs adjacent to the pREA400 *traA* should yield at least some candidates that may be involved in plasmid conjugation, whose functions can immediately be addressed by using the targeted gene-disruption method I've devised. My preliminary results indicate that one such gene, an ORF ca. 3 kb upstream of the *traA* gene, may be an attractive candidate. This 1.9 kb gene, tentatively named *traB*, encodes a 650 amino acid protein with a N-terminal Sir2-like domain (CDD 27970). It's known that a Sir2-like protein called CobB in the bacterium, *Salmonella enterica*, performs deacetylation reactions to modulate the function of a substrate protein, acetyl-CoenzymeA synthetase (13). However, it was recently proposed that not all Sir2-like proteins in bacteria function in protein deacetylation, given that 1) some of these proteins lack the cysteine residues required for deacetylation, and 2) the coding sequence proximity to the genes encoding FtsK/HerA superfamily of ATPases, and the presence of a conserved carboxy-terminal DxH motif, suggested a novel nuclease function for these Sir2-like proteins (6). Although I have not detected this DxH motif in the predicted sequence of *Rhodococcus* TraB, it is intriguing that at least some eukaryotic Sir proteins have been implicated in DNA binding and partitioning (1). It is tempting to suggest that TraB may be the pREA400-encoded factor necessary for pAN12 transfer (see Chapter 4), as the pAN12 FtsK-like Div protein is strikingly similar to the Tra protein involved in *Streptomyces* pIJ101 conjugation, although no evidence yet support this claim. Genetic experiments should clarify whether *traB* is necessary for either pREA400 or pAN12 transmission.

Genes critical for the conjugation of the F-factor, such as *traI*, *traY,* and *traM*, were initially isolated with traditional genetic methods involving mutagenesis. This may also be the means to isolate *tra* factors for pREA400 conjugation in *Rhodococcus*. As mentioned in Chapter 1, either irradiation or chemical mutagenesis is effective to generate *Rhodococcus* mutants. Both methods will be used to mutagenize the transposon-tagged AN12-1F6 starting strain. A screen will be conducted for mutants which can no longer transfer the pREA400-associated kanamycin marker to *Rhodococcus erythropolis* SQ1, outlined in Figure 5.2. These mutant donors may bear mutations in either megaplasmid-encoded genes, or chromosomal genes. The former subset of the

mutants exhibiting the pREA400 conjugation defect should be genetically complemented with a wild type copy of the affected gene, which can be a clone from the same megaplasmid-specific DNA library generated earlier to identify the pREA400 *ori*T. Alternatively, a new genomic library biased for larger-sized insert, and complementation of potential non-lethal chromosomal mutations in genes important for megaplasmid conjugation.

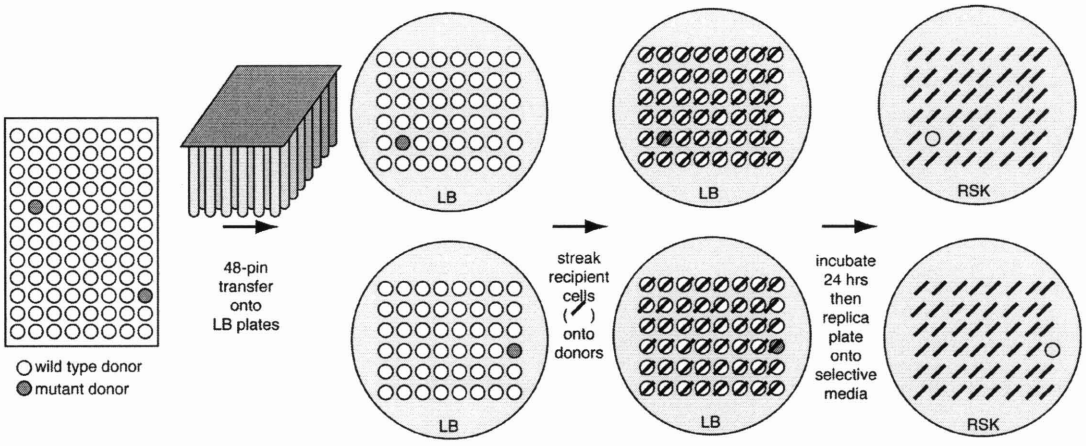
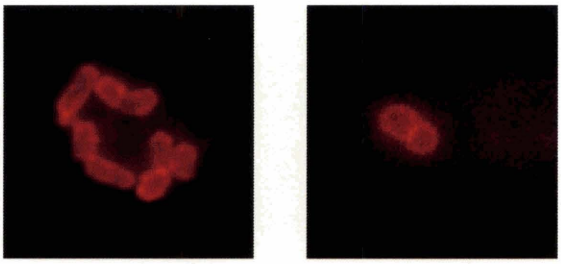


Figure 5.2. High-throughput screening strategy to identify *Rhodococcus* AN12 mutants deficient for pREA400 megaplasmid transfer. *R. erythropolis* AN12PL-1F6 strains will be used for EMS or UV mutagenesis. Mutants will be arrayed onto a 96 well-format in selective media. A 48-pin stamping device will be used to transfer mutant donors onto non-selective plates (LB), then recipients (*R. erythropolis* SQ1) will be streaked across donors. Plates will be incubated at 30°C for 24 hours, then replica-plated onto selective Rf Sm Km (RSK) media. Mutant donors that fail to give rise to SQ1 transconjugants (blank circle) after 4-5 days will be analyzed further.

5.2. Spatial and temporal regulation of *Rhodococcus erythropolis* pAN12 transfer. I have found an unexpected pREA400-specific dependence and co-regulation for the mobility of the cryptic plasmid, pAN12. Specifically, recipient cells selected for pREA400 conjugation are three-times more likely to harbor pAN12 than recipient cells selected for pREA250 conjugation. This is especially intriguing since these two plasmids apparently mobilize in mechanistically different ways; pAN12 does not depend on the TraA relaxase for its transfer and is hypothesized to undergo a dsDNA transfer, while relaxase function is crucial for pREA400 transfer as a ssDNA molecule.

There are three possible and compatible scenarios that could give rise to this phenomenon. First, the cytoplasmic localization of pAN12 could overlap that of

pREA400, and not that of pREA250, during cell division or conjugation, leading to a stochastically more favored co-transfer with the former megaplasmid. There is some evidence that plasmids of different replication origins and machinery occupy distinct cytoplasmic domains in *Sinorhizobium meliloti* (7). I propose to determine the subcellular localization of all *Rhodococcus erythropolis* extrachromosomal replicons using fluorescence *in situ* hybridization (FISH) to determine whether pAN12 occupy similar domains as pREA400 and/or pREA250. Differentially-labeled sequence specific oligonucleotide probes (Cy3 or Cy5) can be generated against pAN12, pREA400 and pREA250 plasmids. Probes will be hybridized to plasmids in fixed wild type *Rhodococcus erythropolis* AN12 cells. My preliminary data suggests that *R. erythropolis* AN12 and SQ1 are amenable to cell biological studies and fluorescence microscopy (Figure 5.3), and that cell boundaries can be distinguished using lipophilic reagents.



R. erythropolis AN12 *R. erythropolis* SQ1

Figure 5.3. Fluorescence microscopy of rhodococci using the Di-I lipophilic dye (Molecular Probes).

Second, the expression of pAN12 and pREA400 plasmid conjugation genes could be transcriptionally coupled, such that conjugation machineries for both assemble at the same time, whereas genes required for pREA250 transfer are not. Certainly, the location of the unidentified pREA400-encoded gene(s) required for pAN12 mobility could be in the same operon as *traA*. This notion could be tested once the genes responsible are identified by genetic and/or biochemical means outlined above, then examining the mRNA expressions of both genes using Real-time RT (reverse transcription) PCR. Our preliminary data suggests that successful transfer of the pAN12 fusion plasmid occurs as early as 4 hours after mating (A. Posada, unpublished data). It's possible to do the same

experiment for both pREA400 and pREA250 to see whether the timing of either megaplasmid transfer event overlaps that of pAN12.

The third potential reason why pAN12 more frequently co-mobilize with pREA400 may be topology/plasmid replication related. My preliminary data strongly suggests pREA250 is a circular plasmid that replicates via a rolling-circle (RCR) mechanism (data not shown). Since pAN12 is also a RCR plasmid, the incompatibility between pAN12 and pREA250 could explain why pREA250 is found in relatively low abundance compared with other AN12 megaplasmids. One could imagine when pREA250 is established in the SQ1 recipient cell following conjugation, the co-transferred pAN12 is frequently eliminated in the absence of selection. This model implies that all plasmid are equally likely to be transferred during conjugation, but may compete for establishment.

5.3. Outlook. *Rhodococcus* is emerging as a genetically, biochemically and cell-biologically tractable microorganism. Further understanding and manipulation of the plethora of metabolic pathways rhodococci encode will benefit current and future industrial processes. Furthermore, characterization of its diverse horizontal gene pool in molecular detail may uncover novel mechanisms in plasmid conjugation that would be of interest to the applied as well as the basic biologist.

5.4. References.

1. **Ansari, A., and M. R. Gartenberg.** 1997. The yeast silent information regulator Sir4p anchors and partitions plasmids. Mol Cell Biol **17:**7061-8.
2. **Datta, S., C. Larkin, and J. F. Schildbach.** 2003. Structural insights into single-stranded DNA binding and cleavage by F factor TraI. Structure **11:**1369-79.
3. **Everett, R., and N. Willetts.** 1980. Characterisation of an *in vivo* system for nicking at the origin of conjugal DNA transfer of the sex factor F. J Mol Biol **136:**129-50.
4. **Grohmann, E., G. Muth, and M. Espinosa.** 2003. Conjugative plasmid transfer in gram-positive bacteria. Microbiol Mol Biol Rev **67:**277-301, table of contents.
5. **Harley, M. J., D. Toptygin, T. Troxler, and J. F. Schildbach.** 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:**6460-8.

6. **Iyer, L. M., K. S. Makarova, E. V. Koonin, and L. Aravind.** 2004. Comparative genomics of the FtsK-HerA superfamily of pumping ATPases: implications for the origins of chromosome segregation, cell division and viral capsid packaging. Nucleic Acids Res **32:**5260-79.
7. **Kahng, L. S., and L. Shapiro.** 2003. Polar localization of replicon origins in the multipartite genomes of *Agrobacterium tumefaciens* and *Sinorhizobium meliloti*. J Bacteriol **185:**3384-91.
8. **Lahue, E. E., and S. W. Matson.** 1990. Purified *Escherichia coli* F-factor TraY protein binds *oriT*. J Bacteriol **172:**1385-91.
9. **Lanka, E., and B. M. Wilkins.** 1995. DNA processing reactions in bacterial conjugation. Annu Rev Biochem **64:**141-69.
10. **Nelson, W. C., M. T. Howard, J. A. Sherman, and S. W. Matson.** 1995. The *traY* gene product and integration host factor stimulate *Escherichia coli* DNA helicase I-catalyzed nicking at the F plasmid *oriT*. J Biol Chem **270:**28374-80.
11. **Nelson, W. C., B. S. Morton, E. E. Lahue, and S. W. Matson.** 1993. Characterization of the *Escherichia coli* F factor *traY* gene product and its binding sites. J Bacteriol **175:**2221-8.
12. **Pansegrau, W., W. Schroder, and E. Lanka.** 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:**2782-9.
13. **Starai, V. J., I. Celic, R. N. Cole, J. D. Boeke, and J. C. Escalante-Semerena.** 2002. Sir2-dependent activation of acetyl-CoA synthetase by deacetylation of active lysine. Science **298:**2390-2.

Appendix A. Colony morphology and solvent tolerance of *Rhodococcus erythropolis* AN12.

A.1. A correlation between *Rhodococcus erythropolis* AN12 morphology and the megaplasmid, pREA100. Expanding upon the earlier observation that AN12 WT cells generally exhibit a mucoid colony morphology during early growth (1-3 days) on LB plates, and that AN12PL or AN12PL-1F6 cells exhibit dry colony morphology during all phases of growth, I asked whether AN12 WT sibling colonies that adopted a dry colony morphony early in growth also have been cured of the small megaplasmid, pREA100. The extrachromosomal contents of three spontaneous mutants (JY759, JY760, and JY761), along with AN12 WT and AN12PL-1F6, were examined by pulsed-field gel (Figure A.1).

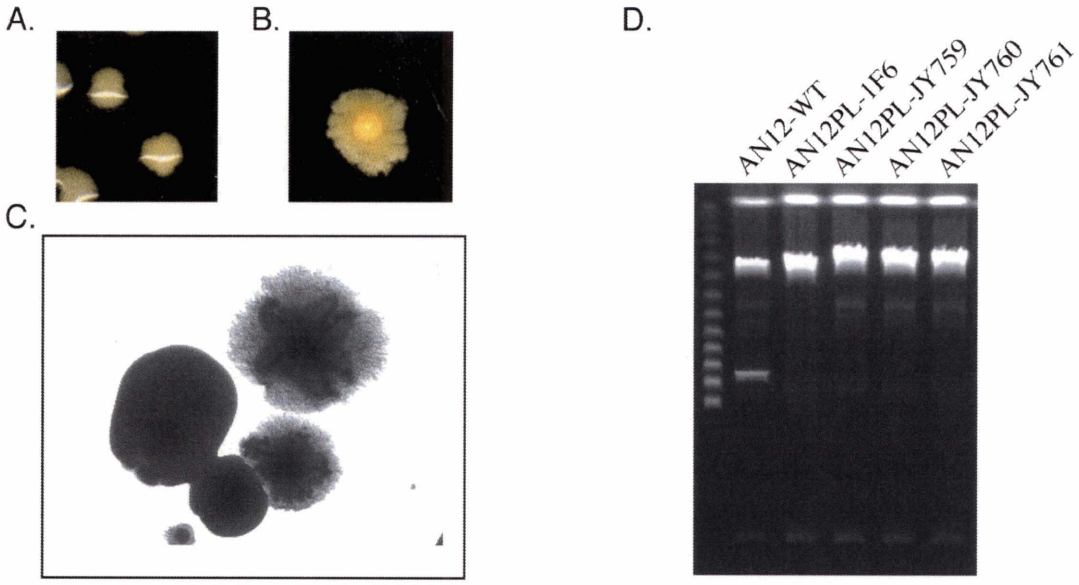


Figure A.1. The absence of megaplasmid pREA100 is correlated with dry colony morphology. Wild-type *R. erythropolis* AN12 (A) and AN12PL-1F6 (B) from LB plates after 3 days (A) and 5 days (B) of growth were photographed using a EPSON flatbed scanner. (C) Morphologically wild-type (mucoid, left) and a spontaneous mutant (dry, right) were isolated from frozen cultures and photographed using brightfield stereo-microscopy. (D) PFG profiles of indicated strains.

Results show that, indeed, the dry colony morphology is a good external indicator of the curing of megaplasmid pREA100, and suggest that the production of the mucoid

substance may depend on genes encoded by the unstable megaplasmid, pREA100. Clearly, this substance is not required for plasmid conjugation in rhodococci, since AN12PL-1F6 is fully capable of mating with *R. erythropolis* SQ1.

Rhodococcus has been shown to be a producer of extracellular polysaccharides (EPS), which aid in biodegradataion. (4, 5, 12, 13). They also produce biosurfactants (1, 9). The mucoid substance associated with wild-type AN12 cells is yet to be structurally determined (HPLC or mass-spectrometry). However, my preliminary data suggests that an optically active substance (OD_{600}) can be extracted using the non-polar detergene, Triton X-100, and is associated with mucoid AN12 cells but not dry AN12 derivatives (data not shown). Furthermore, the production of these compound does not confer any advantages to *R. erythropolis* AN12 when grown on solvents such as hexadecane (data not shown). Further characterization of pREA100 and the substance may elucide a genetic pathway for either EPS or biosurfactant production in *R. erythropolis* AN12.

A.2. Innate and gain of solvent tolerance in *Rhodococcus erythropolis* AN12. A genetic screen for gain of solvent tolerance abilities using both *Escherichia coli* and the *Rhodococcus erythropolis* strain, AN12 is proposed below.

Our preliminary evidence indicates that AN12 can grow in media supplemented with as much as 3% ethanol (Figure A.2.). Much of this tolerance is likely due to aspects of AN12 physiology (cell wall composition, efflux pumps etc.), and the extent of its tolerance is limited by its genome. I propose to exploit the metagenome of culturable and unculturable microorganisms in the following experiment to increase the inherent AN12 ethanol tolerance potential. The vast majority of microbes on our planet remain unidentified and uncharacterized. Their genomes, or collectively, the metagenome (3) represents a treasure trove of genetic potential.

The method outlined in Figure 3 is based on similar strategies used successfully to produce novel antibiotic compounds in *E. coli* (2, 10). Genomic DNA will be prepared from the various sources including environmentally isolated bacteria (which can be from contaminated sites or from pristine soil then enriched subsequently in the laboratory for alcohol tolerance), and the budding yeast, *Saccharomyces cerevisiae*, some strains of which have been reported to be tolerant of greater than 20% ethanol.

An vector based on the well characterized pBeloBAC11 (7), will be adapted to generate an ideal bacterial artificial chromosome (BAC) library backbone, which will feature not only an *E. coli* origin of replication, *Rhodococcus* AN12 compatible drug resistance markers (such as antibiotic resistance genes to gentamicin or kanamycin), but also a cassette homologous to a conjugative AN12 megaplasmid, pREA400. This last BAC library plasmid feature will allow not only for 1) efficient recombination (up to 1×10^4 CFUs/μg for 10 kb plasmids) and maintenance of purified BAC library plasmids by the AN12 genome, 2) direct assay of ethanol tolerance in AN12 without further subcloning, but also 3) the transfer of any novel function associated with the integrated library clone to more recalcitrant strains rhodococci via bacterial conjugation.

AN12 transformed with library clones will be screened for viability at 4% (or higher) concentration of ethanol, which is lethal for wild type AN12 (Figure A.2). A similar growth curve will be performed for the suitable *E. coli* BAC clone host to define critical concentrations, and gain of ethanol tolerance. Likewise, *E. coli* bearing the BAC clones can be assayed directly for gain of ethanol tolerance, by comparison to a strain harboring the empty BAC vector alone as a negative control. We anticipate generating multiple libraries depending on site or specific organism from which a (meta)genome was isolated. Once these libraries have been constructed, arrayed on 96 well plates, it will remain an useful tool to the screen for rhodococci and *E. coli* tolerance to other solvents.

Candidate that survive the ethanol selection will then be tested for even higher tolerance. The most interesting candidates are those which confer the greatest solvent tolerance potential, and will be among the first to be characterized. Genomic DNA will be isolated *Rhodococcus* and *E. coli* BAC transformants to address whether this gain-of-function is genetically due to the heterologous DNA from the BAC clone. This will involve plasmid rescuing the BAC clone from AN12 transformant genomic DNA, or straight forward plasmid purification from the *E. coli* candidate. Because sequencing the entire BAC insert of interest at this point can be costly and time-consuming, the BAC insert will be cleaved into smaller inserts of approximately 10 kb using at least two different restriction enzymes, and each of these subclones will be tested again for the ability to confer ethanol tolerance to either wild type AN12 or *E. coli*. We need only sequence positive subclones, using an inexpensive on-site sequencing facility (MIT

Biopolymers Laboratory). From the sequence information, we can define putative ORF using bioinformatics software such as GENEMARK (Georgia Institute Technology), then annotate the gene function of using the software, CAPASA, which we have developed in the lab to facilitate the I24 *Rhodococcus* genome sequencing project in our collaboration with the Cambridge-MIT Institute. A variety of methods will then be used to identify and characterize the heterologous gene(s) of interest, such as confirming gene activation in the presence of alcohol by fusing its promoter sequence to reporter genes (*lacZ*, *gusA*, or *GFP*) or by generating a non-sense mutation which concomitantly abolishes ethanol tolerance.

We anticipate the identification of genes whose role in solvent tolerance are already characterized, as well as genes whose role in solvent tolerance were not understood. The isolation of genes from the first class will give us important information regarding general mechanisms by which many soil microorganisms withstand environmental stress. Genes in the second class, though more challenging to characterize, will ultimately be the most rewarding as they represent novel, or unexpected, pathways by which organisms adapt to ethanol.

There are a few theoretical considerations for the above experiments. First, as the gain of novel solvent tolerance by either *E. coli* or *R. erythropolis* AN12 is dependent upon the expression of genes encoded by library DNA, the capacity of either of these microorganisms to express exogenous DNA must be addressed. The *E. coli* potential is already evident by the recombinant expression of a myriad of gene products from virtually all organisms studied to date, not the least of which is the gain of antibiotic producing capabilities (2) from what would be similar microbial metagenome DNA. Likewise, we do not anticipate the heterologous gene expression potential of *R. erythropolis* AN12 to be limiting, as another strain of *Rhodococcus erythropolis* has demonstrated capabilities to express recombinant proteins successfully, if not exceeding the capabilities of *E. coli* (8). The second limitation of this screening strategy is that gene clusters required for solvent tolerance must not exceed 40 kb, the average size of a BAC clone insert. Though genes we recover from this screen will be biased towards those that operate in smaller operons (~10 kb), we do not anticipate this to greatly affect the number of positive clones, since genes which affect solvent tolerance identified from studies in

Gram-negative bacterial systems are often single gene products, such as the *glpC* gene, which encodes a glycerol 3-phosphate dehydrogenase subunit C important for organic solvent tolerance of *E. coli* (11) and the heat shock protein, Lo18, which functions in ethanol tolerance of the malolactic bacteria, *Leuconostoc oeni* (6).

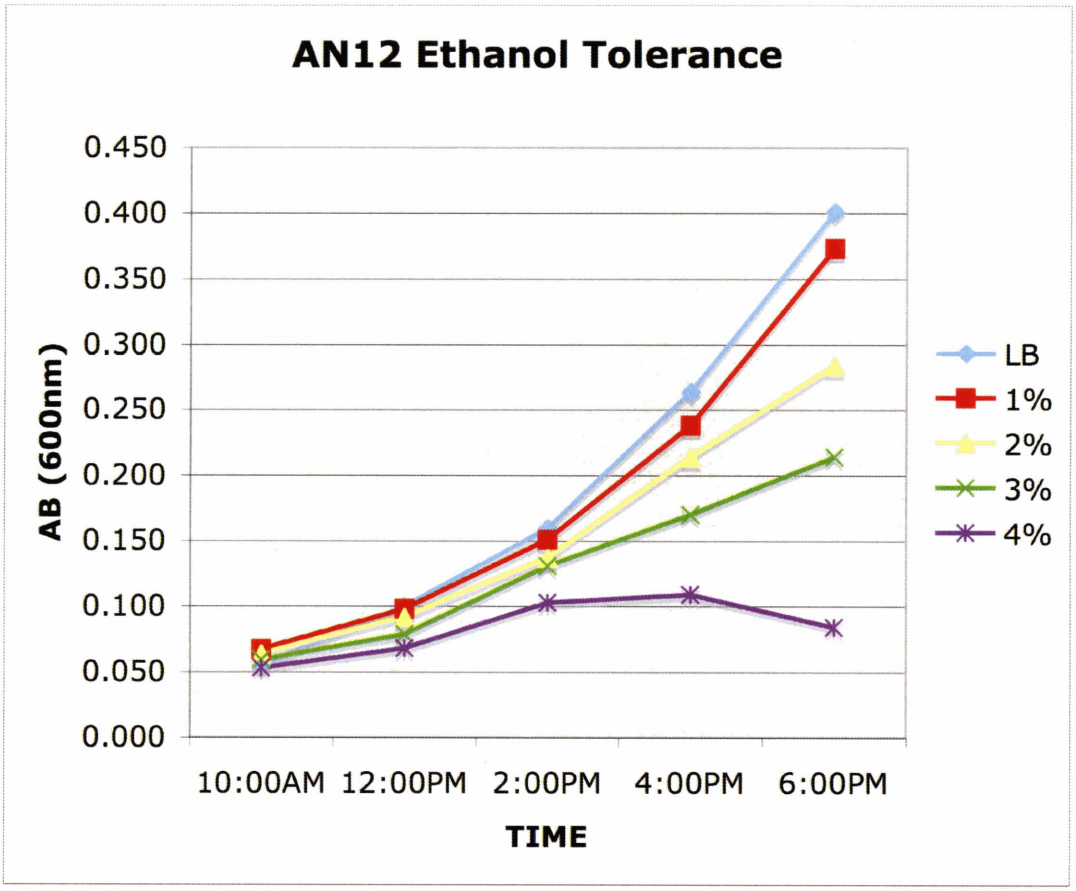


Figure A.2. Inherent ethanol tolerance of *Rhodococcus erythropolis* AN12. Plain LB or LB media with 1%, 2%, 3% or 4% (final concentration, v/v) ethanol were used to cultivate AN12, starting with equal 1:100 dilutions of the same overnight culture (grown in 2.5 ml LB) as inoculums. Growth of cells in agitated baffle flasks after every 2 hours was monitored using a spectrophotometer.

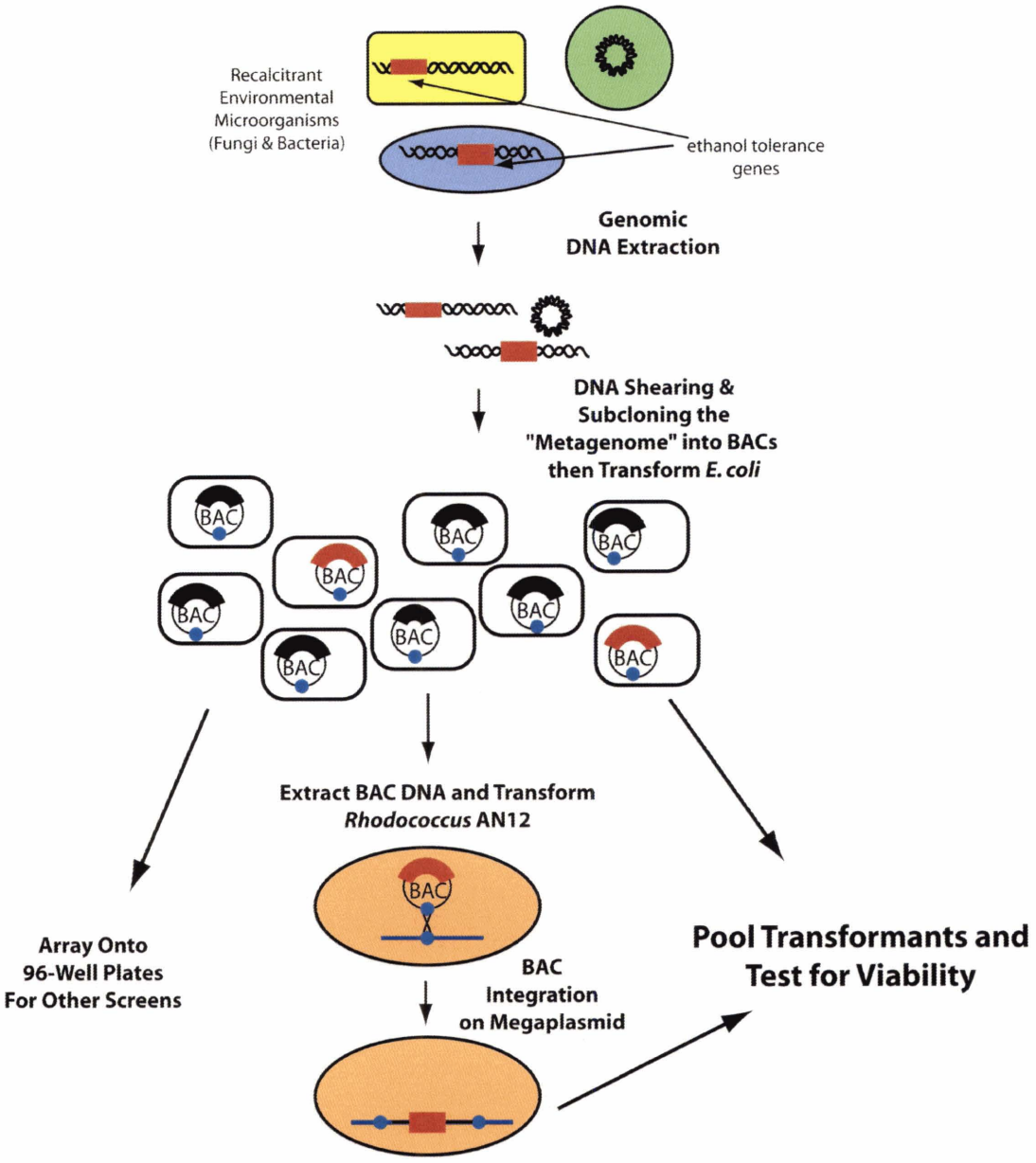


Figure A.3. Strategy for the genetic screening of the metagenome BAC library for the gain of ethanol tolerance in *E. coli* and *Rhodococcus* AN12.

A.3. References.

1. **Bryant, F. O.** 1990. Improved Method for the Isolation of Biosurfactant Glycolipids from *Rhodococcus* sp. Strain H13A. Appl Environ Microbiol **56:**1494-1496.

2. **Gillespie, D. E., S. F. Brady, A. D. Bettermann, N. P. Cianciotto, M. R. Liles, M. R. Rondon, J. Clardy, R. M. Goodman, and J. Handelsman.** 2002. Isolation of antibiotics turbomycin a and B from a metagenomic library of soil microbial DNA. Appl Environ Microbiol **68:**4301-6.
3. **Handelsman, J., M. R. Rondon, S. F. Brady, J. Clardy, and R. M. Goodman.** 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chem Biol **5:**R245-9.
4. **Iwabuchi, N., M. Sunairi, H. Anzai, M. Nakajima, and S. Harayama.** 2000. Relationships between colony morphotypes and oil tolerance in *Rhodococcus rhodochrous*. Appl Environ Microbiol **66:**5073-7.
5. **Iwabuchi, N., M. Sunairi, M. Urai, C. Itoh, H. Anzai, M. Nakajima, and S. Harayama.** 2002. Extracellular polysaccharides of *Rhodococcus rhodochrous* S-2 stimulate the degradation of aromatic components in crude oil by indigenous marine bacteria. Appl Environ Microbiol **68:**2337-43.
6. **Jobin, M. P., D. Garmyn, C. Divies, and J. Guzzo.** 1999. Expression of the *Oenococcus oeni trxA* gene is induced by hydrogen peroxide and heat shock. Microbiology **145 (Pt 5):**1245-51.
7. **Kim, U. J., B. W. Birren, T. Slepak, V. Mancino, C. Boysen, H. L. Kang, M. I. Simon, and H. Shizuya.** 1996. Construction and characterization of a human bacterial artificial chromosome library. Genomics **34:**213-8.
8. **Nakashima, N., and T. Tamura.** 2004. A novel system for expressing recombinant proteins over a wide temperature range from 4 to 35 degrees C. Biotechnol Bioeng **86:**136-48.
9. **Noda, K., K. Watanabe, and K. Maruhashi.** 2003. Isolation of the *Pseudomonas aeruginosa* gene affecting uptake of dibenzothiophene in n-tetradecane. J Biosci Bioeng **95:**504-11.
10. **Rondon, M. R., P. R. August, A. D. Bettermann, S. F. Brady, T. H. Grossman, M. R. Liles, K. A. Loiacono, B. A. Lynch, I. A. MacNeil, C. Minor, C. L. Tiong, M. Gilman, M. S. Osburne, J. Clardy, J. Handelsman, and R. M. Goodman.** 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl Environ Microbiol **66:**2541-7.
11. **Shimizu, K., S. Hayashi, T. Kako, M. Suzuki, N. Tsukagoshi, N. Doukyu, T. Kobayashi, and H. Honda.** 2005. Discovery of *glpC*, an organic solvent tolerance-related gene in *Escherichia coli*, using gene expression profiles from DNA microarrays. Appl Environ Microbiol **71:**1093-6.
12. **Sivan, A., M. Szanto, and V. Pavlov.** 2006. Biofilm development of the polyethylene-degrading bacterium *Rhodococcus ruber*. Appl Microbiol Biotechnol.
13. **Urai, M., T. Aizawa, H. Anzai, J. Ogihara, N. Iwabuchi, B. Neilan, I. Couperwhite, M. Nakajima, and M. Sunairi.** 2006. Structural analysis of an extracellular polysaccharide produced by a benzene tolerant bacterium, *Rhodococcus* sp. 33. Carbohydr Res **341:**616-23.