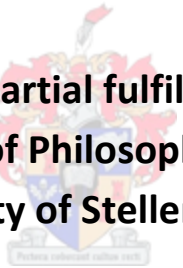
INSIGHTS INTO THE EVOLUTION OF IncQ PLASMIDS DERIVED FROM STUDIES ON pRAS3

By

Wesley Loftie-Eaton



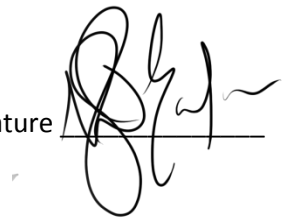
Dissertation presented in partial fulfilment of the requirements for the Degree of Doctor of Philosophy in Microbiology at the University of Stellenbosch.

Supervisor: Prof DE Rawlings

December 2010

Declaration

I, the undersigned, hereby declare that the work contained in this dissertation is my own original work and that I have not previously in its entirety or in part submitted it at any university for a degree.

Signature  31 August 2010

Abstract

Two isogenic plasmids, pRAS3.1 (11,851-bp) and pRAS3.2 (11,823-bp), were identified as tetracycline resistance plasmids occurring within *Aeromonas salmonicida* subsp. *salmonicida* and atypical *A. salmonicida* subsp. *salmonicida* strains that were isolated from salmon aquaculture farms in Norway (L'Abee-Lund and Sørum, 2002). Although sequence analysis showed that, except for the *repC* gene, the replication and mobilization genes of the two pRAS3 plasmids are similar to that of the two IncQ-2 plasmids pTF-FC2 and pTC-F14, incompatibility testing during the course of this study revealed that the replicons of the two pRAS3 plasmids were compatible with the replicons of the IncQ-1α, β and γ plasmids RSF1010, pIE1107 and pIE1130, as well as with the IncQ-2α and β plasmids, pTF-FC2 and pTC-F14, respectively. Through sequence analysis it was suggested that the *repC* gene of the ancestral pRAS3 plasmid was probably acquired during a gene exchange event with a yet to be identified plasmid. The difference in the RepC of the pRAS3 plasmids compared to that of the other IncQ-like plasmids against which the pRAS3 plasmids were tested for incompatibility was thus suggested to be a likely reason for the compatibility of the two pRAS3 plasmid replicons with these IncQ-1 and IncQ-2 plasmids.

Two previously unidentified genes, encoding two small 108 and 74-aa proteins distantly related to the PemIK (Bravo *et al.,* 1987; Tsuchimoto *et al*., 1988) and MazEF (Masuda *et al.,* 1993) TA systems, were found to be present between *repB* and *repA* genes of the two pRAS3 plasmids. Cloning of these two genes onto an unstable pOU82-test vector increased the stability of the vector from 35 to 98% after ~72 generations, thus suggesting that like the PasABC and PasAB systems of pTF-FC2 and pTC-F14, these two genes encode proteins which function as a toxin-antitoxin (TA) system. Although located in a similar position on the plasmids, the TA system of the two pRAS3 plasmids and the Pas systems of pTF-FC2 and pTC-F14 are unrelated, suggesting that these two types of TA systems were acquired independently from each other. Based on the sequence similarity and genetic organization of pRAS3 compared to the IncQ-2α and β plasmids pTF-FC2 and pTC-F14, respectively, but given that the pRAS3 plasmids were compatible with both pTF-FC2 and pTC-F14, as well as other IncQ-like plasmids, it was suggested that the two pRAS3 plasmids be classified into a new IncQ-2γ subgroup.

A comparison of the sequences of the two pRAS3 plasmids to each other by L'Abee-Lund and Sørum (2002) revealed that, apart from a number of point mutations within the *tetAR* tetracycline resistance genes of the two plasmids, the only other differences between them are that pRAS3.1 has 4 tandem copies of 22-bp iteron repeats within its origin of vegetative replication (*oriV*), and 5 tandem copies of CCCCCG 6-bp repeats near the origin of transfer (*oriT*), while pRAS3.2 has only three and four copies of each of the two repeated sequences, respectively. As the two pRAS3 plasmids are likely to have arisen from the same ancestor, this raised the question of how the copy numbers of these two different types of repeat sequences affected the ability of pRAS3.1 and pRAS3.2 plasmids to compete within a host cell as well as within a population of host cells, and therefore, why both of these isogenic plasmids have managed to persist in the environment. The plasmid copy numbers (PCN) of pRAS3.1 and pRAS3.2 were estimated to be 45 ± 13 and 30 ± 5 plasmids per chromosome, respectively. By creating a series of pRAS3.1 derivative plasmids with 3 to 7 copies of the 22-bp iterons and 4 or 5 copies of the 6-bp repeats, it was shown that an increase in the number of iterons brought about a decrease in PCN, probably due to an increased ability to bind RepC, while an increase in the number of 6-bp repeats from 4 to 5 brought about an increase in *repB* transcription, and the higher levels of RepB resulted in an increase in PCN. Thus the reason for pRAS3.1 having a ~1.5-fold higher PCN than pRAS3.2, even though it has 4 × 22-bp iterons compared to the 3 × 22-bp iterons of pRAS3.2, was that it had a higher level of *repB* transcription due to having 5 × 6-bp repeats in its *mobB-mobA/repB* promoter region compared to the 4 × 6-bp repeats of pRAS3.2. The differences in the number of iterons and 6-bp repeats, and hence PCN, did not have an effect on the stability of the two wild type (WT) plasmids or their derivatives even when the TA system was neutralized by having a copy of the TA genes present on a vector *in trans* and it was argued that the relatively high PCN of the two pRAS3 plasmids was sufficient to ensure plasmid stability through random distribution. As the two pRAS3 plasmids were mobilized at similar frequencies difference in PCN and *mobB-mobA/repB* transcription did not seem to affect their mobilization frequency. When pRAS3.1 and pRAS3.2 were competed intracellularly as coresident plasmids, pRAS3.1 was able to displace pRAS3.2 from 98% of the host cells within ~20 generations. The displacement of pRAS3.2 by pRAS3.1 was found to be as a result of pRAS3.1 having 4 × 22-bp iterons, which enabled pRAS3.1 to titrate of the communal pool of available RepC initiator proteins. Plasmids with 5 or 7 × 22-bp iterons, were however less effective at displacing a plasmid with 3 iterons, and it was speculated that

plasmids with more than 4 × 22-bp iterons within their *oriV* were less successful at initiating replication than was a plasmid with 3 iterons within its *oriV*. A direct correlation was found between the PCN of a pRAS3 plasmid and the metabolic burden it imposed on its host. Thus pRAS3.1, as a result of its ~1.5-fold higher PCN than pRAS3.2 placed a small but significantly higher (~2.8%) metabolic load on its host compared to pRAS3.2. It was concluded that pRAS3.1 had a competitive advantage over pRAS3.2 when these plasmids were coresident within a single host (as would have been when the two plasmids first diverged from each other) as it was able to displace pRAS3.2. However, as a result of pRAS3.2 having a lower PCN, it placed a smaller metabolic burden on an isogenic host and this resulted in pRAS3.2 having an advantage over pRAS3.1 at the population level. Sequence remnants of pRAS3.2 from horizontal gene transfer suggested that pRAS3.2 was the original pRAS3 plasmid and thus that pRAS3.1 evolved from pRAS3.2. As the pRAS3.1 derivative plasmids that were constructed during the course of this study are likely to have been intermediates in the evolution of pRAS3.1 from pRAS3.2, I was able to speculate on the stepwise evolution of pRAS3.1 from pRAS3.2 based on the characteristics of these plasmids, and thus, how both macro- and microevolutionary events have contributed to the evolution of these two plasmids.

Opsomming

Die twee isogeniese plasmiede, pRAS3.1 en pRAS3.2, was geïdentifiseer as tetrasiklien weerstandbiedende plasmiede wat in *Aeromonas salmonicida* subsp. *salmonicida* en nie-tipiese *A. salmonicida* voorkom (L'Abee-Lund and Sørum, 2002). DNS volgorde analise deur L'Abee-Lund en Sørum (2002) het gewys dat die gene verantwoordelik vir replisering (uitsluitend die *repC*) en mobililisering naverwant is aan die van twee IncQ-2 plasmiede, pTF-FC2 en pTC-F14. Eksperimente tydens hierdie studie het egter gewys dat die repliserende sisteme van die twee pRAS3 plasmiede versoenbaar is met die repliserende sisteme van die IncQ-1α, β and γ plasmiede RSF1010, pIE1107 en pIE1130, sowel as die IncQ-2α en β plasmiede, pTF-FC2 and pTC-F14, onderskeidelik. Analise van die aminosuur volgorde van die pRAS3 RepC proteïen het gedui daarop dat die proteïen taamlik verskil van die RepC proteïene van die naverwante plasmiede pTF-FC2 en pTC-F14, sowel as die van die IncQ-1 tipe plasmiede, en daar was voorgestel dat die voorsaat pRAS3 plasmied moontlik die *repC* geen bekom het vanaf 'n ander, nog onbekende, plasmied deur middel van horisontale geen uitruiling. Die verskil in die RepC van die pRAS3 plasmiede teenoor die van die ander IncQ plasmiede waarteen hulle getoets was vir onversoenbaarheid, was waarskynlik die rede waarom die pRAS3 plasmiede versoenbaar was met die IncQ-1 en IncQ-2 plasmiede. DNS volgorde analise tydens hierdie studie het die teenwoordigheid van twee, vantevore ongeidentifiseerde, klein 108 en 74 aminosuur proteïene onthul wat vêr langs verwant is aan die PemIK (Bravo *et al.,* 1987; Tsuchimoto *et al*., 1988) en MazEF (Masuda *et al.,* 1993) toksien-antitoksien sisteme. Die gene wat kodeer vir hierdie toksien-antitoksien proteïne kom tussen die *repB* en die *repA* gene van die twee pRAS3 plasmiede voor. Klonering van die toksien-antitoksien gene van die pRAS3 plasmiede op 'n ander onstabiele plasmied het die stabiliteit van hierdie plasmied verhoog van 35 tot en met 98% na ~72 generasies. Hierdie experiment het dus bevestig dat, soos die PasABC en PasAB sisteme van pTF-FC2 en pTC-F14 onderskeidelik, die twee gene 'n toksien-antitoksien sisteem kodeer wat die stabiliteit van 'n plasmied binne 'n bakteriese populasie kan verbeter. Alhoewel die toksien-antitoksien gene van pRAS3 op 'n soortgelyke posisie op die pRAS3 plasmiede voorkom as wat die *pasABC* en *pasAB* gene op hulle onderskeidelike pTF-FC2 en pTC-F14 plasmiede voorkom, is hulle nie verwant nie en dus was dit voorgestel dat die twee tipe toksien-antitoksien sisteme onafhanklik van mekaar verkry is. Aangesien die DNS volgorde en genetiese rangskikking van

pRAS3 teenoor die IncQ-2α en β plasmiede pTF-FC2 en pTC-F14, onderskeidelik, soortgelyk is, asook die feit dat die pRAS3 plasmiede versoenbaar was met pTF-FC2 en pTC-F14, sowel as ander IncQ tipe plasmiede, word dit voorgestel dat die twee pRAS3 plasmiede in 'n nuwe IncQ-2γ subgroep ingedeel word.

'n Vergelyking van die DNS volgorde van die twee pRAS3 plasmiede deur L'Abee-Lund and Sørum (2002) het gewys dat, behalwe vir 'n paar puntmutasies binne die *tetAR* tetrasiklien weerstandsgene, verskil die twee net in die opsig dat pRAS3.1 het 4 agtereenvolgende kopieë van 22-bp 'iteron' herhalings wat geleë is binne sy replikasie oorsprong en 5 kopieë van 'n CCCCCG 6-bp herhaling wat naby sy oorsprong van oordrag geleë is, terwyl pRAS3.2 net 3 en 4 kopieë het van elk van die onderskeie volgorde herhalings. Dus die bestaan van twee plasmiede met verskillende kopiegetalle van die twee verskillende tipe DNA volgorde herhalings, maar wat vermoedelik afkomstig is vanaf dieselfde stam plasmied, bring die volgende oorhoofse vrae aangaande die plasmiede na vore: hoe beïnvloed die DNS volgorde herhalings die vermoë van die twee plasmiede om binne 'n enkele gasheersel te kompeteer vir die beskikbare plasmied repliserings masjinerie, en hoe beïnvloed dit die plasmied-gasheersel verhouding en dus hulle vermoë om te kompeteer op die populasie vlak, en laastens, hoekom het beide weergawes van die plasmied bly voortbestaan in die omgewing? Die plasmied kopiegetalle van pRAS3.1 en pRAS3.2 was eksperimenteel beraam by ongeveer 45 ± 13 en 30 ± 5 plasmiede per chromosoom in *E. coli*, onderskeidelik. Deur 'n reeks van pRAS3.1 derivate te skep met 3 tot 7 'iteron' herhalings en 4 of 5 kopieë van die 6-bp herhalings was dit bewys dat 'n toename in die hoeveelheid 'iterons' 'n afname in die plasmied kopiegetal veroorsaak, vermoedelik deur 'n verbeterde vermoë om RepC te bind, terwyl 'n verhoging van 4 tot 5 kopieë van die 6-bp herhaling 'n afname in die kopiegetal te weeg gebring het. Die *repB* geen van 'n plasmied met 5 × 6-bp herhalings was ~2-voud hoër uitgedruk as die van 'n plasmied met 4 × 6-bp herhalings, en dit was verder bewys dat 'n verhoogde vlak van *repB* transkripsie vanaf 'n L-arabinose induseerbare promoter *in trans* van 'n pRAS3 plasmied met 4 × 6-bp herhalings het 'n ~2-voud verhoging in plasmied kopiegetal teweeg gebring. Die rede dat pRAS3.1 'n ~1.5-voud hoër plasmied kopiegetal gehad het as pRAS3.2, was as gevolg van 'n hoër vlak van *repB* uitdrukking weens die feit dat pRAS3.1 5 × 6-bp herhalings in die *mobB-mobA/repB* promoter area het terwyl pRAS3.2 net 4 van die 6-bp herhalings in dieslefde posisie het. Sou pRAS3.1 4 × 22-bp `iterons' gehad het, maar saam met 4 × 6-bp herhalings

soos pRAS3.2, dan sou die plasmied kopiegetal 23 ± 2 plasmiede per chromosoom gewees het. Die verskil in die hoeveelheid `iterons' en 6-bp herhalings, en dus die plasmied kopiegetal, het nie 'n effek op die stabiliteit van die wilde tipe plasmiede of hulle derivate gehad nie, selfs al was die toksien-antitoksien sisteem geneutraliseer deurdat daar 'n kopie van die toksien-antitoksien sisteem op 'n ander plasmied *in trans* van die pRAS3 plasmiede en hul derivate geplaas was. Die relatiewe hoë plasmied kopiegetal van die pRAS3 plasmiede, wat moontlik hoog genoeg was om plasmied stabiliteit deur middel van toevallige uitdeling te verseker, was voorgestel as die rede vir die hoë mate van plasmied stabiliteit. Soortgelyke frekwensies van mobilisasie vir pRAS3.1 en pRAS3.2 (0.032 ± 0.014 en 0.021 ± 0.013 transkonjugate per donateur, onderskeidelik) was waargeneem. Dus het dit geblyk dat die verskil in uitdrukking van die *mobB-mobA/repB* operon, sowel as die plasmied kopiegetal van die twee pRAS3 plasmiede, nie die mobiliserings frekwensie beïnvloed het nie. Intrasellulêre kompetisie tussen pRAS3.1 en pRAS3.2 het gewys dat pRAS3.1 die vermoë gehad om binne ~20 generasies pRAS3.2 vanuit 98% van die gasheerselle te skop. Daar was gewys dat die teenwoordigheid van 4 × 22-bp `iterons' in die oorsprong van replikasie van pRAS3.1 die rede was vir die vermoë van hierdie plasmied om pRAS3.2 uit te kompeteer binne die gasheersel, moontlik deurdat die 4 × 22-bp `iterons' beter in staat was om die RepC proteïn te bind. Die vermoë van plasmiede met 5 of 7 × 22-bp `iterons' om te kompeteer met 'n plasmied met net 3 × 22-bp `iterons' was toenemend swakker in vergelyking met die van 'n plasmied met 4 × 22-bp `iterons', en hierdie waarneming het gelei tot die voorstel dat plasmiede met meer as 4 × 22-bp `iterons' nie so suksesvol was om replikasie te inisieer soos 'n plasmied met 3 × 22-bp `iterons' nie. 'n Direkte korrelasie was gevind tussen die plasmied kopiegetal van 'n pRAS3 plasmied en die metaboliese lading wat die plasmied op die gasheersel geplaas het. Dus het pRAS3.1, met 'n plasmied kopiegetal van ~1.5-voud hoër as die van pRAS3.2, 'n effens hoër (~2.8%) metababoliese lading op die gasheersel as pRAS3.2 geplaas. In gevolge van die inter- en intrasellulêre kompetiesie eksperimente, was dit ge-argumenteer dat pRAS3.1 'n mededingende voordeel bo-oor pRAS3.2 binne 'n gasheersel (soos wat dit sou gewees het kort nadat die twee plasmiede van mekaar uiteengevloei het) gehad het omdat dit in staat was om pRAS3.2 vanuit die gasheersel te skop. Aan die ander kant het pRAS3.2 'n laer plasmied kopiegetal en dus 'n laer metaboliese lading op die isogeniese gasheersel geplaas het, en daardeur het pRAS3.2 weer op die populasievlak die kompeterende voordeel bo-oor pRAS3.1 gehad. Die eienskappe van pRAS3.2 was meer soortgelyk aan die van ander IncQ-tipe

plasmiede as wat die eienskappe van pRAS3.1 was, en dus word dit voorgestel dat pRAS3.1 vanaf pRAS3.2 afkomstig was. Omdat die derivaat plasmiede wat geskep was vanaf pRAS3.1 tydens hierdie studie moontlike tussengangers in die ontwikkeling van pRAS3.1 vanaf pRAS3.2 was, kan gespekuleer word, gebaseer op die eienskappe van hierdie plasmiede, oor die "stapsgewyse manier" waarmee pRAS3.1 vanaf pRAS3.2 ge-evolueer het, en dus hoe beide makro- en mikro-evolusionere gebeurlikhede bygedra het tot die evolusie van genoemde plasmiede.

Acknowledgements

I am deeply grateful to my supervisor, Professor D. E. Rawlings, for giving me the opportunity and the freedom to learn, experiment and make this thesis my own. Through his knowledge and understanding of science and of life he has truly managed to create an environment where any eager and willing young scientist can come into his own. His enthusiasm for science and his dedication to his students makes him a true inspiration to any student who has had the privilege of being taught by him.

The people that have been instrumental in the success of this project and who deserves a special mention are Dr Gwynneth Matcher, Dr Shelly Deane and Lonnie van Zyl. We have spent countless hours standing next to a white board with marker in hand (or sitting around a table with a beer in hand) discussing and deliberating aspects of this project and of science in general. It was during those moments that I learned how science really works and what it is truly about. I not only have the highest level of respect for them as professionals, but also as my close friends.

I give thanks to all my friends that I have accumulated throughout my time in the Department of Microbiology. Their encouragement, compliments and also criticisms kept me focused especially in the time I spent writing this thesis.

I would like to thank the National Research Foundation, BHP-Billiton, and the University of Stellenbosch for financial support.

Most of all, my deepest love and gratitude goes to my sister Jeanice, my father Eugene and my mother Theresa. Through living with me my sister experienced firsthand all my successes and tribulations during this project, and she was always ready to help around the house when I was simply too busy to do so. Without her support this experience would not have been as pleasant. My father once said to me that no matter what your qualifications are, you have not truly graduated unless you have graduated from the university of life. This has resonated within me throughout my studies and has kept me humble and meek. He has always supported me in who, what and wherever I wanted to be, and in so doing he has been a true mentor. Through my mother's devotion and love for her family she has taught me diligence,

patience and sacrifice, all of which are qualities without which I would not have been able to complete this project.

Finally, my thoughts go to my grandfather whom passed away months before seeing me graduate. Retelling his stories passed on to me always provided a source of laughter in the laboratory. I miss him dearly.

Table of Contents

Abbreviations

A	adenosine
A	alanine
Å	Årmström units
A + T-rich	adenosine and thymidine rich sequence
aa	amino acids
Ap	ampicillin
ATP	adenosine-5'-triphosphate
bp	base pairs
°C	degrees Celcius
C	cystein
C	cytosine
C-terminal	carboxyl-terminus
cDNA	complimentary DNA
Cm	chloramphenicol
C_T	threshold cycle
ddd.H_2O	double distilled deionized water
D	aspartic acid
DNA	deoxyribonucleic acid
DNA Pol I	DNA polymerase I
DNA Pol III	DNA polymerase III
DR	direct repeats
dsDNA	double stranded DNA
Dtr	DNA transfer and replication system
E	glutamic acid
EDTA	ethylenediaminetetraacetic acid
F	phenylalanine
g	gram
G	guanine
G	glycine
G + C-rich	guanosine and cytosine rich sequence
gDNA	genomic DNA
H	histidine
I	isoleucine
IHF	integration host factor
Inc	incompatibility

IPTG	isopropyl β-D-1-thiogalactopyranoside
IR	inverted repeats
K	lysine
Kb	kilobase pairs or 1000-bp
kDa	kilodaltons
Km	kanamycin
L	leucine
L	liter
M	methionine
M	Molar
ml	milliliters
mM	millimolar
MOPS	3-[N-Morpholino]propanesulfonic acid
Mpf	mating pair formation
mRNA	messenger RNA
N	asparagine
N	normal
N-terminal	amino-terminus
nt	nucleotide
ORF	open reading frame
oriT	origin of transfer
oriV	origin of vegetative replication
P	proline
PCN	plasmid copy number
PCR	polymerase chain reaction
Q	glutamine
qPCR	quantitative PCR
R	arginine
R^2	correlation coefficient
RBS	ribosomal binding site
Rep	replication proteins
RNA	ribonucleic acid
RT-PCR	reverse transcriptase PCR
S	serine
SDS	sodium dodecyl sulfate
Sm	streptomycin
SSB	single-stranded DNA binding protein
ssDNA	single stranded DNA

ssi	single stranded initiation site
Su	sulfonamide
T	threonine
T	thymine
TE	Tris EDTA buffer
Tet	tetracycline
Tris	Tris (hydroxymethyl) aminomethane
TA	toxin-antitoxin system
TSS	transcriptional start site
μg	microgram
μl	microliter
V	valine
v/v	volume/volume
W	tryptophan
WT	wild type
w/v	weight/volume
Y	tyrosine
X-gal	5-bromo-4-chloro-3-indolyl-b-D-galactopyranoside

Chapter 1

Literature review

Table of Contents

1.1 INTRODUCTION

Plasmids are seen as selfish DNA entities that propagate and spread throughout bacterial, yeast and fungal communities. As such, plasmids are not only mediators of evolution, but are in themselves also subject to evolutionary processes (Sýkora, 1992). As with prokaryotic chromosomes, plasmid evolution is a continuous process driven by selective pressure and is mediated by insertions, deletions, rearrangements and base pair substitutions. Mutational events that lead to the formation of new plasmid incompatibility groups, or plasmid families, are referred to as macroevolutionary events, while micromutation refers to all mutations that do not cause a change in incompatibility or replicon specificity. Only mutational events that satisfy the criteria imposed by selective pressures lead to evolutionary success (Bergstrom *et al.,* 2000). Such criteria are to maximize benefit to the host, minimize burden on the host, and to increase the efficiency of spread throughout the environment.

1.2 CRITERIA FOR SUCCESS

1.2.1 Benefit to the Host

With the use of mathematical models, Bergstrom and coworkers (Bergstrom *et al.,* 2000) have shown that plasmids cannot be maintained indefinitely in single-clone populations of bacteria within the environment, even if selection favours the genes that they carry. They hypothesize that genes that are highly advantageous will eventually be incorporated into the host's chromosome and the plasmid will eventually be lost. Plasmids therefore have to find a compromise between carrying genes that are too useful, and carrying genes that are selfish in nature. The benefits conferred by plasmids include the broadening of phenotypic characteristics, increased gene dosage and increased gene transfer (Thomas, 2004).

1.2.2 Cost to the Host

Plasmid-carriage imposes an additional metabolic load on the host as a result of replication, gene expression and translation of the elements required for maintenance of the extrachromosomal DNA (Glick, 1995). Therefore, in the absence of selection, plasmid-containing hosts run the risk of being outcompeted by the otherwise isogenic less burdened, therefore faster-growing plasmid-free segregants and thus the plasmid will eventually be completely lost from the population (Lenski and Bouma, 1987; Lenski, 1992). A variety of

strategies exists and are used in combination by plasmids to minimize the additional metabolic load imposed on the host. Such strategies include organization of genes involved in similar functions into operons that are subject to autoregulation, regulation of gene expression at both the transcriptional and translational level, and strict maintenance of plasmid copy number (PCN) so as to prevent runaway replication or plasmid loss (Bingle and Thomas, 2001; Thomas, 2000).

1.2.3 Plasmid Spread

Efficient spread throughout the environment requires stability during vertical inheritance as well as the ability to infect other species by means of horizontal transfer. Plasmids have adopted a variety of strategies in order to counter segregational loss. Such strategies include partitioning functions that ensure even distribution between daughter cells, multimer resolution systems that maximize the number of copies available for segregation or postsegregational killing (PSK) systems that eliminate plasmid-free cells (Bignell and Thomas, 2001; Cooper and Heinemann, 2000; Gerdes *et al.*, 1985).

The ability of plasmids to spread horizontally by means of conjugation or mobilization enables plasmids to spread between individual cells or between different species within a population. The advantages are that it increases the available replication space, provides the ability to evade the dangers of becoming extinct when the environmental conditions become unfavourable for the host and enables the plasmid to reinfect cells from which it has become lost (Bergstrom *et al.,* 2000; Norman *et al*., 2009).

Another, and often less considered, requirement for plasmid stability is its compatibility or incompatibility with other plasmids within the same population or cell. Multiple plasmids are often found within a single host and thus have to compete for the biosynthetic capabilities of the host (Thomas, 2004). If two competing plasmids within the same host are unable to coexist, as would be the case when a plasmid-containing host is infected by another related plasmid, or when a sister plasmid within a host diverges from its parent plasmid, then one of the plasmids stands a chance of being displaced by the other. Therefore, the model for the diversification of plasmid species is based on the requirement to generate diversity among replication and replication control systems in order for plasmids to become compatible (Sýkora, 1992; Thomas, 2004).

1.3 IncQ PLASMIDS

As will become evident in this chapter, plasmids belonging to the Incompatibility group Q (IncQ) family are prime candidates for plasmid-based evolutionary studies. There exist examples of well-characterized plasmids within the family that have diverged sufficiently to have formed two unique subgroups, IncQ-1 and IncQ-2. Within each of the subgroups there also exists examples of plasmids that have diverged enough to have formed even smaller α, β and γ groups of related plasmids, as well as plasmids that have undergone evolutionary changes but not to such an extent that the replicon specificity or incompatibility status has changed. Thus, several evolutionary lineages exist within the IncQ-like plasmid family and both the micromutation and macroevolutionary events can easily be identified.

The genetic organization of the IncQ plasmids is such that the plasmid backbone can be divided into three regions, namely the replicon, mobilization region and accessory DNA. The replicon (discussed in detail in section 2.3) consists of three genes and a corresponding *cis*-acting DNA-binding locus for each of the respective protein products required to initiate replication. The mobilization region (discussed in detail in section 2.7) is of two types, hence the IncQ-1 and IncQ-2 subgroups. The mobilization region of the IncQ-1 plasmids consists of three genes and a *cis*-acting locus for transfer, while that of the IncQ-2 plasmids consist of five genes in addition to the transfer locus. Even though the mobilization systems are evolutionarily unrelated they are functionally linked in both cases to the replicon *via* a gene-fusion. The mobilization system and replicon of both types of IncQ plasmids also share an essential regulatory region and thus demonstrate a genetic organization adapted to conform to the requirements of their unique type of single-strand displacement mechanism of replication. The accessory DNA-region (refer to section 2.1 and Table 1.1) may be cryptic or contain remnants of transposon activity or various antibiotic resistance genes. This, together with the broad-host-range (BHR) capability of IncQ plasmids and the diversity of ecological environments from which they have been isolated demonstrates the active participation of these plasmids within the horizontal gene pool.

1.3.1 Diversity and Ecology

Several plasmids with replicons that bear similarity to IncQ plasmids have been discovered across the world, although only a few have been well characterized and it is only these that will be discussed in detail within this chapter. A non-exhaustive list summarizing the size

and accessory DNA of IncQ-like plasmids, as well as the host and the geographical location from which the host was isolated is given in Table 1.1. The list was extracted and updated from a review that was written by Rawlings and Tietze in 2001.

The most extensively studied members of the IncQ-family are RSF1010 and the almost identical R1162. RSF1010 was isolated from a colicinogenic *Escherichia coli* strain during 1973 in Wisconsin, U.S.A, and conveys resistance to sulfonamides and streptomycin (Guerry *et al.,* 1974; Niedenzu *et al*., 2001). R1162 on the other hand, was isolated from *Pseudomonas aeruginosa* strain 1162 during 1972 in Alberta, Canada (Bryan *et al*., 1972). Sequence alignments from the available sequences in the NCBI database shows that it differs from RSF1010 in the region of the origin of replication due to four point mutations, two single base pair deletions and one 4-bp deletion (see Figure 1.3), and by 1-bp in the origin of transfer (see Figure 1.12). R300B, also near to identical to RSF1010, was isolated from *Salmonella typhimurium* serovar Typhimurium in London, U.K., during 1974 (Barth and Grinter *et al*., 1974). All of the bacterial strains were clinical isolates from different countries and continents, which early on emphasized the wide-spread nature of these plasmids, their host-diversity and economical importance.

RSF1010 (or R1162 and R300B, which will be referred to interchangeably from here on) is regarded as the prototype plasmid for IncQ plasmids and consists of 8.7-kb of DNA with a G+C mole ratio of 61%. It expresses 11 genes which can be divided into the three modules responsible for replication, mobilization and accessory functions as depicted in red, yellow and white, respectively, in Figure 1.1 (Scholz *et al.*, 1989). Each of these modules can function independently of each other provided all the components of the respective module is provided (Derbyshire and Willetts, 1987; Meyer *et al*., 1982; Scherzinger *et al*., 1991). The ability of RSF1010 and its equivalents to confer sulfonamide and streptomycin resistance to bacterial hosts is coded for by the *sulII* and *strAB* genes, respectively, which have inserted between the *repC* and *oriV* sequences (Scholz *et al.*, 1989).

The IncQ-like plasmids pIE1107, pIE1130, pIE1120 and pIE1115 were all isolated from uncultured bacterial communities found within a piggery manure slurry in Germany by means of biparental matings using *E. coli* and *P. putida* as the recipient hosts (Smalla *et al*., 2000; Tietze, 1998). They were identified from the pool of plasmids that were captured as being IncQ-like by dot blot Southern hybridizations using a probe made from the origin of vegetative replication (*oriV*) of RSF1010. Furthermore, PCR-based detection of IncQ-like plasmids in total community DNA extracted from different soil and piggery manure samples

indicated that these plasmids are very common in such environments and further screening of the pool of plasmids that were isolated from the piggery manure slurries revealed a high degree of diversity in antibiotic resistance and restriction profiles (Gotz *et al*., 1996; Smalla *et al*., 2000).

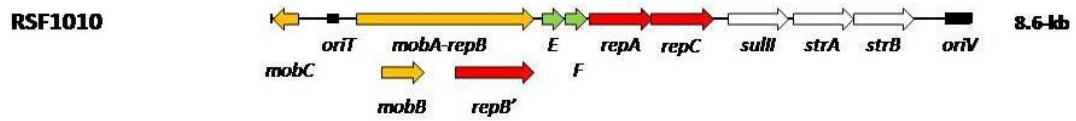


FIG. 1.1. A schematic representation of RSF1010. The genes involved in replication, mobilization and antibiotic resistance (accessory DNA) are indicated by red, yellow and white arrows, respectively. The origins of transfer and replication are indicated by black boxes. ORFs E and F, the functions of which are discussed in section 2.4, are indicated by green arrows.

Sequence analysis of pIE1130, pIE1120 and pIE1115 revealed *strA* and *strB* resistance genes that are arranged on the plasmid similarly as for RSF1010. Like RSF1010, pIE1130 and pIE1115 also carry *sulII* resistance genes, however, that of pIE1130 is located differently to that of RSF1010. In addition to the streptomycin and sulfonamide resistance genes, pIE1130 also carries genes conferring resistance to chloramphenicol (*catIII*) and kanamycin (*aph(3)-I*), while pIE1120 carries a novel *tetY* tetracycline resistance gene. In contrast to the streptomycin and sulfonamide resistant plasmids, pIE1107 contains a *sulII* resistance gene but it is truncated by 35 amino acids (aa) thus rendering it non-functional (Tietze, 1998). It does, however, have functional kanamycin (*aph(3)-I*) and streptothricin (*sat3*) resistance genes.

Not all the IncQ-like plasmids, however, are equipped with antibiotic resistance genes. The implications are that many more IncQ-like plasmids may exist and as a result of a lack of selectable markers may go unnoticed. The IncQ-like plasmid pDN1, for example, is a small (5.1-kb) cryptic IncQ-like plasmid that was only identified upon analysis of the genome sequence of *Dichelobacter nodosus*, a bacterium that was isolated in Australia and which, along with a consortium of other bacteria, is responsible for foot rot in sheep (Whittle *et al.,* 2000). Analysis of the pDN1 sequence revealed that it is 95% similar to pIE1107 over its length, but does not contain any antibiotic resistance genes.

TABLE 1.1. List of IncQ-like plasmids, the host from which they were isolated and accessory DNA

Plasmid[A]	Size (bp)	Source from which isolated	Genes and ORFs in addition to backbone[B]	Reference(s)
P89S	±8,180	*E. coli* (clinical)	Su	(Saano and Zinchenko, 1987)
pAZ1	±8,000	*S. enterica* serovar Typhimurium type 179	Su, Tp (DHFR type III)	(Fling *et al.,* 1988)
PB165	±11,900	*E. coli* (UK)	Sm, Su	(Barth and Grinter *et al*., 1974; Grinter and Barth, 1976)
pBRST7.6	7,621	*Aeromonas hydrophila* strain AO1	*qnrS2*	EU925817[D]
pCHE-A	7,560	*Enterobacter cloacae* (Canada)	*bla*GES-5, integron mobilization unit (IMU)	(Poirel *et al.,* 2009)
pDN1	5,112	*Dichelobacter nodosus* (Australia)	None	(Whittle *et al.,* 2000)
pFM202	±7,100	*Neisseria gonorrhoeae* (Spain)	Ap	(Rotger and Nombela, 1983)
pFM739	±9,450	*N. sicca* (Spain)	Ap, Sm, Su	(Rotger *et al.,* 1986)
pGNB2	8,469	Activated sludge (Germany)	*qnrS2*, Tn1721	(Bonemann *et al.,* 2006)
pHD148	±7,500	*Haemophilus ducreyi* (Kenya)	Su	(Albritton *et al*., 1982)
pHD8.1	±8,100	*Actinobacillus pleuropneumoniae* (Canada)	Sm, Su	(Willson *et al*., 1989)
pIE1107	8,520	Piggery manure (Germany)	*aph*(3')-Id, *sat*3, *sulII*[C]	(Tietze, 1998)
pIE1115	10,687	Piggery manure (Germany)	*linB*-like, *strAB, sulII*	(Smalla *et al*., 2000)
pIE1120	±9,100	Piggery manure (Germany)	*tetA*(Y), *strAB*	(Smalla *et al*., 2000)
pIE1130	10,687	Piggery manure (Germany)	*aph*(3')-I, *catIII, strAB, sulII*	(Smalla *et al*., 2000)
pIE639	±11,100	*E. coli* O20:H-	*aph*(3')-Id, *sat3*, *strAB*, *sulII*	(Tietze *et al.,* 1989)
pIE723	±9,500	*E. coli* O147:K88	*ant*(2")-Ia, *strAB, sulII*	(Tietze *et al.,* 1989)
pJA8102-2	11,823	*A. salmonicida* M28102 (Japan)	*tetAR*(C)	(Aoki and Takahashi, 1986)
pQ7	9,042	*E. coli* strain 7 (Switzerland)	*bla*GES-1, *bla*OXA/aac(6')-Ib, *int3*	FJ696404[D]
pRAS3.1	11,851	*A. salmonicida* subsp. *salmonicida* (Norway) and *A*. *salmonicida* subsp. *salmonicida* MT361 (Scotland)	*tetAR*(C)	(L'Abee-Lund and Sørum, 2002)
pRAS3.2	11,823	Atypical *A. salmonicida* (Norway)	*tetAR*(C)	(L'Abee-Lund and Sørum, 2002)
pTC-F14	14,155	*A. caldus* (South-Africa)	*tnp*, ORF13, ORF20.8, ORF17.2, ORF33	(Gardner *et al*., 2001)
pTF-FC2	12,184	*A. ferrooxidans* (South Africa)	*grx, merR*-like, ORF43, *tnpR*[c]	(Rawlings *et al*., 1984)
R1162	8,684	*P. aeruginosa* (Canada)	*sulII, strAB*	(Bryan *et al*., 1972)
R300B	8,684	*S. enteric* serovar Typhimurium (UK)	*sulII, strAB*	(Barth and Grinter *et al*., 1974)
R678	±14,000	*S. enteric* serovar Dublin (Denmark)	*Sm, Su*	(Barth and Grinter *et al*., 1974; Grinter and Barth, 1976)
R684	±9,500	*Proteus mirabilis*	*Sm, Su*	(Barth and Grinter *et al*., 1974; Grinter and Barth, 1976)
RSF1010	8,684	*E. coli* strain 3 (USA)	*sulII, strAB*	(Guerry *et al.,* 1974)

[A] Plasmids are listed in alphabetical order.

[B] If known the exact gene was given, otherwise the type of antibiotic resistance is given. Ap, ampicillin; Cl, clindamycin; Cm, chloramphenicol; Lm, linomycin; Km, kanamycin; Sm, streptomycin; Su, sulfonamide; Tc, tetracycline; To, tobramycin; Tp, trimethoprim. Ap resistance conferred by *bla*; Cl/Lm conferred by *linB*; Cm conferred by *catIII;* gentamycin/Km/To conferred by *ant*(2")-Ia; Km/Nm/Tb conferred by *ant*(2")-Ia; Km/Nm conferred by *aph*(3')-Id; quinolone resistance conferred by *qnrS2;* streptothricin resistance conferred by *sat3;* Sm resistance conferred by *strAB;* Su resistance conferred by *sulII*; Tc conferred by *tetA*(Y) and *tetAR*(C).

[C] Gene truncated

[D] Unpublished, NCBI accession number for DNA sequence

Two other cryptic IncQ-like plasmids, pTF-FC2 and pTC-F14, have also been isolated from *Acidithiobacillus ferrooxidans* and *Acidithiobacillus caldus* (both previously had the genus name *Thiobacillus*) during 1984 and 1998 (personal communication, S.M. Deane), respectively (Gardner *et al*., 2001; Rawlings *et al*., 1984; Rawlings *et al*., 1986). Both of these extremophiles form part of a consortium of organisms responsible for the leaching of metals at a bioleaching plant in South Africa. As antibiotics are not used in this environment it does not come as a surprise that there are no antibiotic resistance genes on these plasmids. A *Tn21*-like transposon has, however, integrated into the genome of pTF-FC2. Although the *tnpR* and *tnpA* genes coding for the resolvase and transposase enzymes, respectively, of this transposon are inactive as a result of mutation, the transposon retains the ability to be resolved when a functional *tnpR* from *Tn21* is supplied in *trans* (Clennel *et al*., 1995). Located between the 38-bp terminal inverted repeats (IR) of the transposon is a functional glutaredoxin-like gene which is able to compliment an *E. coli trxA*$^-$ mutant, a *merR*-like gene which potentially encodes a MerR regulator but no *merA*-like gene, as well as ORF8 and ORF43 both of which do not have any sequence similarity to any known proteins in the NCBI database. ORF43, however, appears to encode a 12-loop transmembrane protein similar to multidrug transporters but whose function is unknown (Rawlings, 2001). No MerR and ORF43 protein products are, however, produced in *E. coli*, but this does not mean that these two ORFs are not expressed in the native host *A. ferrooxidans.* pTC-F14 has 5 ORFs of unknown function in the same location but they are unrelated to those of pTF-FC2 (Gardner *et al*., 2001).

An excellent example of the ability of IncQ-like plasmids to spread and persist in the environment is the isolation of three tetracycline resistance plasmids, that are either

identical or close to identical, from three locations around the world. pJA8102-2 was isolated from *Aeromonas salmonicida* M28102 in Japan during 1981 (Aoki and Takahashi, 1986). Approximately 20 years later two nearly identical plasmids, pRAS3.1 and pRAS3.2, of which pRAS3.2 has an identical restriction pattern to pJA8102-2, were isolated from multiple strains of *Aeromonas salmonicida* subsp. *salmonicida* and atypical *Aeromonas salmonicida*, respectively, in Norway (L'Abee-Lund and Sørum, 2002). The same researchers also isolated pRAS3.1 from an *A. salmonicida* subsp. *salmonicida* MT361 strain originally from Scotland. In each of the instances the *A. salmonicida* strains were isolated from aquaculture farms where they are responsible for causing furunculosis in the salmon. Furthermore, a tetracycline resistance genomic island bearing 99% sequence identity to pRAS3.2 over 10.1-kb of shared DNA was discovered on the chromosome of *Chlamydia suis* R19, an obligatory intracellular pathogen that was isolated from pigs in the United States (Dugan *et al.*, 2004). The size difference between the integrated DNA and the 11.8-kb pRAS3.2 is as a result of three deletions. A 1.7-kb and a 44-bp deletion removed the *mobB-repB* gene of the integrated plasmid and two out of the three iterons (see section 2.3.2 for a discussion on iterons and genes essential for replication), respectively, thus rendering the replicon inactive. An 8-bp deletion within the promoter region of the *tetA*(C)-*teR*(C) tetracycline resistance genes resulted in constitutive expression of the *tetA*(C) gene even in the absence of tetracycline (Dugan *et al.*, 2004). As *C. suis* is an obligatory intracellular pig pathogen and *A. salmonicida* a fish pathogen with an optimum growth temperature of 20°C, how an equivalent of pRAS3.2 was acquired by *C. suis* is unknown but again illustrates the mobility of IncQ-like plasmids.

1.3.2 Reasons for Broad Host-Range

The diversity and geographic location of hosts which have been found to harbour IncQ-like plasmids emphasizes the ability of IncQ-like plasmids to spread and be maintained in a wide variety of bacteria. In addition to the three different species from which the near identical RSF1010, R1162 and R300B were isolated, RSF1010 and its derivatives have also been shown to replicate in at least 31 different Gram negative species as summarized by Frey and Bagdasarian (1989) and includes organisms such as *Aerobacter aerogenes*, *Caulobacter cresentus, Erwinia carotovora, Gluconobacter* sp., *Hypomicrobium* sp*., Klebsiella pneumonia, Paracoccus denitrificans* and *Vibrio cholerae* to name a few. It has also been mobilized from *E. coli* to Gram positive bacteria such as *Streptomyces lividans* and *Mycobacterium*

smegmatis where it was stably maintained for one week and expressed high levels of streptomycin and sulfonamide resistance (Gormley and Davies, 1991). In addition, RSF1010 has also been shown to replicate and be mobilized by the *virB/D4* and *avhB* type IV secretion systems of *Agrobacterium tumefaciens* which are responsible for delivering oncogenic DNA to plant cells and conjugal transfer of the cryptic pAtC58 plasmid between Agrobacteria, respectively (Cascales *et al*., 2005; Chen *et al*., 2002). RSF1010 can also replicate in *Legionella pneumophila* and be mobilized by its *dot* virulence secretion which is responsible for Legionnaire's disease (Vogel *et al*., 1998).

The most important contribution to the broad-host-range success of IncQ-like plasmids in addition to its ability to be readily mobilized by a large variety of conjugative plasmids and host-encoded Type IV secretion systems, is its ability to initiate its own replication independently from the host (Meyer, 2009). Initiation of replication at the *oriV*, explained in detail in section 2.3, is mediated by the three plasmid encoded proteins, RepB, RepA and RepC (Scherzinger *et al*., 1984). The RepC is a plasmid-specific replication initiator, while the RepB primase and RepA helicase perform functions that are analogous to those performed by the host encoded DnaG primase, DnaB helicase and DnaC helicase loader, respectively (Honda *et al.*, 1991; Miao *et al*., 1995; Scherzinger *et al*., 1991; Scherzinger *et al*., 1997). The lack of requirement for these host-encoded functions was demonstrated in separate *in vitro* experiments by the ability of RSF1010 DNA, provided with RepC, RepA and RepB to replicate while the RNA polymerase was inhibited by rifampicin and streptomycin, the DNA polymerase II was inhibited by aphidicolin, the DnaB, DnaG and i (*dnaT* product) proteins were inactivated by specific antibodies, or in the presence of inactivated DnaA and DnaC temperature sensitive mutant proteins (Scherzinger *et al*., 1991). Replication of the RSF1010 DNA in the *E. coli* enzyme fraction was, however, dependent on ATP, dNTPs, Mg^{2+}, DnaZ, the γ-subunit of the DNA polymerase III holoenzyme, and DNA gyrase for supercoiling. No secondary priming sites capable of initiating replication using host-encoded proteins are present on RSF1010, and the single-stranded mechanism of replication negates the need for coordinated leading and lagging strand synthesis at the replication fork, thus making it independent of many host-functions and allowing replication to be more robust and versatile in different hosts (De Graaff *et al*., 1978; Honda *et al*., 1992; Meyer, 2009).

1.3.3 Mechanism of Replication

One of the defining features of plasmids is their ability to initiate replication autonomously from the host chromosome, and it is this ability that allows plasmids to exist and evolve as individual entities within diverse microorganisms. There are currently three general mechanisms for plasmid replication, namely theta-type, single-strand displacement and rolling circle replication as was reviewed extensively by Del Solar *et al.* (1998). Irrespective of the exact mechanism, replication occurs in three stages - initiation, extension and termination (Krüger *et al.,* 2004).

As mentioned previously, initiation of replication in IncQ plasmids requires the concerted effort of the *oriV* and the plasmid-encoded RepC, RepA and RepB proteins (Scherzinger *et al*., 1984). Once initiation of replication has occurred, the plasmid DNA is replicated by the DNA polymerase III holoenzyme continuously in either one direction initially (L strand), or both directions simultaneously (L and R strands) (De Graaff *et al*., 1978). A D-loop is initially formed during replication and as the single-stranded R-strand is displaced, a theta-type (θ) intermediate is visible by electron microscopy (Scherzinger *et al*., 1991). Termination of replication occurs when both the strands have been fully replicated, i.e. at the same position at which replication was initiated, and thus there are no specific termination signals (Kok *et al*., 1989). This single-strand displacement mechanism of replication, which lacks Okazaki fragment synthesis, is unique to IncQ-like plasmids. The mechanism by which IncQ plasmid replication at the *oriV* is initiated has, therefore, received considerable attention and the crystal structures of the RepA and RepB have been determined (Geibel *et al*., 2009; Niedenzu *et al*., 2001).

1.3.3.1 *The Functions of RepC, RepA and RepB*

RepC is a 31-kDa double-stranded DNA (dsDNA) binding protein responsible for initiation of replication at the *oriV* by inducing localized melting of the DNA strands (Kim and Meyer, 1985). It occurs as a dimer in solution but binds to the DNA as a monomer (Sakai and Komano, 1996). The RepA is a helicase protein responsible for unwinding of the dsDNA in a 5' $\rightarrow$ 3' direction (Scherzinger *et al*., 1997). It functions as a ring-shaped hexamer (29-kDa per subunit) with a 6-fold rotational symmetry when assembled around single-stranded DNA (ssDNA) (Fig. 1.2A). Binding of ssDNA stimulates hydrolysis of ATP to ADP which causes rotational forward movement and thus further unwinding of the dsDNA to produce more

ssDNA (Honda *et al.*, 1991; Niedenzu *et al*., 2001). Unlike the *E. coli* helicase and primase chromosomal equivalents, namely DnaB and DnaG respectively, the RepA helicase of RSF1010 does not interact with the RepB primase (Scherzinger *et al*., 1997).

The primase is responsible for synthesizing short complimentary oligonucleotides at specific initiation sites on the ssDNA which are extended by the DNA polymerase III holoenzyme during complimentary-strand DNA synthesis (Geibel *et al*., 2009). Two molecular forms, a smaller 38-kDa and larger 70-kDa protein, of the RepB are produced from the same coding sequence on RSF1010, but from different start codons. The large RepB protein is part of a relaxase-primase fusion protein and is, therefore, also referred to as MobA-RepB and is encoded by the *mobA/repB* fusion gene. The MobA-RepB fusion protein is required only during mobilization. The smaller primase, designated RepB', or *repB'* when referring to the gene, is only required during replication. The two primases can, however, substitute for each other during replication (Scherzinger *et al*., 1991). Interestingly, the RepB', which appears dumbbell-shaped as illustrated in Figure 1.2B, bears more structural similarity to archael Pri-type primases, which consist of a small catalytic PriS and large PriR subunit, rather than the multidomain DnaG-type primases from bacteria and bacteriophages (Geibel *et al*., 2009).

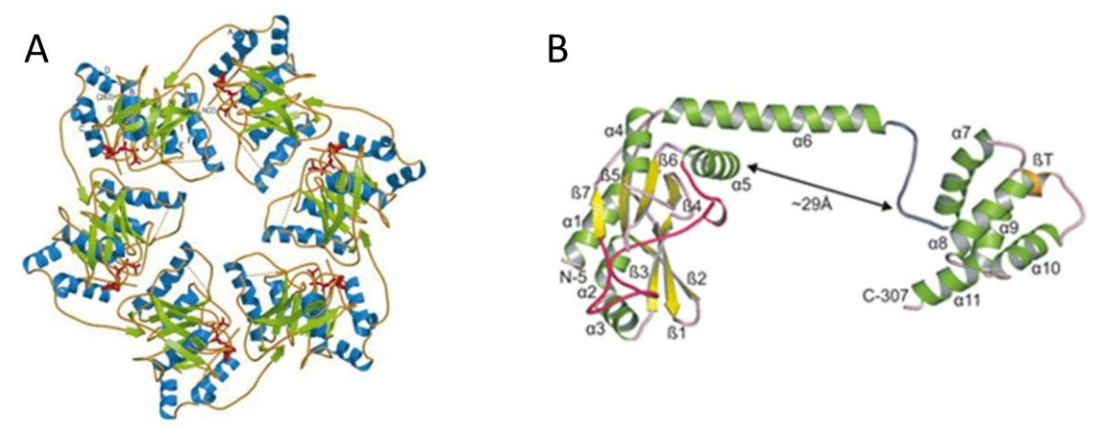


FIG. 1.2. Ribbon diagrams of the RSF1010 RepA helicase and RepB primase modelled from the three-dimensional crystal structures as determined by means of X-ray diffraction. (A) The crystal structure of the hexameric RepA was determined at a 2.4 Å resolution and shows that the monomers assemble to form an annular structure with a six-fold rotational symmetry around a central hole. The central hole has a diameter of ~17 Å and is too small to accommodate a dsDNA helix, and thus suggests that only ssDNA can thread through the hole. The ATP (red stick models) molecules required for helicase activity are located in the cleft between the monomers which are linked to each other at

the N-terminus. The figure was reproduced from Niedenzu *et al.* (2001). (B) The crystal structure of the smaller molecular form of the RepB' primase, which is only required during replication, was determined at a 2.0 Å resolution and shows the large dumbbell-shaped domain at the N-terminal linked to the small catalytic domain at the C-terminal by a 14 amino acid (aa) flexible tether. Comparing the structure of the RepB' to other primases revealed that the N-terminal domain is structurally more closely related to the PriS domain of the archael primase of *Pyrococcus furiosus* than it is to bacterial DnaG-type primases even though it lacks a zinc-binding motif and shares very low sequence homology (comparison and sequence alignment not shown). The figure was reproduced from Geibel *et al.* (2009).

1.3.3.2 *Architecture of the IncQ* oriV

The functional elements within the *oriV* of RSF1010 are three 20-bp direct repeats (DR) with 2-bp spacers, also known as iterons, followed by 25-bp of a G+C-rich region, 31-bp of an A+T-rich region, and a 152-bp palindromic region containing two inverted repeats (IR) known as single-strand initiation sites A and B (*ssiA* and *ssiB,* respectively) (Fig. 1.3) (Scholz *et al.*, 1989). When cloned, all of the RSF1010 DNA up to the first complete iteron can be deleted without affecting replication. Deletion of an additional 19-bp into the iteron results in an inability to replicate. On the right-hand-side of the cloned *oriV* DNA, deletion of the right-arm of the 152-bp palindrome resulted in a 2-fold reduced ability to replicate, and replication was completely abolished when an additional two thirds of the left-arm of the palindrome was deleted (Scherzinger *et al*., 1991). Later sequence comparisons of the RSF1010 *oriV* to that of other sequenced IncQ-like plasmids revealed that there is a 14 – 15-bp highly conserved region following the A+T-rich region, the function of which, however, remains cryptic (Rawlings and Tietze, 2001).

The iterons are the primary binding sites recognized by the RepC replication initiator and are unique to each plasmid in order to ensure plasmid identity amongst competing plasmids (Lin and Meyer, 1986; Rawlings and Tietze, 2001). When two plasmids with very similar or identical iteron sequences are co-resident within the same host, the replication proteins are unable to distinguish between the two *oriV*s during initiation of replication. This results in copy number fluctuations for one or both of the plasmids and thus the inevitable loss of one of the plasmids (Novick, 1987). The 20-bp directly repeated sequence of each iteron in an *oriV* is always fully conserved while the 2-bp spacers are not always conserved. A single

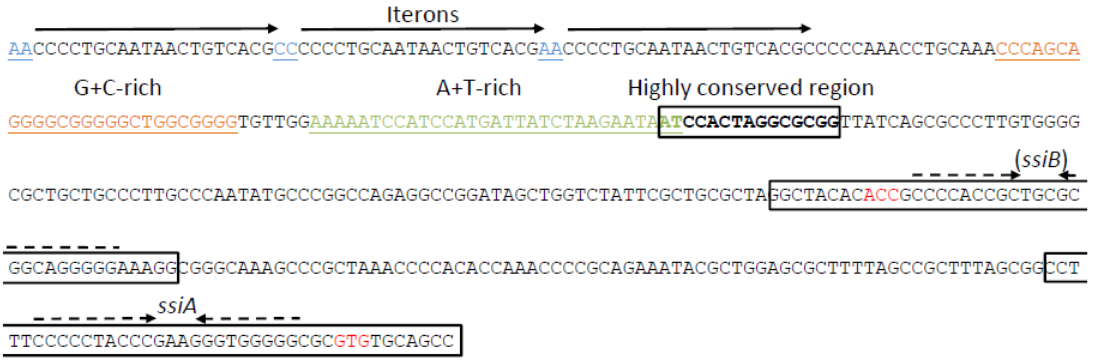


FIG. 1.3. Elements within the *oriV* of RSF1010. The conserved iterons are indicated by arrows while the variable 2-bp spacer sequences are blue and underlined. The G+C-rich area, indicated by brown underlined sequence, contains 23 G and C nucleotides within a 26-bp region, and the A+T-rich region, indicated by green underlined sequence, contains 24 A and T nucleotides within a 31-bp region. The highly conserved region is identical for RSF1010/R1162, pIE1107, pIE1130, pTF-FC2 and pTC-F14, except for one G-to-T transversion in pTC-F14 (Rawlings and Tietze, 2001). The exact position at which complimentary DNA synthesis is initiated (priming) at each of the *ssi*-sites is indicated in red. Priming at *ssiB* occurs on the complimentary strand and is therefore indicated in brackets.

point mutation within one of the three 20-bp conserved repeats of an R1162 *oriV* cloned into pBR322 resulted in an inability or significantly reduced ability, depending on where and in which iteron the mutation was located, to exert incompatibility against R1162 when placed *in trans*. Furthermore, when each of the mutated iterons was cloned as part of a functional R1162 *oriV* linked to a kanamycin resistance marker and tested for their ability to initiate replication when provided with RepC, RepA and RepB *in trans*, they were unable to do so except for one clone which contained a C·G-to-T·A transition in the 6th position of the outermost iteron (Lin *et al*., 1987). Similarly RSF1010 miniplasmids with either single or multiple point mutations within the iterons were unable to replicate when provided with the replication proteins *in trans*, except for two miniplasmids each of which had a single point mutation, T·A-to-G·C and A·T-to-C·G in the 16th and 20th position of the middle iteron, respectively. The RepC did, however, have a reduced binding affinity for the two mutant iterons which were still able to support replication, while it did not bind at all to those iterons incapable of supporting replication. Furthermore, when the two adenine residues of two of the 2-bp spacers in the RSF1010 iterons were substituted with two guanine residues the iterons were still able to support plasmid replication (Miao *et al*., 1995). Sequence conservation within the iterons is, therefore, critical for functional interaction with the RepC. However, individual nucleotides may have a unique, unequal, functional contribution.

Binding of RepC proteins to the iterons induces bending of the DNA, and the degree of binding is dependent on the total number of iterons bound by RepC proteins. Mobility shift assays showed that binding of a single iteron results in a 45 ° bend, binding of two iterons in a 65 ° bend and three iterons in a 75 ° bend, and that the conformational change induced by binding of all three iterons is a structural pre-requisite for melting of the double-stranded DNA (dsDNA) at the A+T-rich area (Miao *et al*., 1995). A specific base sequence within the A+T-rich region is needed for proper functioning as two mutants, one with a T·A-to-C·G transition at bp 22 and one with a G·C-to-C·G transversion at bp 26, exhibited increased thermal instability and could not be maintained at 42°C. The G·C-to-C·G transversion also caused a decreased replication frequency as the mutant replicon had a plasmid copy number (PCN) ~20% of that of the wild-type (WT) R1162 plasmid. A glycine-to-serine amino acid substitution within the RepC, however, resulted in restoration of the PCN. These results suggest that the RepC possibly also interacts with the A+T-rich region, although such an interaction would be weak as binding of RepC to this region could not be detected in gel retardation assays (Kim and Meyer, 1991).

Melting of the dsDNA at the A+T-rich region stimulates the RepA helicase monomers to assemble around the exposed ssDNA and catalyze further unwinding of the DNA in the 5' → 3' direction on both strands. It is thought that the G+C-rich region located between the iterons and the A+T-rich region, in addition to having a structural role in the bending of the DNA, momentarily halts unwinding of the R-strand until both the *ssiA* and *ssiB* initiation sites are exposed (Kim and Meyer, 1991). A stem-loop structure recognized by the RepB' primase and containing the start point for DNA synthesis at its base is formed by the IRs of the *ssi* sites (Fig. 1.4) (Honda *et al*., 1993; Miao *et al*., 1993). Although the IR sequence for the two *ssi* sites are not identical, the nucleotide sequences at the base of each of the stem-loop structures are highly conserved (Sakai and Komano, 1996). The two *ssi* signals could be exchanged, and one *ssi* signal placed on both strands in opposite orientations also yielded a fully functional replicon. Deletion of one of the *ssi* signals, however, resulted in a decrease in PCN and stability as a result of predominantly unidirectional plasmid DNA replication and the accumulation of ssDNA dimers (Tanaka *et al*., 1994; Zhou and Meyer, 1990). It has also been determined that *ssiA* of RSF1010 is more efficient than *ssiB* as double-stranded M13 phage DNA containing *ssiA* was accumulated more efficiently from the incoming ssDNA than when it contained *ssiB* (Yoichi *et al*., 1988). No sequences other than *ssiA* and *ssiB* that are analogous to the 9-bp *E. coli dnaA*-box, the G4-type DnaG recognition site or the θX174 primosome assembly site have been recognized in RSF1010 (Scherzinger *et al*., 1991). Either

one or both of the *ssi* signals could, however, be substituted with the *dnaA*-box, DnaG recognition site, ϕX174 or pACYC184 primosome assembly sites (Honda *et al.*, 1991; Taguchi *et al*., 1996). Initiation of plasmid replication is thus dependent on RepC, RepA and RepB', and replication of either strand is independent of the other. The RepB can, however, be substituted, provided the *ssi* signals were substituted with the corresponding primosome assembly sites.

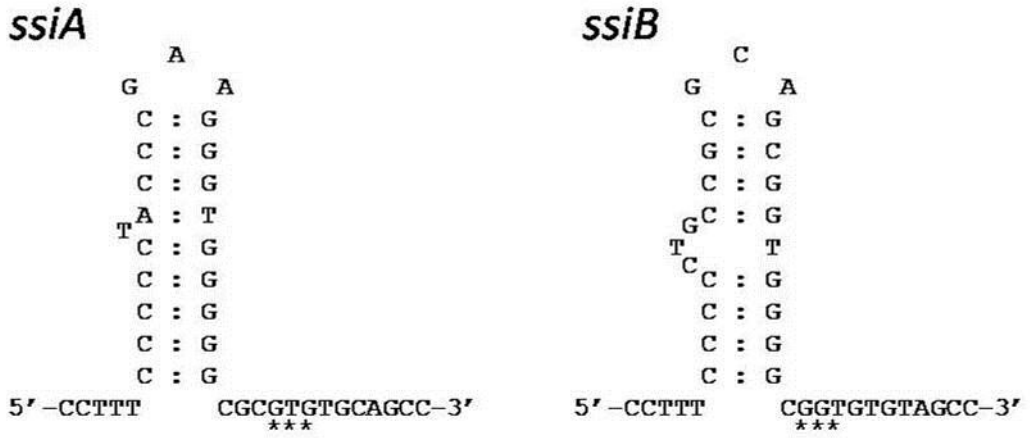


FIG. 1.4. Potential secondary stem-loop structures formed by the *ssiA* and *ssiB* sites within the *oriV* of RSF1010. The sites at which complimentary DNA synthesis is initiated are indicated by asterisks. The figure was adapted from Sakai and Komano (1996).

In summary and as illustrated in Figure 1.5, replication of IncQ plasmids is initiated when RepC monomers bind to the iterons. This results in a bend in the DNA structure that breaks the intramolecular forces between the base pairs at the A+T-rich region and exposes ssDNA (Miao *et al*., 1995, Sakai and Komano, 1996). The RepA monomers are then able to assemble into a hexameric helicase around the ssDNA and catalyze unwinding of the L-strand away from the iterons initially until both the *ssiA* and *ssiB* signals are exposed, then the R-strand as well (Niedenzu *et al*., 2001). The secondary stem-loop structures formed by the *ssi* signals are recognized by the RepB which then binds to and synthesizes complimentary strand primers at the initiation sites (Bryan *et al*., 1972; Yoichi *et al*., 1988). The DNA polymerase III holoenzyme assembles and synthesizes a complimentary DNA strand starting at the free 3'-OH groups provided by the primers on either of the strands, and continues in both directions until the entire strand has been replicated as described before (De Graaff *et al*., 1978; Sakai and Komano, 1996).

1.3.4 Regulation of Plasmid Copy Number

Strict regulation of PCN is essential for maintenance and spread during vegetative growth of the host. In the absence of a partitioning system, plasmids may not be inherited by daughter cells when the PCN is too low. A system that quickly initiates replication under low copy number conditions is therefore needed. If the PCN is too high the additional metabolic load carried by the host might reduce the growth rate and render it uncompetitive, thus also resulting in eventual loss from the population, and therefore, a system that prevents runaway replication is also needed (Bingle and Thomas, 2001). Each family of plasmids is thus maintained at a characteristic PCN by means of a functional balance between positive regulation exerted by the initiators at the *oriV* and negative regulatory feedback loops under control of various protein and or RNA regulatory elements at the promoters of the replication genes (Chattoraj, 2000; Del Solar *et al.*, 1998; Del Solar and Espinosa, 2000).

IncQ plasmids generally maintain their copy number at around 10 to 16 plasmids per chromosome (Rawlings and Tietze, 2001). There may or may not be differences in different bacterial hosts. The copy number of pTC-F14, for example, is 12 to 16 plasmids per chromosome in both *E. coli* and *A. caldus* while that of the RSF1010 equivalent R300B is 10 to 12 in *E. coli* and 29 to 34 in *P. aeruginosa* (Barth and Grinter *et al*., 1974; Gardner *et al*., 2001; Lewington and Day, 1986). Furthermore, a mutation rendering P_2, one of three promoters in the *oriT* region of RSF1010 negatively regulated by MobC and MobA-RepB, inactive results in an increase in PCN in both *E. coli* and *P. putida.* The change in copy number in *P. putida*, however, was not as pronounced as it was in *E. coli* (Frey and Bagdasarian, 1989). These results therefore demonstrate the likelihood of functional differences within IncQ-like plasmids in different hosts. As most regulatory studies on IncQ-like plasmids have been carried out in *E. coli*, it perhaps also demonstrates the need for comparative studies in different hosts. The copy number regulatory system of IncQ-like plasmids has not yet been completely elucidated but a number of regulatory systems, such as the promoter system and its negative regulators mentioned above, have been identified and are discussed below and are illustrated in Figure 1.6.

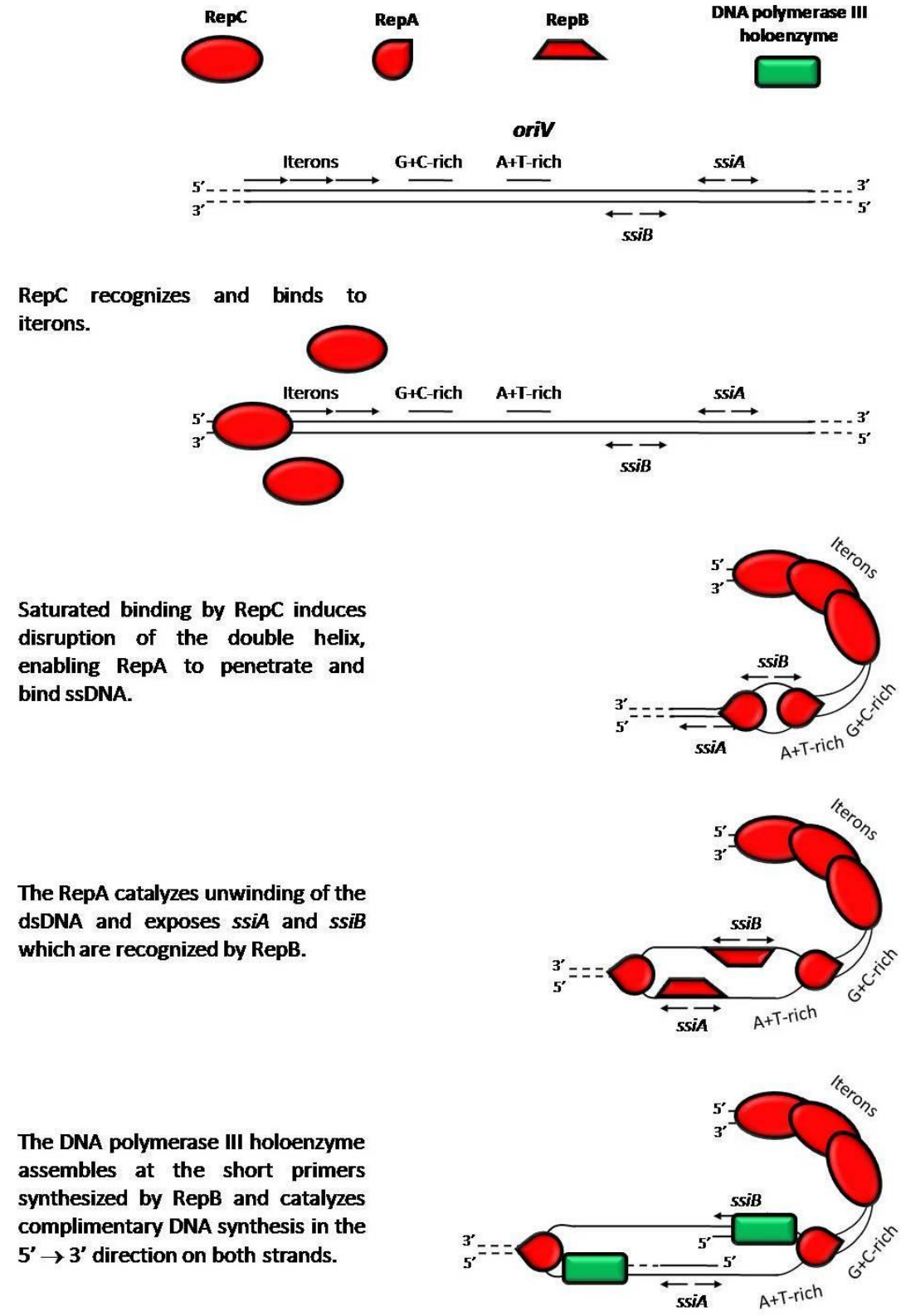


FIG 1.5. Initiation of replication at the *oriV* of IncQ plasmids. Plasmid replication is initiated at the *oriV* by the plasmid-encoded proteins (red). Complimentary strand synthesis by the host-encoded DNA polymerase III holoenzyme (green) takes place on both strands in opposite directions until both strands have been replicated.

The replication of IncQ-like plasmids is initiated and thus positively regulated by the binding of RepC monomers to the iterons as has been discussed earlier. RepC is co-transcribed together with RepA and it has been shown that over expression of the R1162 *repAC* genes *in trans* from behind a P_{Taq} inducible promoter results in a 2 – 3-fold increase in PCN (Kim and Meyer, 1985). Negative feedback regulation of the RSF1010 *repAC* operon at the upstream P_4 promoter is coordinated by a small 7.2-kDa protein encoded by *F*, otherwise known as *cac* for control of *repA* and *repC* (Maeser *et al.,* 1990). The *F* gene is the second of two genes that are located between P_4 and the 5'-end of the *repA* gene (Scholz *et al.*, 1989). The first gene, *E*, is also expressed, however, the function of the small 5-kDa protein remains undefined. A promoter fusion of P_4 to a β-galactosidase reporter system showed that repression of the P_4 promoter requires the presence of only F and not E. It was also shown that F forms a dimer of 17-kDa under saturating conditions. The P_4 promoter region contains two distinct, almost symmetrical 19 – 21-bp regions which are separated by 3 – 4-bp. Each of the regions contains an inverted repeat of variable length as well as an identical 10-bp direct repeat and it overlaps the putative +1, -10 and -35 RNA polymerase binding region. It is therefore likely, as shown by β-galactosidase assays and DNA footprint analysis of the P_4 promoter region, that when the concentration of F is high it forms dimers capable of binding the recognition sequences in the P_4 promoter and as such represses further transcription from P_4, and hence transcription of *repA* and *repC* (Maeser *et al.,* 1990).

Expression of RepA and RepC is not only negatively regulated at the transcriptional level by the activity of F at P_4, but also at the translational level by a small 75-bp silencing RNA. The small RNA is produced from a region overlapping the 5'-terminus of *repA* and is complimentary to the equivalent region on the mRNA transcript from which RepA and RepC is translated. A clone producing the small RNA inhibited the translation of a β-galactosidase reporter gene when it was fused to the P_4 promoter at the 8^{th} codon of the RepA. Conversely, a mutation 13-nt upstream of the first base of the small RNA, but which did not result in an amino acid substitution in RepA, resulted in decreased transcription of the small RNA and a consequent 33 – 44% increase in the copy number of R1162. The small RNA also protected a 75-bp piece of radio-actively labeled DNA which overlapped the ribosomal binding site (RBS) of *repA* during S1 nuclease mapping. It is therefore likely that at high concentrations (which would occur when the PCN is high and provided transcription levels of the small RNA is dosage dependent) of the small RNA effectively competes with the ribosomes for binding of the *repAC* mRNA transcript thus inhibiting translation of the two replication genes (Kim and Meyer, 1986).

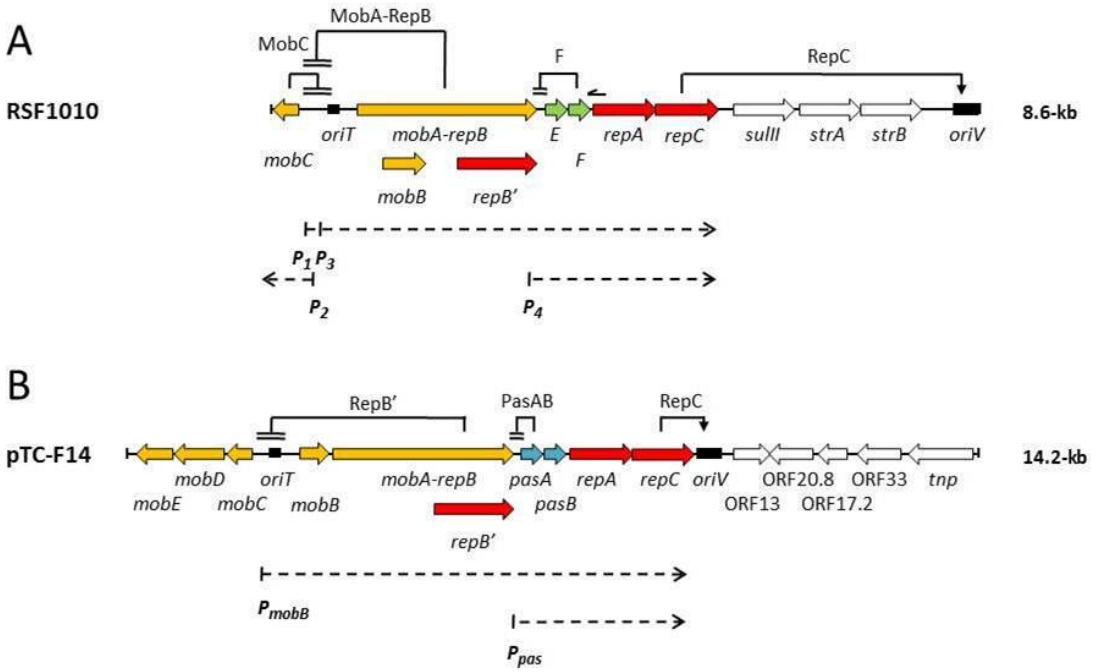


FIG. 1.6. Regulatory networks within the IncQ-1 and IncQ-2 plasmids (A) RSF1010 and (B) pTC-F14, respectively. The mRNA transcripts and the direction in which they are transcribed are indicated in broken lines below the genes included in the transcript. The negative regulators and the approximate regulatory positions at which they bind are indicated by double horizontal lines. The small silencing RNA produced by RSF1010 that binds and inhibits translation of the complimentary *repA* mRNA is indicated by a reversed half arrow. Positive regulators and the approximate position at which they bind are indicated by pointed arrows. Figure 1.6A was adapted from Meyer (2009) and Figure 1.6B from Gardner (2003).

The replication genes of RSF1010 are also regulated at the cluster of three partially overlapping promoters in the *oriT* region (Frey *et al*., 1992). Promoters P_1 and P_3 are responsible for transcription in the direction of the *mobB-mobA/repBB'* and *repAC* operon while P_2 is responsible for transcription of *mobC* (Derbyshire and Willetts, 1987; Scholz *et al.*, 1989). Deletion mutations affecting the negative regulatory activities at either P_1/P_3 or P_2 resulted in a 3-fold increase in PCN, and deletion of the mobA-*repB* (but not *repB'*) and *mobC* genes also resulted in a 4-fold increase in PCN (Frey *et al*., 1992). When RSF1010 or various deletion mutants were placed *in trans* of a β-galactosidase reporter system, in which a *Hin*fI DNA fragment containing the P_1 to P_3 promoter cluster was fused in either orientation to a *lacZ* reporter gene, it was found that both the MobA-RepB fusion protein and MobC are needed at the same time for maximal repression of transcription from the promoter cluster (Frey *et al*., 1992).

Expression of the replication genes of RSF1010 is therefore strictly controlled at two different regulatory regions. However, only the promoter cluster at the *oriT* is essential for plasmid maintenance. RSF1010 was able to replicate only if an artificial P_{lacUV5} promoter was provided for transcription in the direction of the *repB* and downstream genes when the mobilization region which includes the *oriT* and overlapping promoters was deleted (Katashkina *et al*., 2007). There are no transcription termination signals between the 5'-terminus of *mobA/repB* and the 3'-terminus of *repC* (Scholz *et al.*, 1989). The 3 to 4-fold increase in PCN as a result of upregulation of P_1 and P_3 as described above was therefore most likely the result of increased transcription of all the replication genes up to and including the *repC* gene, especially seeing as overproduction of RepB increased the PCN by only 1.6-fold (Frey *et al*., 1992; Haring *et al*., 1985; Meyer, 2009). Further evidence that the RepA and RepC proteins of IncQ-like plasmids are translated from the same transcript as the RepB was provided by studies on pTC-F14 and pTF-FC2. These two IncQ-2 plasmids have a functional toxin-antitoxin (TA) system, consisting of *pasAB* and *pasABC* respectively, between the *repB* and *repAC* genes instead of genes similar to *E* and *F* of RFS1010 (Deane and Rawlings, 2004; Smith and Rawlings, 1997). In spite of this, northern blot hybridizations and reverse transcriptase PCR (RT-PCR) analysis showed that pTC-F14 produces a large ~5.7-kb polycistronic mRNA transcript that originates upstream of the *mobB* (which precedes the *mobA/repB* fusion gene) and terminates downstream of the *repC*. The promoter from which the ~5.7-kb transcript originates was maximally repressed (57% expression relative to unregulated expression levels as determined by a β-galactosidase reporter enzyme assay) when the RepB domain was present *in trans* compared to 92% when the mobilization proteins, excluding the RepB domain of MobA-RepB, were present. Therefore, the RepB-domain rather than the MobA-domain of the fusion protein is the negative regulator of the *mobB* promoter (Gardner and Rawlings, 2004). The *repAC* operon of pTF-FC2 (and pTC-F14), like that of RSF1010, does not have its own promoter immediately upstream of the *repA* and deletion of the entire *pasABC* operon did not affect the plasmid's ability to replicate nor did it cause a change in PCN (Matcher and Rawlings, 2009).

Even though the promoters of the two TA systems are not required for expression of the *repAC* replication genes, they have an important function with regards to plasmid maintenance in addition to their function in plasmid stability (discussed in section 2.6.2). Reporter gene studies on the various promoters of pTC-F14 showed that the two strongest promoters are the promoters of *pasAB* and *mobB*, and with the *pasAB* promoter being 100-fold stronger than the *mobB* promoter (Gardner and Rawlings, 2004). Titration of the pTF-

FC2 PasA-PasB complex by having the promoter region present *in trans* on a high copy number vector resulted in upregulation of P_{pasABC} and a 2.7-fold increase in expression of *repAC*. Also, when the *repAC* genes of pTF-FC2 were deleted and instead supplied *in trans* from behind a concentration-dependent arabinose inducible promoter, it was found that the PCN increased rapidly in response to small increases in induction and quickly reached an upper limit (Matcher and Rawlings, 2009). For these reasons it is thought that at a low copy number such as after cell division or mobilization to a new host, autorepression of the *pasABC* operon is relieved, thus resulting in an increased production of RepA and RepC and a corresponding increase in the frequency of initiation of replication (Matcher and Rawlings, 2009). In support of these findings, it was noticed that the PCN of R1162 remained unchanged when the P_4 promoter was deleted, but that the plasmid was less successful at becoming established after transformation into a new host (Meyer, 2009). It therefore seems likely that the promoter between the 3'-terminus of *repB* and 5'-terminus of *repA* in these IncQ-like plasmids exists not to control the upper copy number limit, but rather to allow for rapid expression of the replication genes when the copy number is too low, thereby allowing the plasmid to become established in a new host and ensuring rapid replication after cell division.

1.3.4.1 *IncQ Plasmid Copy Number Regulation Differs From Other Iteron-Containing Plasmids*

The most extensive work on the copy number control mechanism of iteron-containing plasmids has been done on the low copy number (1 – 2 plasmids per chromosome) theta-replicating plasmids F, R6K and prophage P1 (McEachern *et al.,* 1985; Tsutsui *et al*., 1983; Sozhamannan and Chattoraj, 1993). As with the IncQ-like plasmids, these plasmids require binding of a plasmid-encoded initiator protein to iterons within the *oriV* for initiation of replication, however, there are no other plasmid-encoded proteins such as, for example, a helicase or primase involved in initiation or replication. The necessary enzymatic activities for replication are supplied by the host bacterium (Masai and Arai, 1989; Sozhamannan and Chattoraj, 1993). Plasmid replication may be initiated from up to three *oriV*s, however, a specific *oriV* may be used preferentially (Banerjee *et al.,* 1992; Crosa, 1980). They also often have a second set of direct repeats within the promoter area of the initiator gene which is referred to as the control locus. The 22-bp direct repeats within the control locus share consensus with the iterons within the *oriV* and also serve as binding sites for the initiator,

thereby allowing for autoregulation (Del Solar *et al.*, 1998). In prophage P1, the iterons of the *oriV* overlap with the operator region of the initiator promoter and therefore have a dual regulatory role (Das *et al.,* 2005).

The copy number regulatory system for both initiation and inhibition of replication of these plasmids is better understood than that of IncQ-like plasmids, especially with regards to inhibition of replication. The upper copy number limit of these plasmids is considered to be regulated by a mechanism known as handcuffing, the general principles of which are illustrated in Figure 1.7. The initiator protein has both DNA and protein binding domains, and occurs as either a monomer or a dimer. If the PCN is low, monomers bind to and saturate the iterons in a cooperative manner, thereby disrupting base-pairing in the double helix similar to IncQ-like plasmids and initiating replication (Bowers and Filutowicz, 2008; Bowers *et al*., 2007). If, however, the PCN is high, two iteron-bound monomers are bound by a dimer and a tetrameric monomer-dimer-monomer protein bridge is formed. The bimolecular pairing, or handcuffing structure, inhibits replication through steric hindrance, and can be formed between iteron-bound monomers within the *oriV* and the control locus, or between iteron-bound monomers on separate *oriV*s. As these plasmids have multiple iterons within the *oriV*, and multiple *oriV*s on a plasmid, a small change in PCN translates to a change in iteron concentration and as such leads to increased sensitivity in the negative feedback loop which maintains the PCN within narrow limits (Chattoraj, 2000).

It was demonstrated that the copy number of a miniP1 plasmid decreased by 75% when a second identical *oriV* was cloned into the plasmid and that this decrease in PCN was not overcome when extra initiator was supplied *in trans*, thus demonstrating an effect explained by handcuffing rather than titration of the initiator by the additional iterons (Park *et al.,* 2001). Similarly, when a second set of iterons was cloned into a novel position on the IncQ plasmid R1162 it resulted in a decrease in PCN. An excess supply of the RepA and RepC proteins by expressing the *repAC* genes from behind an IPTG-inducible promoter *in trans*, in contrast to what was observed with miniP1, resulted in a 6-fold increase in PCN. Furthermore, an excess supply of the RepA and RepC proteins also enabled R1162 to overcome an incompatible phenotype when an identical *oriV* was cloned on a multicopy vector *in trans* (Kim and Meyer, 1985). These results differ from the predicted behaviour of plasmids that control their copy number by means of handcuffing, and suggest that the upper copy number limit of IncQ-like plasmids is controlled by rate-limitation of the

synthesis of the three replication proteins rather than steric hinderance by means of handcuffing.

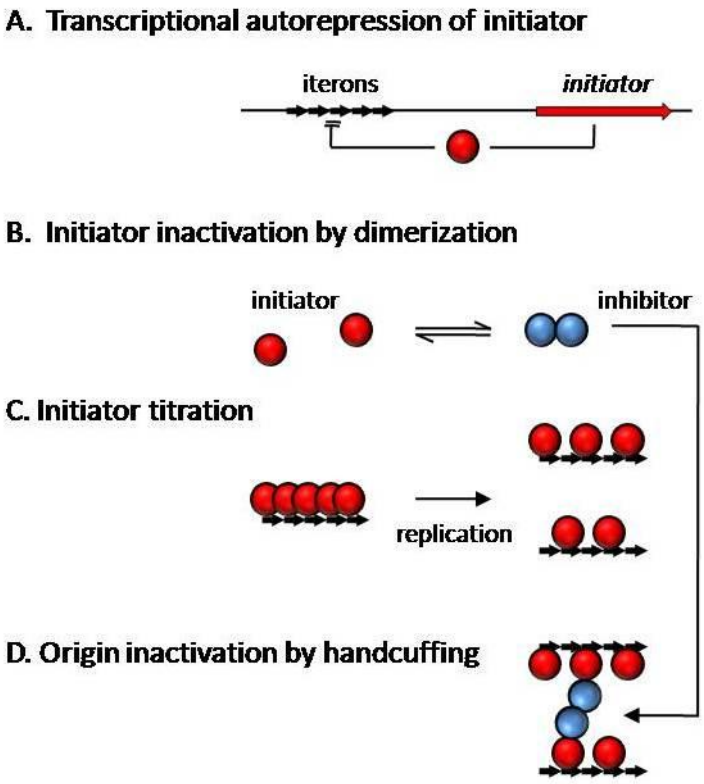


FIG. 1.7. The principles of handcuffing. (A) In plasmids like prophage P1 the initiator promoter overlaps with the origin iterons and monomer binding to the iterons thus serves a dual role in autorepression of the initiator promoter and initiation of replication. (B) At a higher concentration monomers dimerize and are thus unavailable for iteron binding. (C) Initiation of replication requires saturated binding of initiator monomers to the iterons. After replication the monomers are titrated by the iterons in the origins of the daughter plasmids such that none of the origins are saturated and replication is thus inhibited. (D) Dimerized initiator proteins bind iteron-bound monomers thus pairing origins in a handcuffing structure which further decreases the probability of replication. The figure was adapted from Paulsson and Chattoraj (2006).

By definition, copy number control by rate-limitation would mean, however, that plasmid replication would continue to take place indefinitely until one of the metabolites or replication components, be it plasmid or host-encoded, became rate-limiting. Copy-up mutants, such as the well-known pUC-cloning vectors in which the negative copy number regulator Rom (also known as Rop) has been deleted, have a copy number in excess of 400 plasmids per chromosome and thus demonstrate that the host-encoded factors can support replication at such high frequencies (Cesareni *et al.,* 1991; Lee *et al*., 2006b; Lin-Chao *et al.,*

1992). As mentioned earlier, deleting the operator regions in RSF1010 responsible for negative regulation of the replication genes resulted in a mere three-fold increase in PCN. Furthermore, an excess supply of RepA and RepC from a source *in trans* of RSF1010 resulted in only a six-fold increase in PCN, although it must be granted that the RepB which was not included and could have been rate-limiting, (Frey *et al*., 1992; Haring *et al*., 1985). These results suggest that another mechanism other than negative regulation at the promoter regions exists in order to control specifically the upper copy number limit. The small silencing RNA identified for R1162 (Kim and Meyer, 1986) is a possible candidate, however, a similar silencing RNA has not yet been identified in any of the other IncQ-like plasmids.

The observation that RepC proteins exist as dimers in solution (Sakai and Komano, 1996) provides a possibility for upper copy number regulation. It could be that at a high intracellular concentration the RepC will dimerize, although without forming handcuffing structures at the *oriV*, and thereby limit the amount of monomers available for replication (see Figure 1.7B and C only). Such a phenomenon would explain why over-expression of the replication proteins induced a relatively small increase in PCN as well as why, unlike for 'handcuffing plasmids', an excess supply of RepC enabled the plasmid to overcome an incompatibility phenotype when an identical *oriV* was present on a vector *in trans*. Finally it would also explain why the copy number of pTF-FC2 reached a plateau fairly quickly and then decreased in response to increasing levels of RepA and RepC (Matcher and Rawlings, 2009). With this in mind, it is evident that the copy number control system of IncQ plasmids at the upper limit (inhibition) is different to that of other iteron-containing plasmids and is poorly understood compared to control at the lower limit (initiation), and thus requires more attention.

1.3.5 Plasmid Incompatibility

Plasmid incompatibility, as summarized by Novick (1987), is a function of the relatedness of plasmid replicons. Plasmids with closely related replicons are unable to coexist within the same host cell in the absence of selection, whereas replicons that have had sufficient time to diverge usually are compatible. The inability to coexist is as a result of the inability of the replication machinery to distinguish between the two plasmids when selecting a template for replication (Fig. 1.8). If one plasmid is selected for replication more frequently than the other, either by chance (symmetrical incompatibility) or due to interference from one of the

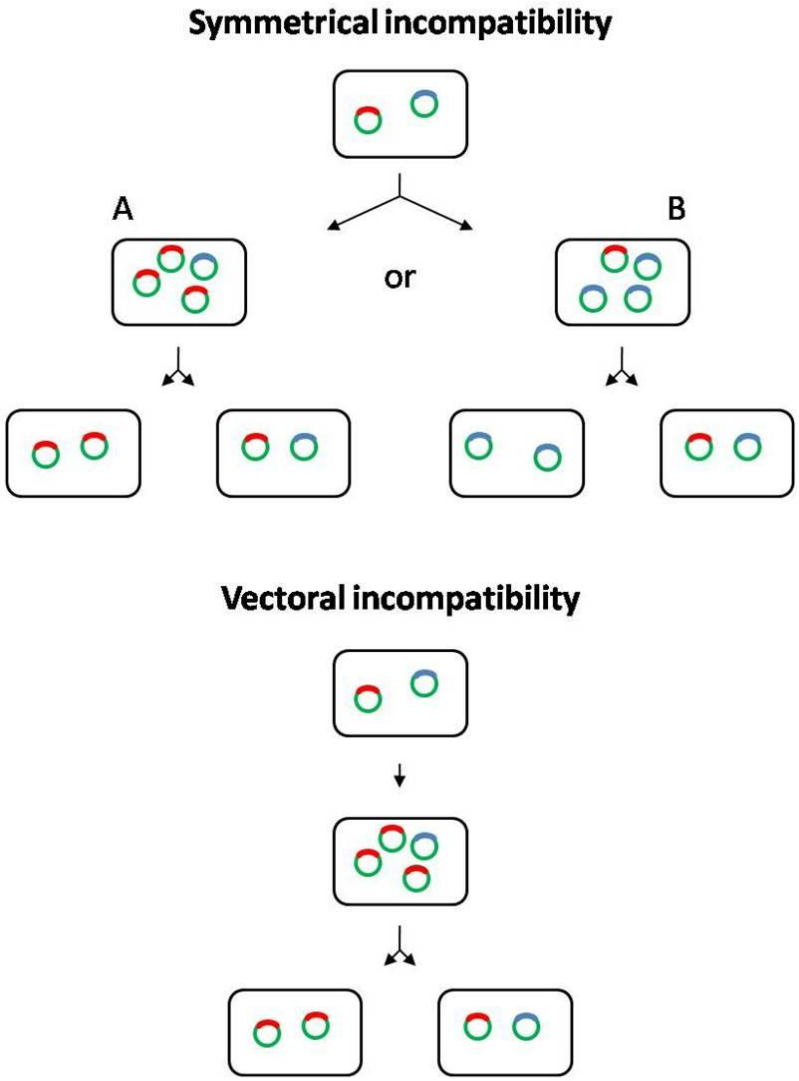


FIG. 1.8. The general principle of replicon-associated plasmid incompatibility. The green circles represent related plasmid replicons within bacterial cells and the red and blue regions represent different accessory DNA for each of the two plasmids. Symmetrical incompatibility occurs when the closely-related replicons are selected at random for replication, thus giving rise to A or B. During cell division the plasmids are segregated to give rise to subpopulations containing either the red plasmid, the blue plasmid or both plasmids. Vectoral incompatibility occurs when one plasmid is replicated preferentially above the other (red plasmid), thus giving rise to two subpopulations, one containing only the red plasmid and one containing both plasmids. The percentage of cells containing both plasmids will become smaller with every generation, thus eventually giving rise to a homogenous population (not shown in figure).

plasmids (vectoral or asymmetrical incompatibility), it may result in loss of the less-frequently replicated plasmid from the host. Incompatibility can also be as a result of the inability of related partitioning systems to distinguish between two coresident plasmids, resulting in an asymmetrical segregation pattern. As will become evident in the next section, incompatibility as a result of related partitioning systems is not relevant to IncQ plasmids and will, therefore, not be discussed any further. Of importance, however, is that a weak form of incompatibility exists between pTF-FC2 and pTC-F14, although it was found to be as a result of cross-talk between the TA systems, and will thus be discussed in the section dealing with the TA systems of IncQ plasmids (see section 2.6.3) (Paulsson and Chattoraj, 2006).

The iterons in iteron-containing plasmids are in most instances the primary incompatibility determinants as they are able to titrate the initiator protein and thus inhibit replication (Chattoraj, 2000; Novick, 1987). To demonstrate that the same principle is valid for IncQ plasmids the iterons from R1162 were cloned into a multicopy vector and placed *in trans* of R1162, whereafter the cells were allowed to replicate in the absence of selection before checking for plasmid retention. In doing so it was also demonstrated that the strength of incompatibility was dependent on the number of cloned iterons *in trans*. A vector containing 3 cloned iterons was able to displace R1162 more effectively than a vector containing only 2 cloned iterons, whereas a vector with 1 iteron was unable to displace a coresident R1162 plasmid. By observing the effect that the cloned iterons had on the relative copy number of R1162 it became evident that the cloned iterons caused a decrease in PCN, and correlating with the incompatibility status, the extent of the decrease depended on the number of iterons. A vector with 3 iterons caused the most significant decrease in the copy number of R1162, whereas a single iteron had only a small but visible effect. A vector containing either partial or no iterons had no effect on the stability or copy number of R1162 (Lin and Meyer, 1986). Similar results were obtained when chemically synthesized 36-bp oligonucleotides containing repeat sequences identical to the R1162 iterons were cloned in multiple copies (1, 2 and 3) into a vector and placed *in trans* of R1162. Furthermore, *in vitro* replication of R1162 was effectively inhibited in the presence of the cloned iterons but not in the presence of vector alone. Again the extent of inhibition depended on the number of cloned iterons *in trans*. A competitive model for the expression of incompatibility was formulated from these results. It holds that the copy number of an IncQ plasmid depends on the number of successful RepC-iteron interactions, that iterons *in trans* reduce the amount of RepC available for replication thus resulting in a decrease in

PCN, and finally that in a portion of the population the PCN becomes so low as a result of the titration that not all daughter cells inherit the plasmid (Lin *et al*., 1987).

As replicon-associated plasmid incompatibility is a function of relatedness it can be used, in addition to DNA sequence homology, as a means to group plasmids into families, hence the term 'Incompatibility group' prior to the letter assigned to plasmid families such as IncF, IncP, IncQ, IncU and IncW, for example. The incompatibility status of most of the IncQ-like plasmids discussed thus far has been determined in relation to other IncQ members (Rawlings and Tietze, 2001). A symmetrical pattern of segregation was observed when pIE1120 was coresident with RSF1010 and similarly when pIE1115 was coresident with pIE1107 (Smalla *et al*., 2000). Plasmids pIE1107, pIE1115 and pIE1130 are unable to coexist with RSF1010. Sequence analysis of pIE1107, however, revealed that it contains a second non-functional *oriV,* referred to as *oriVa*, that is similar to the *oriV* of RSF1010, and when deleted, to give pIE1108, the two plasmids were compatible (Smalla *et al*., 2000; Tietze, 1998). The non-functional *oriV* contains three iterons that are identical to the iterons of RSF1010, however, a 19-bp region has been deleted that was found in RSF1010 to be essential for replication (Tietze, 1998). It was later found that both pIE1130 and pIE1115 also contain the same non-functional RSF1010-like *oriV* and that deletion of this *oriV* renders the plasmids compatible with RSF1010 (Rawlings and Tietze, 2001). The pIE1107 derivative pIE1108, as an incoming plasmid in reciprocal transformation experiments, rapidly displaced pIE1130, and pIE1130 was not able to establish in a host containing a resident pIE1108 (Smalla *et al*., 2000). This unidirectional displacement is different to the symmetrical incompatibility described for pIE1115 and pIE1108. However, it was not tested whether this aggressive and unidirectional incompatibility phenotype was as a result of a titrative effect of the pIE1108 iterons (*ori-Qb* iterons of pIE1107) only, or whether there were other contributing factors such as non-productive binding of pIE1108 proteins to replication or regulatory regions on pIE1130.

The two IncQ-2 plasmids pTF-FC2 and pTC-F14 are compatible with each other when coresident. Incompatibility phenotypes were, however, observed when these plasmids were coresident with pIE1108 (pIE1107) or RSF1010. Both resident pTF-FC2 and pTC-F14 plasmids were displaced by pIE1108 as an incoming plasmid, but neither were able to displace a resident pIE1108 during reciprocal transformation experiments. RSF1010, as an incoming plasmid, was able to displace pTC-F14, but not pTF-FC2, and neither were able to displace RSF1010 (Gardner *et al*., 2001). It was later demonstrated during complementation studies

using RSF1010 and pTC-F14 that it was probably non-productive binding of the RepA or RepC proteins of RSF1010 rather than its *oriV* that was responsible for the incompatibility phenotype with pTC-F14 (Gardner and Rawlings, 2004). Whether this is true for the combinations which included pIE1108 and pTF-FC2 is not known as it was never tested.

1.3.5.1 *Evolution of Incompatibility Groups*

The paradox with regards to the evolution of new incompatibility groups from existing ones is that the genes or loci involved in incompatibility specificity are normally involved in essential plasmid maintenance functions and are thus intolerant to mutations (Sýkora, 1992). To demonstrate this point (as discussed in the section 2.3.2 dealing with the structure and function of the *oriV*), it was shown that single and multiple point mutations within the iterons of R1162 and RSF1010 rendered not only just the one iteron, but the entire *oriV* inactive. Mutations that did not render the iterons completely inactive were the exception and although such mutations did not abolish replication they did adversely change the binding affinity for RepC (Lin *et al*., 1987; Miao *et al*., 1995). Point mutations within the AT-rich region of the R1162 *oriV* also had adverse effects on replication at higher (42°C) temperatures. The effects of one specific mutation, which also caused a 20% relative decrease in copy number, could, however, be suppressed by a second-site mutation within the *repC* gene (Kim and Meyer, 1991). Non-lethal mutations within critical regions can thus be tolerated as long as they are complimented by a second-site mutation either in the *cis*-acting locus or *trans*-acting gene.

With this in mind, the second non-essential *oriV* in each of pIE1107, pIE1130 and pIE1115 presents an interesting scenario for the evolution of IncQ incompatibility groups (Tietze, 1998; pIE1130 and pIE1115 unpublished sequence). As mentioned previously, the second *oriV* of pIE1107, *ori-Qa*, is presumed to be non-functional as deletion of the same region in a cloned RSF1010 *oriV* rendered the *oriV* inactive (Tietze, 1998). Sequence analysis shows that the equivalent *oriV* in pIE1115 contains a number of point mutations further downstream, while that of pIE1130 is identical to the *oriV* of RSF1010, respectively (alignments not shown). Whether these two *oriV*s can support replication in the presence of the RSF1010 replication proteins remains to be tested. As these non-essential iterons in each case are identical to that of RSF1010 they probably were, however, responsible for titrating the RSF1010 RepC during the incompatibility studies. The functional replicon-specific iterons of

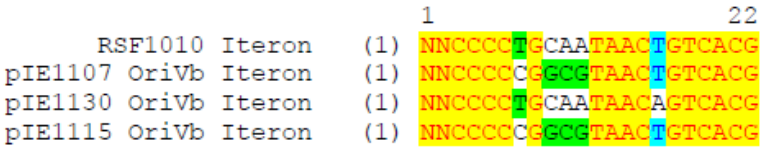
pIE1107, pIE1130 and pIE1115, on the other hand, differ from the 20-bp conserved region of the RSF1010 iterons by 4, 1 and 4-nt, respectively, and the respective RepC proteins share 91.5, 88.4 and 91.5% aa sequence identity. This explains, at the nt and aa sequence level, why these plasmids have been assigned to incompatibility groups different from that of RSF1010 (Fig. 1.9). The functional iterons of pIE1115, however, are identical to that of pIE1107 and so too are the aa sequences of the two RepC proteins, although at the nt level they share only 98.9% (9 point mutations) identity (data not shown). This demonstrates at the sequence level why the two plasmids exhibit symmetrical incompatibility when co-resident and is the reason why pIE1107 and pIE1115 have been grouped into the same IncQ-1β-subgroup.

In each case, the essential *oriV* is separated from the non-essential *oriV* by different antibiotic resistance genes. It is, therefore, not hard to imagine that acquisition of new antibiotic resistance genes in the ancestral plasmid by means of recombination could have resulted in duplication of the *oriV*. From here on point mutations could have accumulated in one of the *oriV*s and were eventually matched by a second-site mutation in the corresponding locus of the *repC* gene, as was demonstrated experimentally for the A+T-rich region and *repC* of R1162 by Kim and Meyer (1991). Duplication of the *oriV* followed by compensatory mutations is, therefore, one of the possible methods by which new incompatibility groups could evolve within the IncQ family.

1.3.6 Stability of IncQ-Like Plasmids

One of the requirements, in addition to replication and conjugal transfer, for the establishment, maintenance and spread of plasmids is the ability to be inherited efficiently by each daughter cell during vegetative cell division (Stewart and Levin, 1977). Large low copy number plasmids such as RK2, with a PCN of 5 – 8 per chromosome, the F plasmid, R1 and prophage P1, all with a PCN of 1 – 2 per chromosome, cannot rely on random distribution to ensure vertical inheritance (Engberg and Nordstrom, 1975; Frame and Boshop, 1971; Prentki *et al*., 1977; Roberts and Helinski, 1992;). Considering a PCN of 5 plasmids per chromosome such as for RK2, the theoretical probability that a plasmid might not be inherited in a daughter cell based on random distribution is 1 in every 16 cells per generation (Williams and Thomas, 1992). These plasmids therefore employ a combination of partitioning, toxin-antitoxin (TA) and multimer resolution systems which work together to

A



B

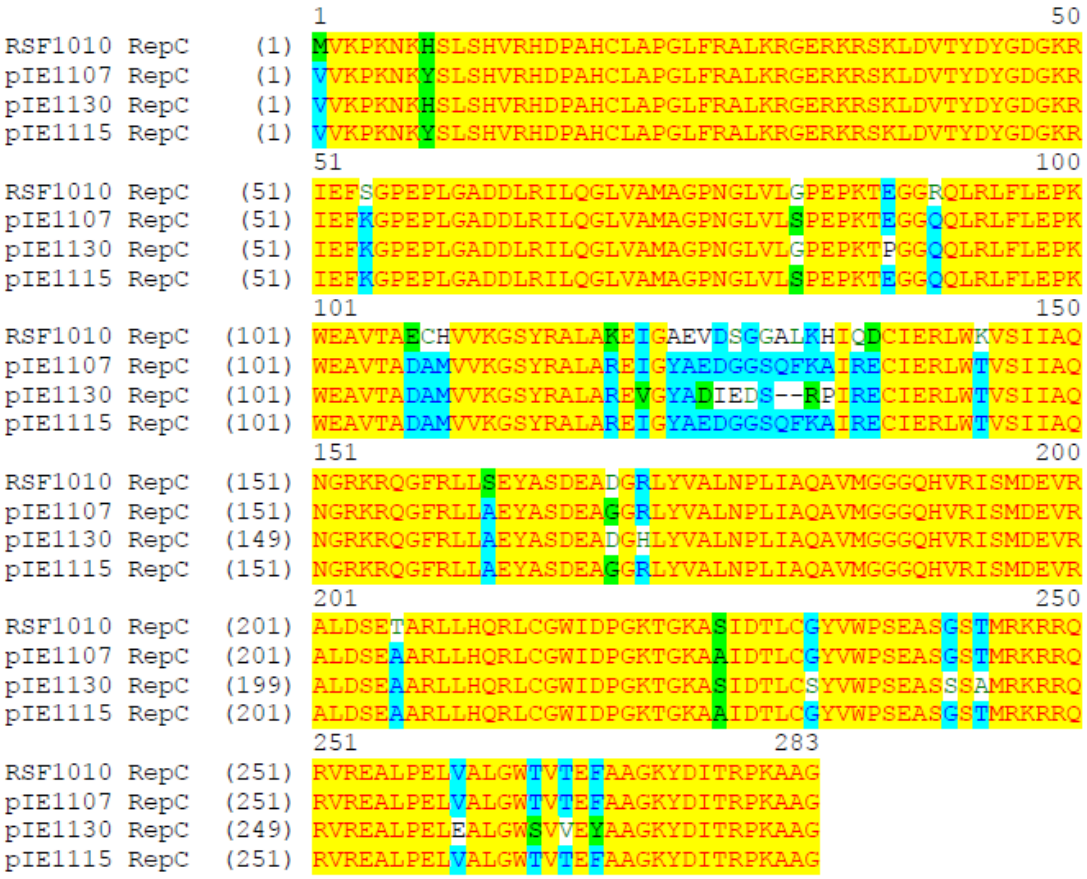


FIG. 1.9. **Sequence alignments demonstrate the differences between the iterons and RepC proteins of closely-related IncQ-1 plasmids.** (A) Only one of the 22-bp functional *oriV*b-iterons of pIE1107, pIE1130 and pIE1115 are aligned to one of the RSF1010 iterons. The *oriV*a iterons of pIE1107, pIE1130 and pIE1115 are identical to the RSF1010 iterons (not shown). The 2-bp spacers are included as NN nucleotides. (B) The start codons of the RepC proteins of pIE1107, pIE1130 and pIE1115 identified by sequence analysis are located as much as up to 32-codons upstream, in the case of pIE1130, of the first aa in the alignment and the sequence in this upstream region is highly variable. Therefore, only the region of aa sequence that bears homology to the RepC of RSF1010 was included in the alignment.

achieve loss frequencies of less than 10^{-6} per cell per generation (Boe *et al.*, 1987; Gerdes *et al.*, 1985; Nordstrom and Austin, 1989; Roberts and Helinski, 1992). Small high copy number plasmids rely mainly on their high copy number and random distribution for stable maintenance during cell division. Other plasmids such as ColEI, pSC101 and pTF-FC2, contain a recombination site for multimer resolution, a partitioning- or a TA system to increase stability within the population, respectively (Nordstrom and Austin, 1989; Smith and Rawlings, 1997).

Thus far there have been no reports of an IncQ-1-like plasmid containing a stability system even though these plasmids are stably maintained in a broad variety of hosts. Their relatively high PCN of between 10 and 16 plasmids per chromosome, depending on the plasmid, is thought to be sufficient for stable inheritance (Rawlings and Tietze, 2001). Assuming a PCN of 10 plasmids per chromosome and that it doubles prior to cell division and assuming that the segregational stability of IncQ plasmids is completely random, then according to the formula $P_0 = 2^{1-n}$ (where n is the number of plasmid copies per cell at cell division), the probability (P_0) of a plasmid-free segregate arising is only once in 5×10^5 cell divisions (Williams and Thomas, 1992). The high PCN, could however, increase the frequency with at which plasmid multimers are formed when the plasmids are resident in *recA*-proficient hosts, although this has not yet been demonstrated for IncQ plasmids. No sites similar in function to the *cer*-site on ColEI plasmids have yet been reported for IncQ-like plasmids. Whether multimers are formed between sister IncQ-like plasmids, and if so, how they are resolved requires investigation.

In contrast to the IncQ-1 plasmids, the two members of the IncQ-2 group each have a *pas* (plasmid-addiction system) TA system located between their *repB* and *repA* genes. Insertion of the *pas* systems upstream of the *repAC* genes in such a way that the regulatory activities of the TA systems contribute to the initiation of replication under low copy number conditions probably provided a significant selective advantage (Matcher and Rawlings, 2009; Rawlings and Tietze, 2001). The *pas* system of pTF-FC2 is different from that of pTC-F14, as well as most other TA systems, in that it consists of three genes, *pasABC*, encoding three proteins, namely the PasA antitoxin, PasB toxin and PasC helper protein, respectively. In spite of the additional PasC of pTF-FC2, these two plasmids are clearly related as the PasA and PasB proteins share 81% and 72% aa sequence identity, respectively (Deane and Rawlings, 2004).

1.3.6.1 *The Mechanism by Which Toxin-Antitoxin Systems Confer Plasmid Stability*

The TA systems that have been discovered thus far have been divided into diverse families, however, they all share a common theme in that they consist of a long-lived toxin and a short-lived antitoxin (Van Melderen *et al*., 2009). Although the functions of chromosomally-encoded TA systems are still subject to debate (Pandey and Gerdes, 2005), at least a large proportion of them have been shown to play a role in the stringent response system of bacteria by inhibiting growth during periods of nutrient limitation. TA systems are not however restricted to chromosomes. As a result of their addictive nature TA systems have been captured by a number of plasmids where the toxins inhibit growth of the host in response to plasmid loss rather than nutrient starvation (Rawlings, 1999).

The TA families are divided into two types. The antitoxin of Type I TA systems is a small 60 – 70 nucleotide (nt) antisense RNA that inhibits translation of the toxin through silencing of the toxin-mRNA. The promoters of these systems are usually constitutively expressed but the toxin proteins are only produced when the silencing RNA is no longer transcribed (Fozo *et al*., 2008; Van Melderen *et al*., 2009). The antitoxin and toxin of the Type II TA systems are both proteins and binding of the antitoxin to the toxin results in inactivation of the toxin (Fig. 1.10). The genes are organized in an operon in which the antitoxin is transcribed first and at a higher frequency than the toxin. Although the antitoxin is the primary repressor its efficiency is increased when it is bound to the toxin (Gerdes *et al*., 2005). In the event of plasmid loss the antitoxin is quickly degraded either as a result of being an unstable RNA or, in the case of the Type II proteic antitoxins, by host-encoded proteases such as Lon or Clp (Gerdes *et al*., 2005). As a consequence the long-lived toxin is no longer inactivated and thus inhibits growth of the host by, depending on the type of toxin, causing damage to the cell membrane, interference with the DNA gyrase and thus causing the formation of double stranded breaks in the DNA, or by inhibiting translation either through cleavage of the mRNA (Zhang *et al*., 2004) or preventing translocation of the peptidyl tRNA (Bahassi *et al*., 1999; Gerdes *et al*., 1986; Jiang *et al*., 2002; Liu *et al*., 2008; Zhang *et al*., 2004).

TA systems, unlike partitioning systems, therefore do not ensure plasmid stability during vegetative growth of the host but rather inhibit the growth of plasmid-free segregants so that they do not outcompete the plasmid-containing hosts and thereby ensure plasmid maintenance within the population. Furthermore, plasmids containing a TA system are also less likely to be displaced from a host population by competing plasmids during horizontal

transferand are able to colonize hosts previously occupied by plasmids lacking a TA system (Cooper and Heinemann, 2000).

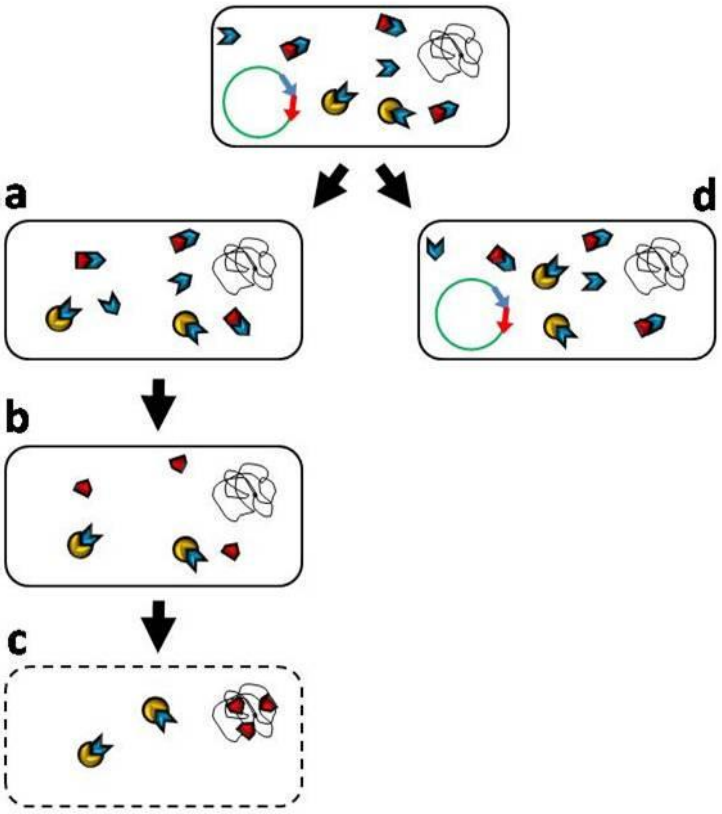


FIG. 1.10. A schematic representation of Type II TA-mediated growth inhibition in response to plasmid loss. The antitoxin (blue) is expressed at a higher frequency than the toxin (red) from the TA-operon (blue and red arrows). The higher concentration of antitoxin is needed as it is continually degraded by protease enzymes (yellow). In the event that a plasmid (green) is not inherited during cell division (a) the antitoxin is rapidly degraded by the protease and is no longer available to neutralize the toxin (b). The toxin is then free to interfere with host-DNA replication or gene expression and results in growth inhibition or cellular death (c). If, however, the plasmid is inherited (d) the antitoxin is continuously expressed and available to neutralize the toxin and growth of the host is not affected. Adapted from Rawlings (Rawlings, 1999).

1.3.6.2. *The Toxin-Antitoxin Systems of pTF-FC2 and pTC-F14*

Both *pas* systems were effective at inhibiting the growth of plasmid-free *E. coli* segregants when cloned on an unstable pOU82 heterologous replicon. The *pasABC* system of pTF-FC2 was more effective than the *pasAB* of pTC-F14 as after ~100 generations 93% of the

population retained the pOU82 plasmid containing the pTF-FC2 *pasABC* system compared to 60% for the *pasAB* system of pTC-F14. Inactivation of the *pasC* gene of pTF-FC2, however, resulted in a decrease in effectiveness to below that of pTC-F14 as only 47% of the *E. coli* population retained the test plasmid and it was suggested that the PasC increases the effectiveness with which the PasA antitoxin neutralizes the PasB toxin (Deane and Rawlings, 2004; Smith and Rawlings, 1997). The efficiency of the pTF-FC2 *pasABC* system was also found to be strain dependent as well as dependent on the presence of Lon protease. The apparent stability of the pOU82 replicon containing *pasABC* was increased 100-fold in *E. coli* CSH50 and only 2.5-fold in *E. coli* JM105 compared to when the TA system was absent and it was completely ineffective in *E. coli* JM107 and JM109. The *pasABC* system was also found to be ineffective in an *E. coli lon*-mutant compared to a *lon*-proficient strain thus suggesting that the Lon protease is involved in degradation of the PasA antitoxin (Smith and Rawlings, 1998a).

As discussed previously, the promoters of both *pas* systems are autoregulated by the PasA-PasB complex (Gardner and Rawlings, 2004; Smith and Rawlings, 1998b). The pTF-FC2 PasA is capable of autoregulating its promoter in the absence of PasB, however, β-galactosidase assays showed that the efficiency of repression by PasA on its own is only a quarter of that of the PasA-PasB complex (Smith and Rawlings, 1998b). The antitoxins, in combination with their respective toxins, of either *pas* system were able to repress the promoter of the other although the PasA of pTF-FC2 was twice as effective at repressing both its own promoter as well as the promoter of the pTC-F14 *pas* system. The PasC of pTF-FC2 did not have a noticeable influence in the cross regulation of the pTC-F14 *pas* promoter. It was also shown that pTF-FC2 is able to displace pTC-F14 when these plasmids are coresident in the absence of selection and that this displacement is mediated by the PasA (Deane and Rawlings, 2004). The displacement of pTC-F14 is probably the consequence of cross-regulation between the *pas* systems and provides pTF-FC2 with a selective advantage over pTC-F14 in the event that these two plasmids are coresident within a single host, a scenario which is not unlikely as these two plasmids were isolated from the same environment.

1.3.7 IncQ Mobilization

Large conjugative plasmids encode all the genes required for efficient plasmid transfer from the donor bacterium to the recipient bacterium. That is, all of the proteins required for

mating pair formation (Mpf) as well as DNA transfer and replication (Dtr) are provided for by the plasmid itself (Lawley *et al*., 2004). IncQ plasmids, however, encode only the Dtr genes and rely on conjugative plasmids to provide the Mpf components. The occurrence of IncQ-like plasmids is therefore often associated with the occurrence of larger conjugative plasmids. For example, the pRAS3 plasmids were always found to be co-resident with a large conjugative IncU plasmid, pRAS1, irrespective of the *A*. *salmonicida* strain from which they were isolated and irrespective of the geographical location (Aoki and Takahashi, 1986; L'Abee-Lund and Sørum, 2002). Two plasmids belonging to the IncN family and one belonging to the IncPα family were isolated together with pIE1107, pIE1120, pIE1115 and pIE1130 from the piggery manure bacterial cultures (Smalla *et al*., 2000). R1162 was mobilized from *P. aeruginosa* 1162 to *E. coli* at low frequencies and although a conjugative plasmid was not isolated at the time, a transfer factor responsible for pilus formation was present (Bryan *et al*., 1972).

Although there was no conjugative plasmid present in the *E. coli* host from which RSF1010 was isolated, it was shown to be mobilized at equally high frequencies from both *E. coli* and *P. aeruginosa* by IncP plasmids, relatively efficiently from *E. coli* by IncF, IncFVI, Inclα, IncM and IncX plasmids, and less efficiently by IncN or IncW plasmids (Cabezon *et al*., 1994; Francia *et al.*, 2004; Guerry *et al.,* 1974; Willitts and Crowther, 1981;. A 27.6-kb plasmid was associated together with pTF-FC2 in the *At. ferrooxidans* host from which it was isolated, however, this plasmid was never captured and it is not known whether it was a conjugative or mobilizable plasmid (Rawlings *et al*., 1984). Nonetheless, even though no conjugative partner plasmids were identified at the time of isolation for pTF-FC2, pTC-F14 and pDN1, they were all efficiently mobilized by RP4, an IncP conjugative plasmid (Van Zyl *et al.*, 2003; Whittle *et al.,* 2000; Rawlings *et al*., 1986).

1.3.7.1 *Two Different Mobilization Systems for the IncQ Family*

The replication systems of all the IncQ plasmids discussed thus far consist of an *oriV* and three conserved genes, the *repB*, *repA* and *repC* encoding the primase, helicase and the initiator, respectively. Their mobilization systems, however, are different in that RSF1010 and the RSF1010-like plasmids such as pIE1107, pIE1130, pIE1120 and pIE1115 have a mobilization system consisting of only three genes, namely *mobA*, *mobB* and *mobC*, while that of pTF-FC2 and pTC-F14 consist of five genes with *mobD* and *mobE* being the extra two

genes. The IncQ-like plasmids which contain a mobilization system similar to that of RSF1010 are grouped as IncQ-1 plasmids while pTF-FC2 and pTC-F14 currently make up the only members of the IncQ-2 group. Common to both these mobilization systems is that (as mentioned previously) a *repB* gene encoding a primase domain which is active in both replication and mobilization, albeit in different forms, is fused to the *mobA* gene (Del Solar *et al.*, 1998). The rest of the mobilization genes are also organized in a similar manner relative to each other. In both systems, the *mobC*, and the *mobD* and *mobE* of the IncQ-2 plasmids, are transcribed divergently to the *mobB* and *mobA/repB* genes from a cluster of promoters at the *oriT*, which is central to both systems (Rohrer and Rawlings, 1992; Scholz *et al.*, 1989; Van Zyl *et al.*, 2003).

1.3.7.2 *Components of The IncQ-1 Mobilization System and Their Respective Functions*

The *oriT* is the site at which the relaxosome, consisting of all the proteins involved in mobilization, is assembled prior to conjugal transfer of the plasmid DNA and is also the site at which transfer is terminated (Brasch and Meyer, 1987; Bhattacharjee *et al.,* 1992). The R1162 *oriT* was characterized as a 38-bp region containing an IR, with each arm consisting of 10-bp separated by 3-bp, followed by an 8-bp conserved region which leads up to the *nick*-site (Fig. 1.11) (Brasch and Meyer, 1987). The *oriT*s of the IncQ-1 plasmids are highly conserved at the *nick*-site, whereas the *oriT*s belonging to the IncQ-2 plasmids, although still similar in structure to that of the IncQ-1 plasmids, belong to a group of *oriT*s which includes the IncP plasmids RP4 and R751 (Rawlings and Tietze, 2001).

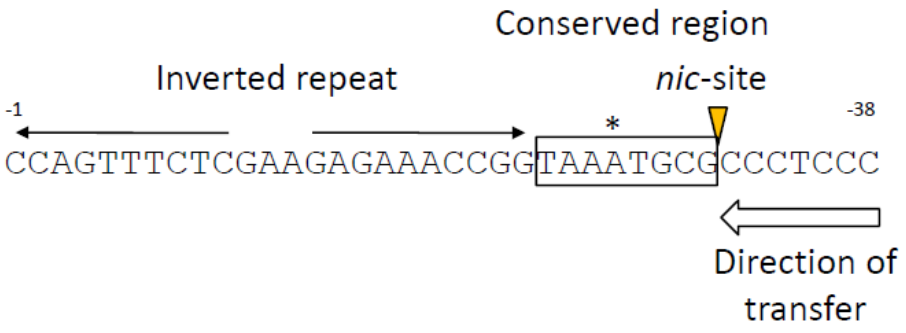


FIG. 1.11. Origin of transfer of R1162. The *oriT* of R1162 is identical to that of RSF1010 except for the unique G-to-A transition at position 27 (marked with an asterisk) in R1162. Adapted from Meyer (Meyer, 2009).

The amino-terminal MobA domain of the MobA-RepB fusion protein is responsible for both nicking of the *oriT* as well as recircularization of the linear ssDNA upon entry into the recipient host (Bhattacharjee and Meyer, 1991; Scherzinger *et al.,* 1993). Only the strand to be transferred is nicked and forms a reversible tyrosyl phosphodiester bond with the MobA at its 5'-end. The 5'-end of the ssDNA is protected from λ-exonuclease digestion in the presence of MobA and remains covalently bound to the single stranded *oriT* DNA during transfer (Bhattacharjee and Meyer, 1991). Although both reactions are thought to be performed by the same MobA molecule, the interactions of the MobA with the *oriT* during nicking and religation are different as the IR, specifically only the outer arm, is required for religation and not for nicking. It was found that only eleven nucleotides on the 5'-end of the RSF1010 *nick*-site, which thus excludes the IR, are required for recognition and cleavage by the MobA (Scherzinger *et al.,* 1993). On the other hand, deletion of the outer arm of the IR of R1162 decreased the mobilization frequency 100-fold, while deletion of the inner arm completely abolished transfer (Brasch and Meyer, 1987). Based on these findings, and that the actual structure of the hairpin-loop formed by the IR was found to be more important than the exact sequence, it was proposed that the role of the hairpin-loop, located on the 3'-end of the transferred strand, is to capture the MobA attached to the 5'-end and thereby ensure ligation after the single stranded linear DNA enters the recipient (Brasch and Meyer, 1987; Bhattacharjee *et al.,* 1992).

The *mobB* gene of the IncQ-1 plasmids is transcribed from within the *mobA* coding sequence, although in a different reading frame, whereas in the case of the IncQ-2 plasmids it is transcribed from an open reading frame upstream on the *mobA/repB* gene and is approximately 100 nucleotides shorter (Rohrer and Rawlings, 1992; Scholz *et al.*, 1989; Van Zyl *et al.*, 2003). An in-frame deletion of the *mobB* gene of R1162, such that it did not affect the overlapping *mobA* gene, resulted in a smaller proportion of nicked plasmids, thought to be as a result of decreased relaxosome stability, and a corresponding 100-fold decrease in the mobilization frequency. A slightly larger deletion removing a portion of the *mobA* gene also resulted in decreased stability (Perwez and Meyer, 1999). Similarly, a frame-shift mutation in the *mobB* gene of pTF-FC2 resulted in a 4,500-fold decrease in the mobilization frequency (Rohrer and Rawlings, 1992). It is therefore thought that the MobB specifically interacts with the MobA to ensure stability of the relaxosome.

Inactivation of the R1162 MobC resulted in a 50-fold decrease in the mobilization frequency and transfer became very sensitive to the levels of gyrase within the cell. When tested in an

E. coli with a thermo-sensitive gyrase mutation the mobilization frequency (at a non-permissive temperature) was not influenced by the levels of gyrase in the cell if the MobC was present. In the absence of MobC, however, the mobilization frequency was decreased more at a non-permissive temperature than it was at a permissive temperature. Primer extension assays were used to show that MobC enhances strand separation at the *nic*-site. In the presence of MobC the *oriT* DNA strands were separated and nicked by permanganate-induced cleavage, resulting in multiple early termination products. In the absence of MobC the *oriT*-DNA remained largely double-stranded and was no longer sensitive to permanganate-induced cleavage, thus only a predicted single termination product was observed (Zhang and Meyer, 1997).

During mating only ssDNA is transferred and thus the missing strand must be replicated before the plasmid can become established (Parker *et al*., 2002). Inactivation of the carboxy-terminal primase domain of the MobA-RepB by insertion mutagenesis of the *mobA*/*repB* fusion gene resulted in a loss of mobilization. Similarly, mobilization was abolished when the *oriV* was entirely deleted as well as when the *ssi*-site on the transferred strand was cloned in the incorrect orientation. These results demonstrate the requirement for the primase during mobilization. Additional experiments in which the RepB primase domain was detached from the MobA relaxase domain also demonstrate that the MobA-RepB fusion was required for optimal mobilization. This was achieved by making various deletions of the *mobB* gene, that is transcribed from within the *mobA* but in a different frame. As the MobB is also required for mobilization it was supplied *in trans*. Detachment of the two domains did not affect the activity of either of the domains but it did result in a 99% decrease in the mobilization frequency compared to when the MobA-RepB fusion protein was present (Henderson and Meyer, 1996). That the MobA-RepB fusion protein is required for optimal mobilization and that it is probably transferred along with the ssDNA was confirmed by showing that RepB in the recipient could not substitute for an absence of MobA-RepB in the donor (Henderson and Meyer, 1999).

1.3.7.3 *Regulation of the Mobilization Genes*

As mentioned previously a cluster of three promoters, P_1 to P_3, located in the *oriT* region of RSF1010, is responsible for expression of the divergently transcribed *mobB-mobA/repBB'* and *mobC* genes and all three promoters are negatively regulated by binding of the MobC

and MobA-RepB at the *oriT* (Frey *et al*., 1992). A regulatory region in the same location was also shown to be present in both pTF-FC2 and pTC-F14 (Gardner and Rawlings, 2004; Matcher and Rawlings, 2009). By means of 2D electrophoresis, it was shown that molecular coupling occurred between two R1162 *oriT*s, especially when a large amount of positively supercoiled DNA was present. Molecular coupling was observed only when all three mobilization proteins were present. Plasmids lacking an *oriT* but still containing the cluster of three promoters were also deficient for coupling. However, when the copy numbers of these plasmids were determined relative to coupling-proficient plasmids no significant difference was found (Zhang and Meyer, 2003). More than one plasmid molecule is transferred during a conjugational event, and therefore, the possibility exists that coupling of the *oriT*s by the relaxosome occurs so as to promote multiple rounds of transfer by localizing the plasmids at the pore (Parker and Meyer, 2002; Zhang and Meyer, 2003).

1.3.7.4 *The IncQ-2 Mobilization System Is Similar to the Tra1 Dtr System Of RP4*

Sequence comparisons of the *oriT* and five genes comprising the mobilization regions of pTF-FC2 and pTC-F14 by Van Zyl *et al* (2003) revealed that these systems are more similar in sequence and organization to the Tra1 Dtr system of RP4 than they are to the unique mobilization system of the IncQ-1 plasmids. The *oriT*s of both pTF-FC2 and RP4, like all other *oriT*s, contain an IR repeat, followed by a highly conserved *nic* site. On average, the MobA, excluding the primase domain, MobB, MobC and MobD mobilization proteins of pTF-FC2 have 44 to 52% similarity but only 25 to 33% identity to the TraI, TraJ, TraK and TraL proteins of RP4, respectively. MobE is the least conserved with 38% similarity and only 15% identity to its TraM counterpart (Rohrer and Rawlings, 1992). The MobA-RepB and MobB proteins of pTC-F14 are 75 and 78% identical, respectively, to the pTF-FC2 equivalents, however, the MobC, MobD and MobE proteins only have 27, 40 and 21% identity to the pTF-FC2 counterparts, respectively, and they have even less homology to their RP4 counterparts (Van Zyl *et al.*, 2003).

The functions of the MobA, MobB and MobC proteins of pTF-FC2 and pTC-F14 are thought to be relatively similar to that of RSF1010 with regards to relaxosome formation and stabilization (Lawley *et al*., 2004; Rohrer and Rawlings, 1992; Van Zyl *et al.*, 2003). The additional MobD and MobE proteins are not essential for mobilization, however, if they are deleted their absence has a negative influence on the mobilization frequency of both

plasmids. A construct containing the pTF-FC2 mobilization region with the *mobDE* genes deleted was mobilized by RP4 at a 1 500-fold lower frequency than a construct containing the WT system (Rohrer and Rawlings, 1992). Similarly, although not as drastically, a construct containing the pTC-F14 mobilization region was mobilized at a 600-fold lower frequency when the *mobD* and *mobE* genes were deleted compared to when the genes were intact (Van Zyl *et al.*, 2003).

1.3.7.5 *General Model for Mobilization*

Of all the conjugative systems tested thus far, IncP conjugation seems to be the most efficient at mobilizing the IncQ plasmids (Willits and Crowther, 1981; Rawlings and Tietze, 2001). For this reason most of the details in this section will be discussed with reference to the RP4 conjugative system. Successful conjugation relies on assembly of the transferosome which is responsible for pilus synthesis, the bacterial sex apparatus within which the plasmid DNA is transferred between donor and recipient cells, as well as processing of the plasmid DNA by the relaxosome (Brasch and Meyer, 1987). A coupling protein connects the relaxosome with the transferosome. The components of the transferosome are encoded by the Mpf genes while the relaxosome components and the coupling protein are encoded by the Dtr genes (Lawley *et al*., 2004).

The Mpf system of RP4 consists of 16 genes, *trbA* – *trbP*. Only *trbA* – *trbL* are required for a functional Mpf system during conjugative transfer of RP4 between *E. coli* cells and the system is similar to the 12-component VirB/VirD4 transport system of the Ti plasmid of *A. tumefaciens*. VirB1 – VirB11 represent the Mpf components and VirD4 the coupling protein (Lessl *et al.,* 1992b; Schröder and Lanka, 2005). As a result of the extensive similarity between the two systems many of the functions of the Trb proteins of RP4, as well as the proteins involved in Mpf of other conjugative or Type IV secretion systems, have been derived from studies on the VirB/VirD4 system and have been summarized by Schröder and Lanka (2005) (Lessl *et al.,* 1992b; Lawley *et al*., 2004). Some of the protein functions include, in no specific order, perforation of the peptidoglycan cell wall, structural components of the pili and outer membrane complex, components responsible for adhesion, an ATP-dependent secretion motor embedded within the secretion machinery, a cytoplasmic NTPase which

TABLE 1.2. Summary of the functions of shared components of the VirB/VirD4 and IncPα Mpf systems as well as of the IncPα and IncQ-2 Dtr systems

Function	Protein component		
	***A. tumefaciens* VirB/D4**[A]	**IncPα (RP4)**[B]	**IncQ-2 (pTF-FC2; pTC-F14)**[C]
Mating pair formation			
Perforation of peptidoglycan cell wall	VirB1		
Pilus structural subunit	VirB2	TrbC	
Outer-membrane pore component	VirB3	TrbD	
Inner-membrane NTPase	VirB4	TrbE	
Adhesin-like	VirB5	TrbJ	
Modulator of secretion channel	VirB6	TrbL	
Lipoprotein connecting pilus to core	VirB7	TrbH	
Periplasmic core component	VirB8	TrbF	
Outer-membrane anchor	VirB9	TrbG	
Channel regulator	VirB10	TrbI	
Cytoplasmic NTPase	VirB11	TrbB	
Acetylase		TrbP	
DNA transfer[D]			
Relaxase	VirD2	TraI	MobA
Facilitates relaxase binding		TraJ	MobB
Strand separation		TraK	MobC
Unknown		TraL	MobD
Unknown		TraM	MobE
Coupling protein	VirD4	TraG	

[A] Function of the VirB/D4 secretion system was reviewed by Shröder and Lanka (2005)

[B] The similarities within different transport systems were compared by Lawley *et al.* (2004).

[C] Only the DNA transfer components with homologs to the IncQ-2 mobilization proteins were included.

[D] Sequence analysis of pTF-FC2 and pTC-F14 genes was done by Van Zyl *et al.* (2003).

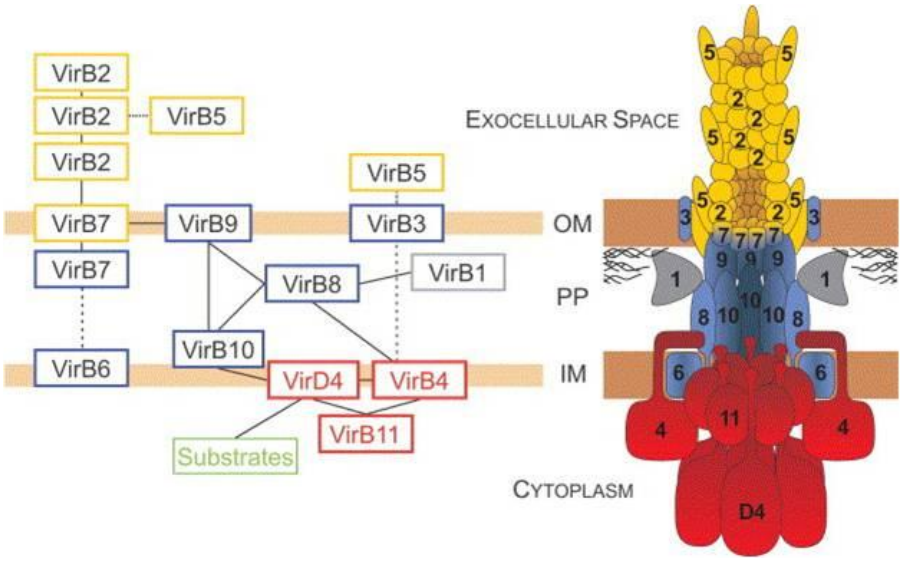


FIG. 1.12. A schematic representation of a Mpf system based on the Type IV VirB/VirD4 secretion system of *A. tumefaciens.* The components responsible for providing energy to the complex are illustrated in red, the core complex components are blue, peptidoglycanases in grey and the outermembrane- and pilus-associated components are yellow. The lines connecting the components in the diagram on the left represent interactions that have been determined experimentally. For a list of the functions of the different components consult Table 1.2. A detailed mechanistic description of the model is given in the review by Schröder and Lanka (2005) from which this illustration was reproduced.

provides additional energy for the secretion machinery, as well as energy and secretion modulators (Table 1.2 and Fig. 1.12). Only the TrbB, TrbC, TrbE, TrbG and TrbL, as well as the TraG coupling protein and TraF of the Dtr system are absolutely essential for mobilization of RSF1010 between *E. coli* cells (Lessl *et al.,* 1993).

Processing of the IncQ plasmid DNA for transfer is mediated by the Dtr components described earlier and involves two steps (Fig. 1.13) (Schröder and Lanka, 2005). First the MobA, together with the MobC, binds to the *oriT* and disrupts the dsDNA to form ssDNA (Zhang and Meyer, 1997). The MobB also binds in order to stabilize the relaxosome. The ssDNA is then nicked by the MobA by means of a transesterification reaction, resulting in MobA being covalently bound to the 5'-end while the 3'-OH group remains free (Bhattacharjee and Meyer, 1991; Scherzinger *et al.,* 1993). Next, the plasmid DNA is unwound as it passes through the pilus complex, provided by the conjugative plasmid, in a 5' $\rightarrow$ 3' direction. At the end of this process the hairpin-loop at the 3'-end hooks the MobA, presumably still attached to the pilus components in the recipient as well as to the 5'-end, and the ends of the linear DNA are covalently linked to each other by means of the ligase

activity of MobA to form a circular single stranded copy of the plasmid (Becker and Meyer, 2002; Parker *et al*., 2002). As the MobA is fused to the RepB primase, primers are synthesized at the *ssi*-site on the transferred strand and the DNA is replicated as discussed previously (Henderson and Meyer, 1996; Henderson and Meyer, 1999; Parker *et al*., 2002).

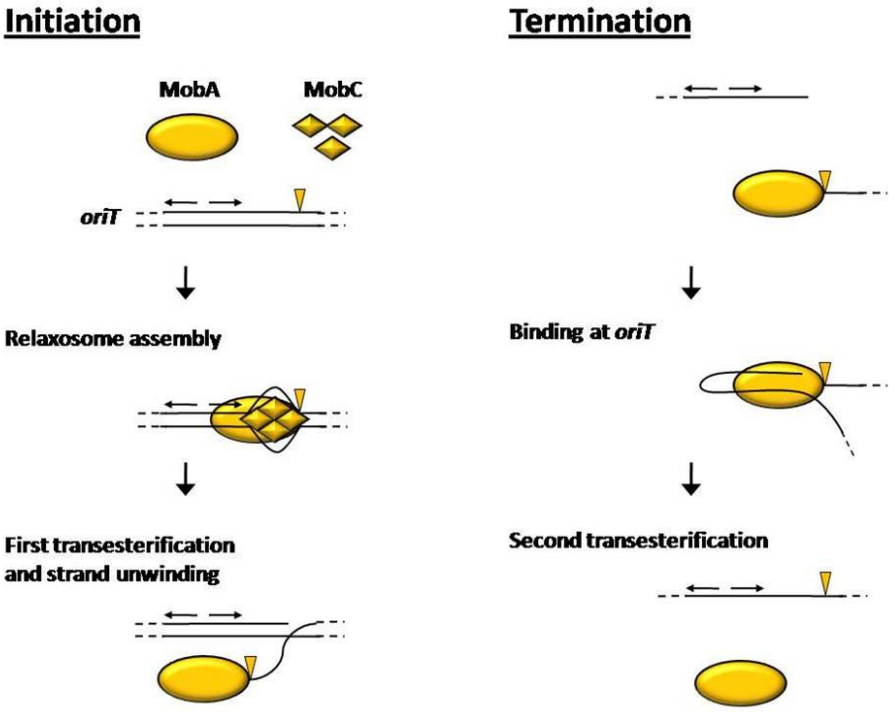


FIG. 1.13. Initiation and termination of mobilization occurs at the *oriT*. During initiation the MobA binds to the dsDNA at the *oriT*, initiates formation of the relaxosome and nicks the strand to be transferred at the conserved *nick*-site, indicated by a triangle, and forms a tyrosyl phosphodiester bond with the 5'-end. The MobC assists in strand separation. The MobB, not shown in the diagram, interacts with the MobA to stabilize the relaxosome. Upon termination after a round of transfer the two arms of the IR, indicated by divergent arrows, hybridize to form a dsDNA complex that is recognized by the MobA, which is still covalently attached to the 5'-end of the ssDNA, and the two ends of the strand are rejoined by a second transesterification reaction. The figure was reproduced from Becker and Meyer (2002).

1.3.7.6 *Evolution of IncQ Mobilization Systems*

The Tra2 Mpf- and Tra1 Dtr systems of RP4, for example, are ~11-kb and ~15-kb, respectively (Lessl *et al.,* 1992a). Due to the large size of conjugative systems the lack of a Mpf system in IncQ plasmids is probably as a result of the size limitation imposed on IncQ

plasmids by their single-strand displacement mechanism of replication (Rawlings and Tietze, 2001; Meyer, 2009). Furthermore, only five of the seventeen Tra1 genes are absolutely essential for the processing and transfer of RP4 (Lessl *et al.,* 1993). Considering the mobilization system of the IncQ-1 plasmids this number can be further reduced to as little as three genes, and again, a reduction in the number of genes, and hence size, was probably selected for by the mechanism of replication.

It was also discussed in an earlier section that the BHR capabilities of IncQ plasmids are as a result of the plasmid-encoded primase that negates the need for host-encoded functions (Meyer, 2009). Coupling of the RepB primase to the MobA relaxase, in both the IncQ-1 and IncQ-2 mobilization systems, in such a way that it is delivered to the recipient cell together with the transferred DNA facilitates establishment of the plasmid within the new host (Parker *et al*., 2002).

The mobilization systems of the two IncQ-2 plasmids are related to each other, although not as closely as that of the IncQ-1 plasmids (Francia *et al.*, 2004; Rohrer and Rawlings, 1992; Van Zyl *et al.*, 2003). The mobilization systems of the two IncQ subgroups, however, are not related. The three-component IncQ-1 mobilization system is unique whereas the five-component system of the IncQ-2 plasmids bears distinct similarity to the Tra1 proteins of the IncP plasmid RP4 (Rohrer and Rawlings, 1992; Van Zyl *et al.*, 2003). These two systems have therefore probably been acquired separately in order to adapt to and enhance the frequency at which the IncQ-like plasmids are mobilized by the conjugative plasmids in their environment.

Becker and Meyer (2003) proposed a model for the evolution of individual mobilization systems wherein it was suggested that the accessory proteins (such as MobB, MobC, MobD and MobE) became secondarily associated with the MobA and its *oriT* as they only serve to enhance stability of the relaxosome (Becker and Meyer, 2003). In support of this, Van Zyl and coworkers (2003) demonstrated the ability of mobilization systems to adapt to different conjugative systems with the following experiments. pTC-F14 was found to be mobilized at a 3000-fold lower frequency than pTF-FC2 by RP4, however, when the two plasmids were co-resident the mobilization frequency of pTC-F14 increased to nearly that of pTF-FC2. By cloning the *mobD* and *mobE* genes of pTF-FC2 and placing them *in trans* of pTC-F14 deletion derivatives, they showed that the MobD and MobE proteins of pTF-FC2 could be substituted for the pTC-F14 MobD and MobE proteins (the sequences of which are very different) with a consequent increase in mobilization frequency. From this they concluded that the pTF-FC2

mobilization system is better adapted to the Dtr system of RP4, and that the mobilization system of pTC-F14 is possibly better adapted to the Dtr system of a different unidentified conjugative plasmid. The results also suggested that swapping of the accessory genes such as is a likely mechanism by which mobilization systems can adapt to different conjugative systems (Van Zyl *et al.*, 2003). In conclusion, the ability of IncQ-like plasmids to adapt to and exploit a variety of conjugative systems for their own mobilization is in part responsible their ability to spread to such diverse hosts and habitats.

1.4 AIMS OF THIS PROJECT

The two pRAS3 plasmids (pRAS3.1 and pRAS3.2) have been isolated from several *A. salmonicida* species found in salmon aquaculture farms in Norway, Scotland and Japan (Aoki and Takahashi, 1986; L'Abee-Lund and Sørum, 2002). Due to their widespread nature they have been referred to as global, non-conjugative tetracycline resistance-bearing plasmids of *A. salmonicida*. Initial analysis of the sequence of these two plasmids by L'Abée-Lund and Sørum (2002) revealed that the two plasmids are near to identical to each other and that they contain replication and mobilization genes similar in sequence and organization to the IncQ-2 plasmids (Fig. 1.14). The authors, however, failed to identify a *repC* and *mobB* gene, even though it has been shown for other IncQ-like plasmids that the RepC protein is essential for replication and that the MobB protein is actively involved in mobilization.

Close inspection of the sequence differences between pRAS3.1 and pRAS3.2 (L'Abee-Lund and Sørum, 2002) revealed that pRAS3.1, unlike the other IncQ-like plasmids, including its isogenic counterpart pRAS3.2, contains 4 fully conserved and correctly spaced 22-bp iterons as opposed to only 3. This is significant as the iterons are essential for initiation of replication and copy number control (Kim and Meyer, 1985; Lin and Meyer, 1986; Lin *et al*., 1987; Rawlings and Tietze, 2001). pRAS3.1 also contains 5 CCCCCG-repeats located near the *oriT* whereas pRAS3.2 contains only 4 of these repeats. The *oriT* region has been identified as an important regulatory region for both IncQ-1 and IncQ-2-like plasmids (see section 1.2.4, Fig. 1.6). Finally, there are a number of point mutations within the *tetA*(C) and *tetR*(C) tetracycline resistance genes of these two plasmids which have resulted in five and ten amino acid substitutions within the TetA and TetR proteins, respectively. Besides these differences, pRAS3.1 and pRAS3.2 are otherwise identical. Apart from being isolated and

sequenced very few biological studies have, however, been performed on these two plasmids.

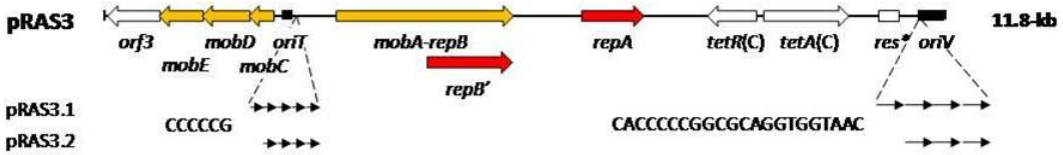


FIG. 1.14. ORFs and sequence differences between pRAS3.1 and pRAS3.2 identified by L'Abée-Lund and Sørum (2002). The number of 6-bp repeats located near the oriTs and 22-bp iterons within the *oriV*s of pRAS3.1 and pRAS3.2 are indicated by short and long arrows, respectively. The sequence of each of the two types of DRs is indicated on the left of the respective repeats. The nucleotide differences within the *tetAR* genes are not indicated in this diagram. The predicted genes involved in mobilization and replication are indicated in yellow and red, respectively, while the accessory genes are indicated in white.

The aims of this project are to; (1) characterize the pRAS3 plasmids with regard to sequence analysis, gene expression, copy number and mobilization; (2) determine their relationship to other IncQ-like plasmids with regards to incompatibility and mobilization; (3) determine the extent to which differences in the number of iterons and possibly gene expression contribute to the copy numbers of the two plasmids; and finally, (4) determine the influence of iteron copy number and PCN on intra- and intercellular competition. In doing so we hope to establish why two versions of the pRAS3 plasmid exist and gain insight into why IncQ-like plasmids tend to have only three perfectly conserved iterons and hence, a specific PCN.

Chapter 2

Characterization of pRAS3.1 and pRAS3.2

Table of Contents

2.1 INTRODUCTION

Plasmids are categorized based on the sequence similarity, genetic organization and the phenotypic properties of their replicons (Fernández-López *et al.,* 2006). To date, at least 28 plasmids have been identified as being IncQ-like (see Chapter 1 section 1.2.1), however, few have been studied in detail. The general characteristics shared by plasmids belonging to the IncQ family include their small size (~5 – 14-kb), broad-host-range, moderate copy number of 10 to 16 plasmids per chromosome and an ability to be efficiently mobilized by a variety of conjugative systems, especially by the transfer system of IncP plasmids.

The family consists of two major subgroups, namely IncQ-1 and IncQ-2. They are grouped into the two respective groups based upon the sequence similarity of their respective *repBAC* genes and *oriV*, their incompatibility relationship with other IncQ-like plasmids and on the basis of the type of mobilization system carried by the plasmids. The IncQ-1-like plasmids (RSF1010/R1162/R300B, pIE1107, pIE1150, pIE1120, pIE1130 and pDN1) have a mobilization system consisting of three genes (*mobABC*) and an *oriT,* while plasmids belonging to the IncQ-2-like subgroup (pTF-FC2 and pTC-F14) have a mobilization system consisting of five genes (*mobABCDE*) and an *oriT*. In contrast to the IncQ-1-like plasmids, members of the IncQ-2 subgroup have a toxin-antitoxin (TA) system located between the *repB* and *repA* replication genes. Although streptomycin and sulfonamide resistance genes are common examples of antibiotic resistance genes carried by IncQ-like plasmids, the accessory DNA can be diverse or cryptic by nature or even absent (refer to Chapter 1 section 1.2.1).

Plasmids pRAS3.1 and pRAS3.2 were isolated from five strains of *Aeromonas salmonicida* subsp. *salmonicida* and four atypical *A. salmonicida* strains*,* respectively, during a screen for plasmids carrying tetracycline resistance genes (L'Abee-Lund and Sørum, 2002). Eight of the nine strains were isolated from Norwegian salmon farms and one of the *A. salmonicida* subsp. *salmonicida* strains was isolated from a salmon aquaculture farm in Scotland. *A*. *salmonicida* subsp. *salmonicida,* and to a lesser extent the atypical *A. salmonicida*, are pathogens causing furunculosis in salmon. Strains of both *A. salmonicida* species carried a 45-kb plasmid, pRAS1, in addition to pRAS3.1 or pRAS3.2. One of the *A. salmonicida* subsp. *salmonicida* strains carried a 48-kb plasmid, pRAS2, as co-resident plasmid with pRAS3.1. Both pRAS1 and pRAS2 are capable of self-transmission between donor and recipient cells, while the two pRAS3

plasmids are not. It was observed, however, that the two pRAS3 plasmids were efficiently mobilized by either of the two conjugative plasmids (L'Abee-Lund and Sørum, 2002). Analysis of the nucleotide sequence of the two pRAS3 plasmids revealed that they are almost identical to each other. Plasmid pRAS3.1 (11,851-bp) is 28-bp larger than pRAS3.2 (11,823-bp) as a result of having 4 × 22-bp conserved iterons within the *oriV* and 5 GCGGGG (6-bp) repeats between the *oriT* and *mobA*, compared to 3 × 22-bp iterons and 4 × 6-bp repeats for pRAS3.2. There are also a number of point mutations within the *tetAR*(C) tetracycline resistance genes of the two plasmids that have resulted in 5 amino acid substitutions in the TetR transcriptional regulator and 10 amino acid substitutions within the TetA tetracycline efflux pump.

Sequence comparisons by L'Abee-Lund and Sørum (2002) of the pRAS3 sequence to that of other known plasmids within the NCBI database revealed that pRAS3 contains replication and mobilization genes that are between 89.5 and 96.2% identical to that of pTF-FC2, that the genes are organized in a similar manner and that pRAS3 contains *oriV* and *oriT* sequences that are IncQ-like. What was surprising in the manuscript by L'Abee-Lund and Sørum, however, was the apparent absence of a *repC* and *mobB* gene in pRAS3. The RepC protein is a plasmid-specific protein that binds to its cognate iterons in order to initiate plasmid replication (Kim and Meyer, 1985). The RepA and RepB proteins alone are unable to initiate replication and RepC is thus essential for plasmid maintenance (Scherzinger *et al*., 1991). The MobB protein forms part of the relaxosome which is assembled at the *oriT* during mobilization (Perwez and Meyer, 1999). Although not known whether the function of the pTF-FC2 MobB is similar to that of RSF1010, inactivation of the *mobB* gene of pTF-FC2 by a frame-shift mutation resulted in a 4,500-fold decrease in the mobilization frequency (Rohrer and Rawlings, 1992). Therefore, although the MobB is not absolutely essential, it is required for efficient mobilization.

It can be said, in evolutionary terms, that all plasmids that exist are successful. Only mutations that increase the ability of the plasmid to replicate, spread and outcompete its parent plasmid will become established (Paulsson, 2002; Watve *et al.,* 2010). Plasmids pRAS3.1 and pRAS3.2 are unique in that they represent two natural (successful) variants of the same plasmid, but which differ from each other in two putative regulatory regions, and have managed to spread and persist in the environment. Apart from identifying the two pRAS3 plasmids as tetracycline resistance determinants of *A*. *salmonicida* and observing that they are readily mobilized by their respective co-resident conjugative plasmids, no biological studies were performed on

pRAS3.1 and pRAS3.2 and their sequences have been only partially annotated. The aim of this chapter was to complete the annotation of the two pRAS3 plasmids, determine their copy numbers and mobilization frequencies, and finally to group the two plasmids into the IncQ plasmid classification scheme (Gardner *et al*., 2001) based on sequence similarity, functional relatedness of their mobilization systems and replicon incompatibility. The comparative analysis of the two pRAS3 plasmids in this chapter lays the foundation for the subsequent chapters wherein the reasons for the differences between the two plasmids and the significance thereof, is addressed.

2.2 MATERIALS AND METHODS

2.2.1 Bacterial Strains, Media and Growth Conditions. *E. coli* strains, cloning vectors and plasmid constructs are given in Appendix B. Detailed plasmid maps of pRAS3.1, pRAS3.2 and pBAD28 are given in Appendix D. In all aspects where pRAS3.1 and pRAS3.2 are identical, the two plasmids will be referred to as pRAS3. All cultures were grown in Luria-Bertani broth (LB) or on Luria-Bertani agar (LA) at 37°C unless otherwise stated (Appendix A). Antibiotics were added as required and are listed in Appendix A. Cultures were washed or diluted in phosphate-buffered saline (PBS, pH7.4, Appendix A).

2.2.2 DNA Techniques, Sequencing and Analysis. Plasmid preparation, restriction endonuclease digestions, gel electrophoresis and cloning were carried out using standard techniques (Ausubel *et al.,* 1993; Sambrook *et al.,* 1989). The polymerase chain reaction (PCR) was used to amplify DNA regions for cloning when no appropriate restriction sites were present. The reaction parameters included an initial denaturation step of 90 s at 94°C, followed by 30 cycles of denaturation (30 s at 94°C), a variable annealing step dependent upon the average primer annealing temperature (Appendix C), and an elongation step at 72°C with the extension time dependent on the amplicon size (Appendix E). PCR was performed in a Sprint temperature cycling system (Hybaid) using the Expand high-fidelity PCR system DNA polymerase (Roche Molecular Biochemicals). The sequences of all PCR products intended for cloning were confirmed by DNA sequencing using the dideoxy chain termination method and an ABI PRISM™ 3100 genetic analyzer.

2.2.3 Plasmid Copy Number Determinations. Plasmid-containing *E. coli* DH5α cultures were grown overnight in the presence of antibiotics, reinoculated at a 1/100 dilution into 50 ml pre-warmed non-selective LB media and grown while shaking at 37 ° to an OD_{600} of ± 0.8. Total genomic DNA (gDNA) was extracted from 1 ml culture using the QIAamp® DNA Mini Kit (QIAGEN). The gDNA was eluted in a final volume of 60 µl by running two volumes of 30 µl elution buffer over the column. The DNA concentration and purity was checked using a NanoDrop (Thermo Scientific) spectrophotometer.

Real-time qPCR amplification was performed using a LightCycler™ (version 2.0) with the LightCycler FastStart DNA Master SYBR Green I kit (Roche Diagnostics) according to the protocol of Lee and coworkers (Lee *et al.,* 2006a). A total of 4 ng gDNA was added to each 20 µl amplification reaction containing 3 mM $MgCl_2$, 0.5 µM of each of the forward and reverse primers and 2 µl of the FastStart DNA Master SYBR Green I reaction mix. The thermal cycling parameters included a pre-incubation at 95°C for 10 min, followed by 35 cycles of denaturation (95°C, 10 s), annealing (56°C, 4 s) and extension (72°C, 11 s). The ramp rate was set to 20°C/s. The melting curve was measured by cycling from an annealing temperature of 65°C through to a melting temperature of 95°C at a ramp rate of 0.1°C/s.

The plasmid copy numbers of pRAS3.1 and pRAS3.2 were determined as the number of plasmids per chromosome according to the method of Lee *et al*. (2006a). The respective plasmid and chromosomal DNA concentrations in each 4 ng total DNA sample were extrapolated from standard curves included in the same run using the LightCycler software (version 3.5). The standard curves (C_T value *vs* Log of the DNA concentration) were created using a 10-fold dilution series (4×10^0 to 4×10^{-5} ng) of plasmid DNA. The *gapA* amplicon (Appendix C) was cloned into pGEM-T®easy to serve as a template in the standard curve. The pGEM-GAPA and pRAS3.1 plasmids were extracted from *E. coli* DH5α using a Nucleobond® AX plasmid DNA purification kit (Machery-Nagel), and their concentrations were determined (6 replicates) using a NanoDrop™ spectrophotometer (Thermo Scientific). The number of molecules in each reaction was calculated using the formula given in Appendix A. The amplification efficiencies for each primer set was determined from the same standard curves.

2.2.4 pOU82 Stability Assays. Plasmid stability assays were performed using the pOU82-test vector (Gerdes *et al.*, 1985) according to a method adapted from Cooper and Heinemann

(2000). Single *E. coli* DH5α colonies containing the respective plasmids were inoculated into 5 ml LB media containing antibiotics and grown at 30°C for 12 h, whereafter ~1,000 cells, obtained by serial dilutions in PBS, were transferred at 24 h intervals to fresh LB media for a total of 4 days in the absence of selection. Each day the serial dilutions containing the cultures were also spread onto LA plates containing 40 μg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal) and incubated at 37°C. The number of plasmid-containing (blue) and plasmid-free (white) colonies was recorded for each ~20-generation sample point and used to calculate the percentage plasmid loss (Gerdes *et al.*, 1985).

2.2.5 Determination of Mobilization Frequency. *E. coli* S17.1 donor cultures containing the respective plasmids and *E. coli* CSH56 recipient cultures were inoculated into 5 ml LB media containing the appropriate antibiotics and incubated overnight at 37°C. Two milliliters of the cells were harvested by centrifugation at 8,000 rpm for 2 min and washed by three successive rounds of resuspension in 2 ml PBS followed by centrifugation before finally being resuspended in 1 ml PBS. The absorbance of each sample was determined at OD_{600} and standardized to an absorbance value of 1. The cultures were then mixed in a donor-to-recipient ratio of 1:100 and 100 μl of each mixture was spotted onto a LA plate and incubated at 37°C for 30 minutes. The agar plugs were excised, resuspended in 10 ml PBS and vigorously shaken for 30 min after which cells from 8 ml of each sample were collected by centrifugation and resuspended in 1 ml PBS. Serial dilutions were spread onto donor- and transconjugant-selective LA plates and the number of transconjugants per donor calculated after incubation at 37°C. Donor and recipient cultures that were not mated were also spread onto the respective antibiotic-containing plates as a control.

2.2.6 Incompatibility Assays. Plasmid-containing *E. coli* DH5α cultures were made chemically competent and transformed with a second plasmid and spread onto LA plates containing antibiotic selection for both plasmids. Single colonies were picked into 5 ml LB media containing the respective antibiotics for both plasmids and incubated overnight at 30°C. Thereafter, for two consecutive days, the cultures were serially diluted in PBS and ~1.0×10^3 cells from each culture were re-inoculated into fresh non-selective LB media and grown for 24 h (~34 generations in total). The PBS-diluted cultures from the selective growth cycle, as well as a PBS-diluted culture from the non-selective growth cycles were spread onto non-selective LA plates and incubated at 37°C. Finally, 50 colonies of each culture were replica-plated from

the non-selective plates onto sets of LA plates containing selection for either plasmid individually or both plasmids together, as well as plates without antibiotic selection. The percentage plasmid retention for each sample was determined by scoring the number of colonies that grew on the respective selective LA plates compared to the non-selective plates. As a control, cells containing the individual plasmids were grown and replica-plated similarly in order to monitor plasmid stability.

2.2.7 Displacement Assays. When the incompatibility was so strong that it was not possible to obtain co-transformed cells, reciprocal transformation experiments were used to determine the ability of an incoming plasmid to displace a resident (established) plasmid in order to verify the incompatibility as symmetrical or unidirectional (Gardner *et al*., 2001). Competent *E. coli* DH5α cells containing a resident plasmid were transformed with a second plasmid and spread onto LA plates containing antibiotic selection for the incoming plasmid only. After overnight incubation at 37°C, sixteen of the colonies were picked and replica-plated onto differential LA plates to select for the resident plasmid, the incoming plasmid, both incoming and resident, as well as non-selective plates in order to account for cell viability. Each transformation experiment also included a parallel transformation with a compatible vector such as pUC19 or pACYC177 as a control for competency. Cells containing each plasmid as a single resident plasmid were also grown on non-selective LA plates for the duration of the assay before being tested for plasmid retention on selective plates as a control for plasmid stability.

2.2.8 Random Knockouts and Screening for an Incompatibility Determinant. The EZ-Tn5™ transposon system (Epicenter®) was used to generate random knockouts of pRAS3.1. The transposon (0.025 pmol) and 1 μl of the transposase were mixed with 0.05 pmol pRAS3.1 DNA in the EZ-Tn5™ reaction buffer and incubated for 2 h at 37°C as per the manufacturer's protocol. The reaction was stopped by heat inactivation at 70°C for 10 min and the salts were removed by dialysis of the reaction mix on a 0.025 μm membrane filter (MILLIPORE) for 30 min. One microliter of the reaction mix was transformed into 50 μl electrocompetent *E. coli* EC100D cells using a Gene Pulser® II electroporation apparatus (Bio-RAD laboratories). The electrical settings were: voltage, 2.5 kV/cm; capacitance, 25 μF; and pulse controller parallel resistor, 200 Ω. Immediately after discharge 950 μl SOC recovery media (Appendix A) was added to the transformed cells, which were then incubated for 60 min at 37°C. The entire transformation mixture was spread onto LA plates containing kanamycin antibiotic selection

and incubated at 37°C overnight. All the colonies were scraped off the plates into 100 ml fresh pre-warmed 100 ml LB media containing antibiotic selection and incubated for 1 h at 37°C. The plasmid DNA was purified using the Nucleobond AX™ plasmid purification kit (Macherey-Nagel®). Twenty four of the colonies that were originally obtained were restreaked separately to obtain clonal cells from which the plasmid DNA was purified so that the frequency of EZ-Tn5 insertion into pRAS3.1 could be determined. The EZ-Tn5 vector (10 pg) as well as a no DNA control was transformed in parallel as controls for cell competency.

The bank of random knockouts was screened for a compatible phenotype by transforming 1 ng of the purified plasmid DNA into electrocompetent *E. coli* EC100D cells containing a resident pTF-FC2Cm plasmid. After recovery, various dilutions of the transformation mixture were spread onto LA plates containing selection for either the incoming mutated pRAS3.1 only or both the incoming and resident pTF-FC2Cm plasmids together and incubated overnight at 37°C. Two hundred of the colonies that appeared to have a compatible phenotype on the double-selective plates were replica-plated onto LA plates containing selection for the incoming pRAS3.1 plasmid only. After two rounds of growth while selecting for pRAS3.1 only, the colonies were replica-plated onto differential selective plates in order to verify retention of the pTF-FC2 plasmid. The plasmid DNA was extracted from 48 of the colonies that displayed a compatible phenotype and analyzed and grouped by restriction endonuclease analysis. One representative from each group was re-transformed into chemically competent *E. coli* DH5α cells and spread onto LA plates with selection for only the pRAS3.1 mutant in order to separate the plasmids. Separation of the pRAS3.1 mutants from pTF-FC2Cm was confirmed by replica-plating the cultures onto differential selective antibiotic-containing media. The plasmid DNA was then purified and the exact positions of insertion verified by sequencing using the R6KKanR fwd and rev primers (Appendix C). The EZ-Tn5 vector was cloned into the SphI sites within the tetracycline resistance genes of pRAS3.1 to serve as a positive control for incompatibility. The resulting construct (pRAS3.2::tetAR) as well as R6K (EZ-Tn5, negative control for incompatibility) was transformed in parallel with the bank of mutant plasmids.

2.2.9 Reverse Transcription-PCR (RT-PCR).

Plasmid-containing *E. coli* DH5α cells from an overnight culture were diluted 1×10^2-fold into fresh pre-warmed LB media and incubated at 37°C while shaking until the cultures reached an

optical density at 600 nm of 0.7 - 0.8. The cells were harvested and total RNA was isolated using a RiboPure™-Bacteria RNA isolation kit (Ambion®), as per manufacturer's instructions, and eluted in 60 μl elution buffer. The concentration of the RNA was determined by measuring its absorbance at 230 nm on a NanoDrop™ spectrophotometer (Thermo Scientific). The RNA was treated with DNAse I deoxyribonuclease (Roche) as per manufacturer's protocol to remove any remaining gDNA. The quality of the RNA was assessed by electrophoresis at 60 V for 60 min on a 1% (w/v) agarose/18% (v/v) formaldehyde denaturing gel in 1 × MOPS electrophoresis buffer (Appendix A). Approximately 1 μg of the purified RNA was converted to cDNA using the Transcriptor First Strand cDNA synthesis kit (Roche) as per manufacturer's protocol. The cDNA was diluted to a concentration of ~50 ng/μl and 100 ng of cDNA was used in each PCR reaction. The PCR reaction parameters were as described in section 2.2.2. An RNA control was included in the PCR reactions to verify that there was no residual DNA prior to cDNA synthesis.

2.3 RESULTS

2.3.1 Reanalysis of pRAS3.1 and pRAS3.2.

The sequences of the two pRAS3 plasmids were previously reported and annotated by L'Abee-Lund and Sørum (2002). Sequence comparisons by these authors revealed that the two plasmids have replication and mobilization genes that are similar in sequence and organization to that of pTF-FC2. Careful reanalysis of the sequence revealed the presence of both the *repC* and *mobB* genes as well as three other genes (*orf3* and the *pemIK*-like genes) and one truncated resolvase-like gene (*res**) which were previously not detected. The size of the proteins encoded on the pRAS3 plasmids as well as their putative translational initiation sites are listed in Table 2.1. The location of the genes on pRAS3 and their similarity to pTF-FC2 and pTC-F14 are illustrated in Figure 2.1. The comparison shows that overall the pRAS3 RepA, MobA-RepB, MobA, MobB, MobC, MobD and MobE proteins are between 89 and 97% identical to the corresponding proteins in pTF-FC2. The RepC protein of pRAS3, however, is much less conserved as it is only 25% identical to its counterpart in pTF-FC2 and even less (16%) to that of pTC-F14.

TABLE 2.1. Putative ribosomal binding sites of the pRAS3 genes and biochemical information of the protein products

Protein	Putative Translational Initiation Sites[A]	pRAS3			pTF-FC2[D]		
		Amino acids	Molecular Mass (KDa)[C]	pI[C]	Amino Acids	Molecular Mass (KDa)	pI
Orf3	ACCTACGGAT**GGAGG**TTGCC*AUG*	250	28.3	5.98	-	-	-
MobE	GGTATT**GGAGG**CCGTAGTCC*AUG*	212	22.9	7.62	213	23.0	7.63
MobD	ACATTTA**AGGAG**TAAGCAAC*AUG*	227	25.4	5.45	227	25.3	5.44
MobC	TGAGCAA**AGGAG**TTAGCAGA*AUG*	118	13.0	9.55	118	13.0	9.55
MobB	GCCCCCGGCGAGA**AGGAG**AC*AUG*	104	11.6	9.05	106	11.6	9.39
MobA	CCAAGAGCCAGACGGCAGAG*AUG*	436	50.2	8.99	409	46.8	9.00
RepB	AGCAGGAAAGGGTGC**AGG**CG*AUG*	353	40.3	9.42	352	40.2	9.46
PemI	ACTTAATT**GGAGG**CTTGGAA*AUG*	74	8.5	5.23	-	-	-
PemK	GACGTGGACGGCTGGACGGA*AUG*	108	11.5	9.03	-	-	-
RepA	GTTCTACAC**AGGAG**AACAGC*AUG*	292	31.6	5.93	290	31.2	6.20
RepC	CAACAAGGGGGGCA**AGG**GCT*AUG*	309	34.4	7.78	299	33.8	8.87
TetR	TCAAGCATGAGAATTCGCGA*AUG*	226	24.5 (24.6)[B]	6.11 (6.44)	-	-	-
TetA	AACGCAGTC**AGG**CACCGTGT*AUG*	396	41.5	9.35 (9.32)	-	-	-
Res*	TATGACTGAGAGTCAACGCC*AUG*	99	-	-	-	-	-

[A] Putative ribosomal binding sites (matching three or more of the AGGAGG consensus RBS) are shown in bold and start codons in italics.

[B] Values in brackets represent the pRAS3.1 values that are different from the pRAS3.2 values.

[C] Molecular mass and isoelectric values were determined using the Compute pI/Mw tool on the ExPASy Proteomics server (Bjellqvist *et al.,* 1993; Bjellqvist *et al.,* 1994; Gasteiger *et al.,* 2005).

[D] Biochemical information of the pTF-FC2 proteins which are in common with the pRAS3 plasmids.

The two small ORFs located between the *repB* and *repA* genes are different to the small genes encoding the *pasABC* and *pasAB* toxin-antitoxin systems of pTF-FC2 and pTC-F14, respectively, however they appear to also encode a TA system that is distantly related to the *pemIK* (*parDE*) (Bravo *et al.,* 1987; Tsuchimoto *et al*., 1988) and *mazEF* (*chpAI* and *chpAK*) (Masuda *et al.,* 1993) TA systems. The *orf3* gene, located immediately after the *mobE* gene, encodes a putative 270-aa protein. Although similar in position, it does not bear any similarity to the *orf3* gene of pTF-FC2 which encodes a putative 170-aa protein.

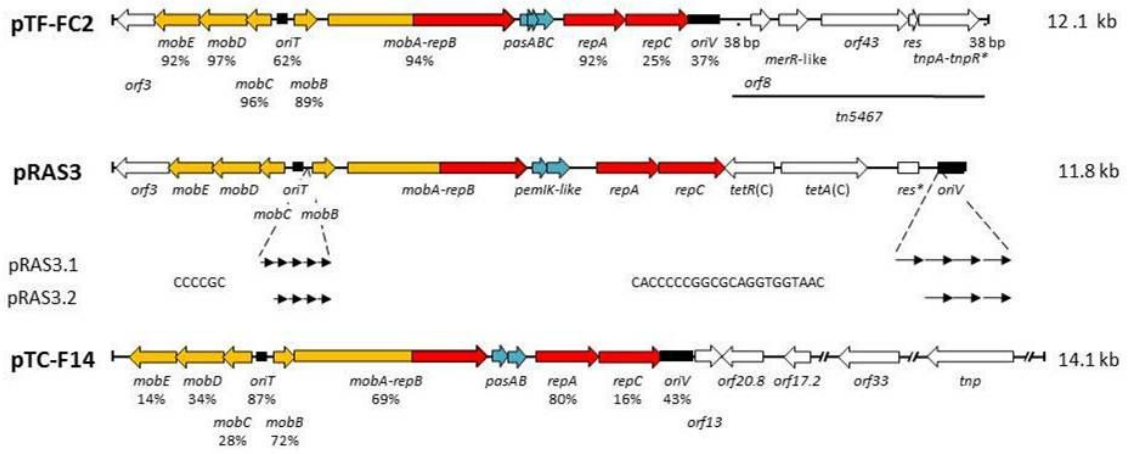


FIG. 2.1. Comparison of pRAS3.1 and pRAS3.2 to each other and to the IncQ-2 plasmids pTF-FC2 and pTC-F14. The differences in the number of 6-bp (left) and iteron (right) repeats and their respective repeat sequences are given below the pRAS3 diagram. The percentage values assigned to the pTF-FC2 and pTC-F14 mobilization (yellow) and replication (red) proteins indicate the percentage amino acid identity to the corresponding pRAS3 proteins. The percentage identity values below the *oriT* and *oriV* regions (black) of the two IncQ-2 plasmids refers to the nucleotide identity to the corresponding pRAS3 sequences. The stability system and accessory proteins are indicated in blue and white, respectively.

The partial resolvase-like gene located between the *oriV* and the *tetA* gene contains the coding sequence for 99-aa of a putative 155-aa resolvase-like protein encoded by plasmid R46. Interestingly, the disrupted end of the partial resolvase-like gene is the beginning of a 2620-bp region on pRAS3.1 (nt 3238 to 5872) and pRAS3.2 (nt 3199 to 5833) that is 98 and 99% identical to a corresponding region on R46 (nt 42063 to 44682) (Brown and Willetts, 1981), respectively (Fig. 2.2). The coding sequences of the *tetAR* genes present in these plasmids are included in this region, however, the region of homology ends abruptly 51-bp before the stop codon of the *tetR* gene of R46. This region was previously described by L'Abée-Lund and Sørum (2002) as a recombinatorial "hot spot". As a result of the recombinatorial activity in this location of the ancestral plasmid, the *tetR* genes of the pRAS3 plasmids encode an additional 7 amino acids and thus the TetR proteins of pRAS3.1 and pRAS3.2 share only 89 and 90% amino acid sequence identity with the TetR of R46, compared to the 96 and 99% identity for the TetA proteins, respectively.

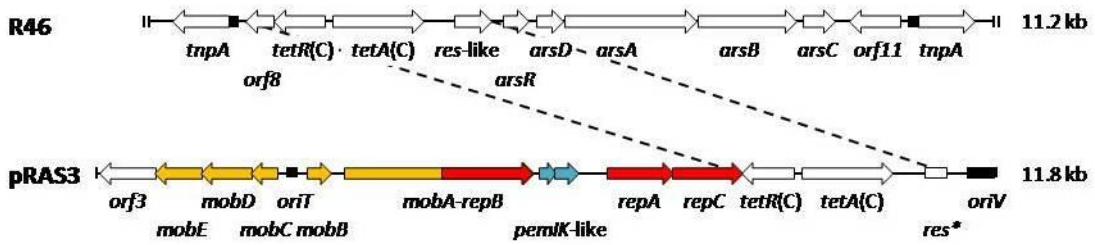


FIG. 2.2 The tetracycline resistance genes of the pRAS3 plasmids are located on a 2635-bp region that is ~99% identical at the nucleotide level to a corresponding 2620-bp region on R46. The region of similarity on the pRAS3 plasmids stretches from 74-bp before the carboxy-terminus of the *tetR* gene and ends 746-bp past the carboxy terminus of the *tetA* gene. A putative gene encoding a resolvase-like protein (*res*-like) is present in full on R46 but is truncated by 56 amino acids on the pRAS3 plasmids (*res**). A small insertion is present in the pRAS3 *tetA* gene 18-bp before the TGA stop codon and is responsible for the 15-bp size difference between the two genes.

2.3.2 The Architecture of the pRAS3 *oriV* Compared to Other IncQ *oriV*s

During their analysis of the *oriV* sequences of pRAS3.1 and pRAS3.2, L'Abée-Lund and Sørum (2002) were able to identify the iterons, a G + C- and A + T-rich region, as well as 3 short

(C)CCCCG repeats together with a AGACACA-tail immediately upstream of the first (C)CCCCG repeat. Similar repeats together with the tail region were also identified in the *mobA*, *mobC* and between the *oriT* and *mobB* sequences. No *ssiA* and *ssiB* sites were, however, reported. Reanalysis of the *oriV* sequence revealed two putative inverted repeat regions as indicated in Figure 2.3. The first repeat forms a 9-bp stem-loop structure with a ΔG value of -2.9 and the second a 6-bp stem-loop with a ΔG value of -2.0. Although the ΔG values are not very high, both repeats contain four conserved bases located at the bottom of the stem-loop structure on the 3'-side which, like the initiation sites of RSF1010 (Sakai and Komano, 1996), end with a TG sequence. These two inverted repeats are thus likely candidates for the *ssiA* and *ssiB* initiation sites. A 15-bp conserved region within the *oriV* regions of IncQ plasmids was previously identified by Rawlings and Tietze (2001) (Fig. 2.3). When the *oriV* sequences of the pRAS3 plasmids and pTF-FC2 are aligned such that the iterons, G + C- and A + T-rich regions overlap it appears as if the 15-bp conserved region is not present. However, upon close inspection of the sequence it is evident that it is present ± 12-bp further downstream and contains eleven of the normally fifteen conserved nucleotides. An alignment of the twenty conserved nucleotides of the iterons of the IncQ-1 representative plasmids RSF1010, pIE1107, pIE130 and IncQ-2 representative plasmids pTF-FC2 and pTC-F14 shows that eleven of the twenty conserved nucleotides are identical for each of the iterons (Fig. 2.4). The pRAS3 plasmids, however, share only five of the eleven highly conserved nucleotides. The architecture of the pRAS3 *oriV*s is thus comparable to that of other IncQ-like plasmids, however the iterons and the 15-bp highly conserved region of pRAS3 are not as conserved as for the other IncQ-like plasmids.

2.3.3 The *repC* Gene Was Possibly Acquired by Gene Swopping

The *repC* gene of pRAS3 encodes a protein with a predicted size of 309-aa, however, it bears only 25% identity to the 299-aa RepC protein of pTF-FC2 and even less (16%) to the 303-aa RepC protein of pTC-F14. The percentage of similar (but not identical) amino acids in the RepC amino acid sequence of pRAS3 is 42% compared to its counterparts in both pTF-FC2 and pTC-F14. Although the percentage similarity is much higher than the percentage identity, it is only half of the overall 82% similarity that is shared between the RepC proteins of the IncQ-1

(RSF1010, pIE1107 and pIE1130) and IncQ-2 plasmids (pTF-FC2 and pTC-F14). The RepA (292-aa) protein of pRAS3 on the other hand bears 92 and 80% identity to the RepA of pTF-FC2 (290-aa) and pTC-F14 (291-aa), respectively. Similarly, the RepB primase domain (352-aa) of the pRAS3 MobA-RepB fusion protein shares 97 and 79% identity with the RepB primase domains of pTF-FC2 and pTC-F14, respectively.

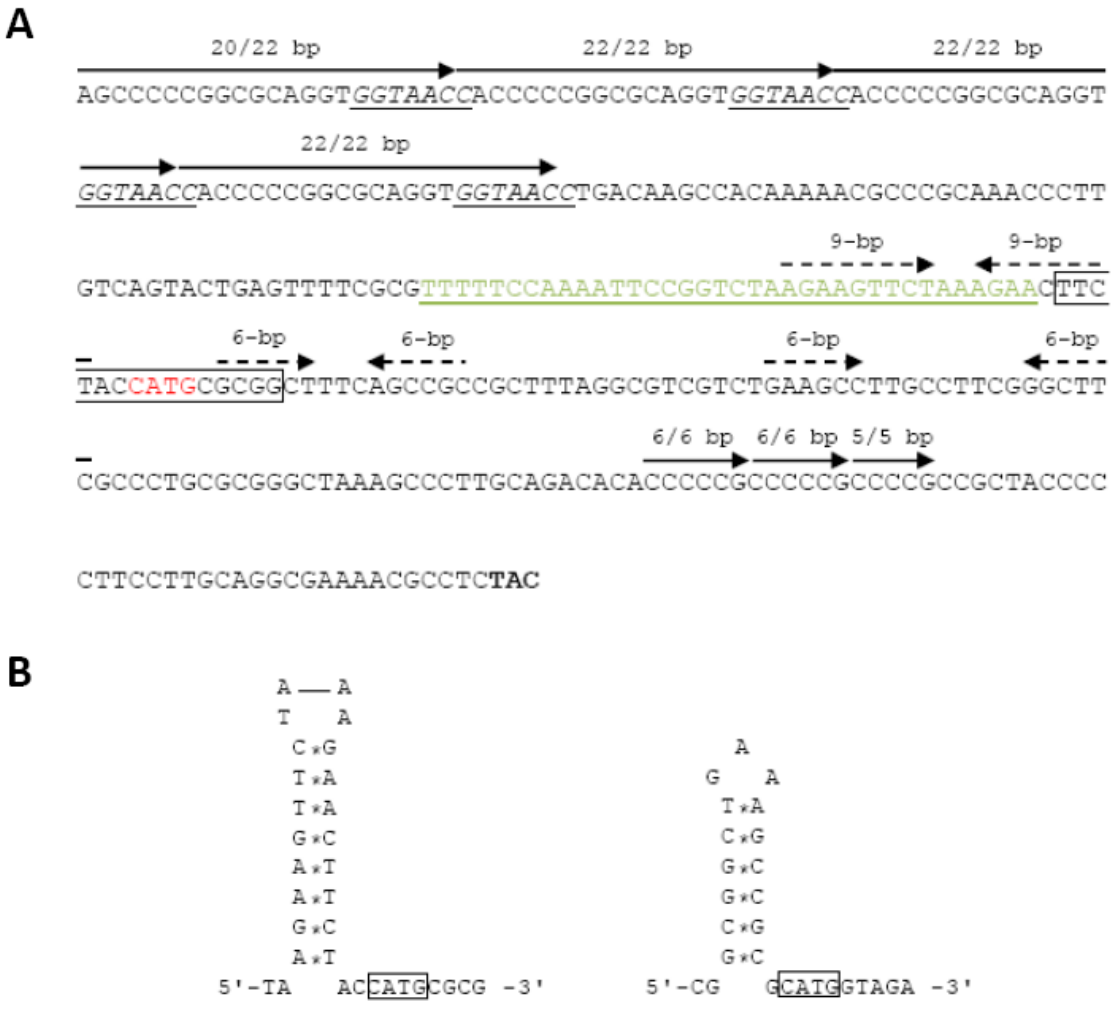


FIG. 2.3. The *oriV* of the pRAS3 plasmids contain all the features that are common to IncQ-like *oriV*s. (A) The features and nucleotide sequence of the *oriV* of pRAS3.1. The *oriV* of pRAS3.2 is similar, however, it contains only 3 × 22-bp iterons (long arrows spanning 22-bp each). The underlined sequences at the end of each of the iterons indicate the position of BstEII restriction sites which are unique to this region of the plasmid. The underlined sequence indicated in green is an A + T-rich region

and is followed by a 15-bp region (boxed) that is highly conserved between the *oriV*s IncQ-like plasmids. The inverted repeats (broken arrows) on the left and right of the conserved region possibly contain the single strand initiation (*ssi*) sites as they each have a conserved consensus sequence (red) on the 5'-side of the putative stem-loop structure. The CCCCCG repeats present in the *oriV* are indicated by short arrows spanning the 6-bp sequences. (B) Putative stem-loop structures containing the *ssiA*- (left) *ssiB*-like (right) sites on the 5'-side of the structure. The conserved sequences of the *ssi*-like sites are indicated by a box.

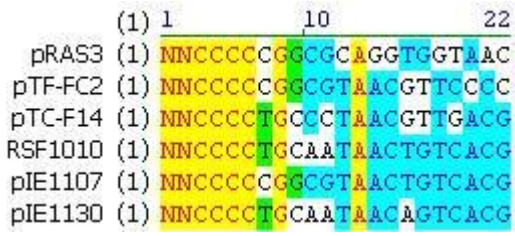


FIG. 2.4. A comparison of a single pRAS3 iteron to the iterons of representative plasmids from the IncQ-1 and IncQ-2 groups. The nucleotides of the 2-bp spacers are indicated by an N, which can be an A, T, G or C.

An ATP/GTP-binding site A signature (P-loop), bacterial histone-like DNA binding signature and DNA polymerase family A signature motif was identified within the RepC of pTC-F14 (Gardner, 2003) by comparing the protein sequence against the motif sequences within the PROSITE database (Hulo *et al.,* 2006) using ScanProsite (http://au.expasy.org/prosite) (De Castro *et al.,* 2006). A search for similar protein motifs within the pRAS3 RepC amino acid sequence revealed that none of these motifs are present in the sequence. A conserved pattern (A-L-N-P-x-[IL]-A-[DEQ]-A-[IV]-[IM]-G-[AEG]-x-[PQ]-H-[TV]-R-I-[ENS]-M-[AD]-E-V-R-x-L-[DEQ]-[ST]-[DE]-[APT]-A-R-L-[IL]-H-Q-R-L-C-G-W-I-D-P-G-K-[ST]-G) is present within the RepC sequences of RSF1010, pIE1130, pIE1107, pTF-FC2 and pTC-F14 and was identified by aligning the RepC sequences of these plasmids using the PRATT 2.1 tool (Jonassen *et al.,* 1995; Jonassen, 1997) available at ScanProsite. This conserved pattern could not be found in the RepC sequence of pRAS3. Therefore, the low level of similarity of the pRAS3 RepC protein to the RepC proteins of pTF-FC2 and pTC-F14 compared to the high degree of similarity of the RepA and RepB proteins between these plasmids, together with the difference in G+C content of the *repC*

gene compared to other pRAS3 genes, suggest that the *repC* gene was obtained during a recombinatorial event in which the previous initiator gene was exchanged for the current gene rather than the accumulation of point mutations.

BLAST analysis of the RepC amino acid sequence of pRAS3 against the non-redundant database at NCBI confirmed that the protein is only distantly related to the RepC proteins of other IncQ-like plasmids that have been characterized. The phylogenetic trees in Figure 2.5 were constructed using all the sequences identified by the BLAST program with an E value greater than 10^{-21}. It shows that unlike the RepA protein of pRAS3, which groups with other IncQ-like RepA proteins, the RepC protein does not cluster together with any of the RepC-like proteins from IncQ-like or other plasmids. The source from which the pRAS3 *repC* gene was possibly obtained, therefore, remains unknown.

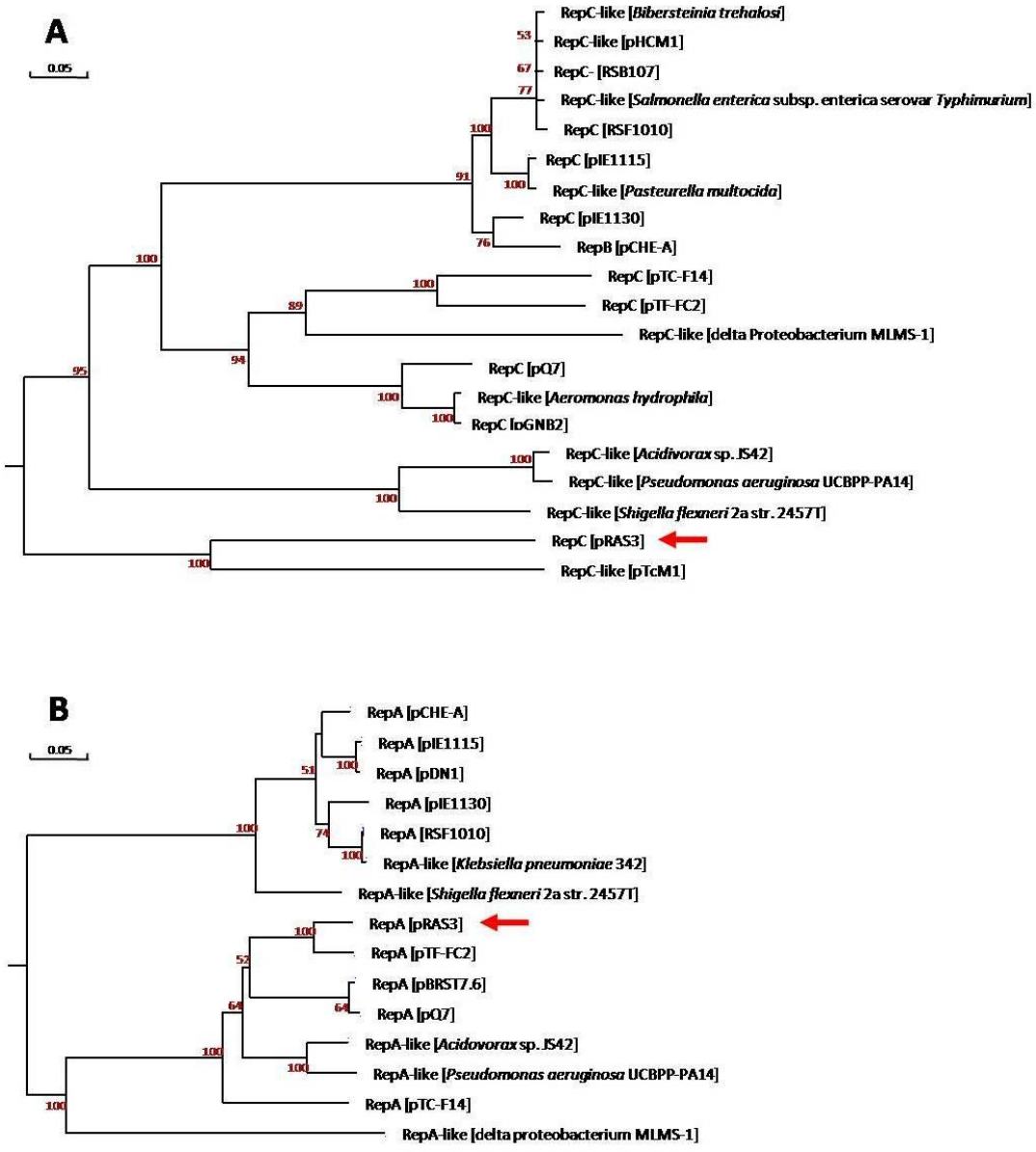


FIG. 2.5. Phylogenetic trees showing the relationship of the pRAS3 (A) RepC and (B) RepA protein sequences to other similar (E > 10^{-21}) protein sequences from the Genbank database. The pRAS3 RepC and RepA proteins are indicated by red arrows. The RepC-like sequences group into multiple clades and from the phyolgenetic tree it is evident that the RepC sequence of pRAS3 does not fall into the same clade as the RepC proteins from other IncQ-like plasmids. Two clades containing the IncQ-1 and IncQ-2-like RepA sequences are formed when these amino acid sequences are aligned. The RepA sequence of pRAS3 groups together with the IncQ-2-like RepA sequences. RepC sequences were as follows: *Acidovorax* sp. JS42 YP_986775; *Aeromonas hydrophila* YP_002221296; *Bibersteinia trehalosi*

YP_512236; delta Proteobacterium MLMS-1 ZP_01290807; *Pasteurella multocida* YP_232871; plasmid pCHE-1 YP_002563153; plasmid pGNB2 YP_003422524; plasmid pHCM1 NP_569413; plasmid pIE1115 NP_065286; plasmid pIE1130 NP_862666; plasmid pQ7 ACY39224; plasmid pRSB107 CAH64749; plasmid RSF1010 NP_044310; plasmid pTC-F14 11344137; plasmid pTF-FC2 AAA27382; plasmid pTcM1 YP_001837324; *Salmonella enterica* subsp. *enterica* serovar *Typhimurium* AAS18379 and *Shigella flexneri* 2a str. 2457T NP_838066. RepA sequences were as follows: *Acidovorax* sp. JS42 YP_986776; delta Proteobacterium MLMS-1 ZP_01290810; *Klebsiella pneumoniae* 342 YP_002237147; plasmid pBRST7.6 YP_002221297; plasmid pCHE-A YP_002563152; plasmid pDN1 NP_073216; plasmid pIE1115; plasmid pIE1130 NP_862665; NP_065285; plasmid RSF1010 NP_044309; plasmid pTC-F14 NP_835380; plasmid pTF-FC2 AAA27381; plasmid pQ7 ACY39225; *Pseudomonas aeruginosa* UCBPP-PA14 YP_789368 and *Shigella flexneri* 2a str. 2457T NP_838067

2.3.4 The RepC Is Essential for Replication

As the amino acid sequence of the pRAS3 RepC is very different from that of other IncQ-like plasmids I tested whether the gene and its corresponding protein product are still required for replication. The kanamycin resistant EZ-Tn5 vector was cloned into the SphI-SphI sites within the *tetAR* genes of pRAS3.1. This was followed by a NheI-NheI deletion which truncated the *repC* gene by 79 codons on the carboxy-terminus side. The resulting construct, pR6K.3.1.repC$^{\Delta}$, was transformed in parallel into *E. coli* EC100D and DH5α. It was able to become established, and thus able to replicate, in *E. coli* EC100D (which contains the *pir* gene required for maintenance of the R6K γ *ori*) but not in *E. coli* DH5α. Deletion of a large portion of the carboxy-terminus of RepC thus inactivated the pRAS3 replicon. Replication of the *repC*-deletion clone (pR6K.3.1.repC$^{\Delta}$) in *E. coli* DH5α was restored when a 1,015-bp fragment containing the *repC* gene was cloned by PCR and expressed *in trans* from behind the arabinose-inducible promoter P_{BAD} of pBAD28. The *repC* gene in pBAD28-repC was, however, only functionally expressed from behind P_{BAD} if it was provided with an artificial AGGAGG consensus RBS (pRAS3REPC Fwd primer, Appendix C) 6-bp upstream of the ATG-start codon. Alternatively, a 2.7-kb SalI-StuI fragment containing both the *repA* and *repC* genes cloned behind the arabinose-inducible P_{BAD} promoter of pBAD28 to give pBAD28-repAC, was also able to compliment pR6K.3.1.repC$^{\Delta}$. Translation of the *repC* gene therefore seems to be dependent on translation of the *repA* gene, but can be separated if a RBS is provided for the *repC* gene.

2.3.5 The Copy Number of pRAS3.1 and pRAS3.2 Differs

The copy numbers of pRAS3.1 and pRAS3.2 were measured using Real-Time quantitative PCR (qPCR) and found to be 45 ± 13 (*n* = 11) and 30 ± 5 (*n* = 4) plasmids per chromosomes, respectively, in *E. coli* DH5α. Primer sets specific to the *tetAR* genes of the pRAS3 plasmids (pRAS3A, Appendix C) and to the *gapA* single copy house-keeping gene of *E. coli* (*E. coli* GAPA, Appendix C) were used in the qPCR reactions to amplify the plasmid and chromosomal DNA, respectively. The amplification efficiencies of the pRAS3A and *E. coli* GAPA primer sets were found to be 2.01 ($R^2 > 0.999$) and 2.04 ($R^2 > 0.999$), respectively. Although the amplification efficiencies of the two primer sets were at the maximum theoretical amplification efficiency (*E* = 2.0), no non-specific amplification products, which could have resulted in an abnormally high amplification efficiency, were observed during the melting curve analysis and thus the standard curve was not discarded (Fig. 2.6). The number of plasmids and chromosomes in the total DNA samples were extrapolated from standard curves that were created by plotting the Ct values obtained for each of the pRAS3.1 and pGEM-GAPA standards against the log of their respective concentrations using the LightCycler 3.5 Software. The range of the serially diluted standards spanned from 7.78×10^8 to 7.78×10^3 and 2.68×10^9 to 2.68×10^4 molecules per reaction for pRAS3.1 and pGEM-GAPA, respectively.

2.3.6 The Toxin-Antitoxin System Is Functional

The two ORFs located between the *repB* and *repA* genes encode two small proteins of 74 and 108-aa, respectively. The putative -35 (TTGCAT) and -10 (TACAAT) sites of the promoter region 30-bp upstream of the first ORF are separated by 17-bp (Appendix E) and the promoter is thus very similar to the *E. coli* σ^{70} consensus promoter (TTGACA - ± 17-bp - TATAAT) (Lisser and Margalit, 1993). Each of the ORFs have a putative RBS located 10 and 9-bp upstream of their

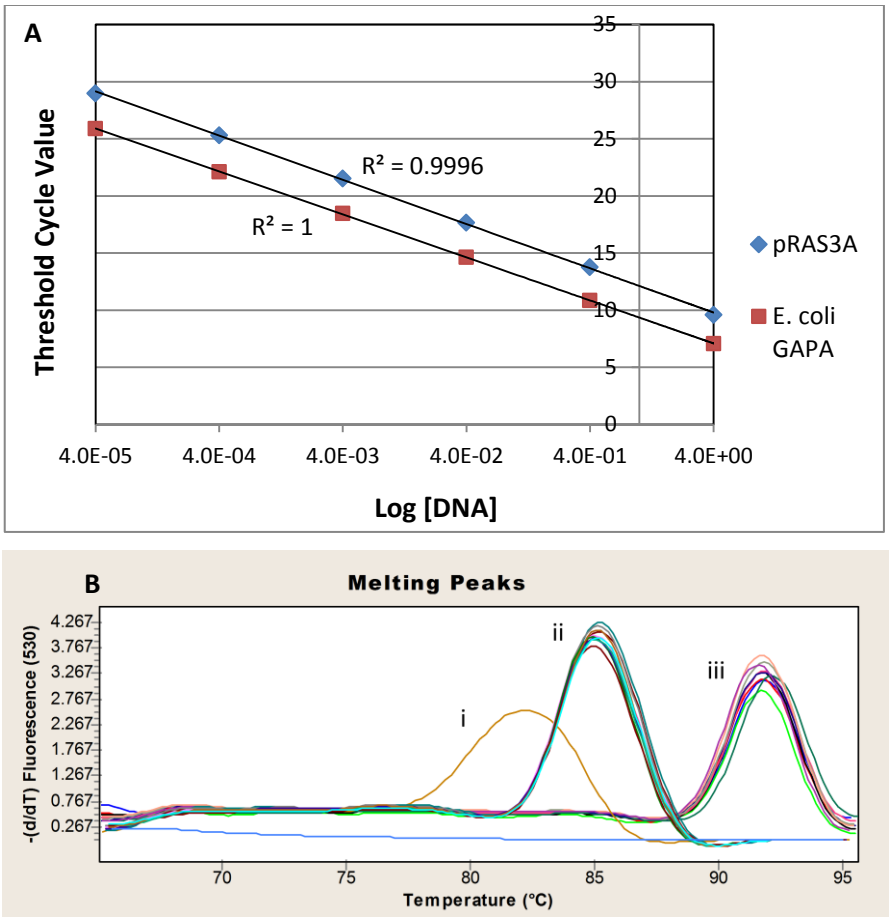


FIG. 2.6. Standard curves for the pRAS3A and *E. coli* GAPA primer sets. (A) The standard curves for both the pRAS3A and *E. coli* GAPA primer sets had R^2-values of 0.999 and greater, indicating a good correlation between the C_T value and DNA concentration. (B) Primer dimers were formed in the no DNA control of Real-Time PCR reactions containing the pRAS3A primers (i) when the number of amplification cycles exceeded 35 cycles. As the pRAS3A primer dimers were never present when 4.0×10^{-5} ng or more of pRAS3 template was included in the reaction (ii), the presence of the primer dimers in the no DNA control was not regarded as a cause of concern. No primer dimers or non-specific products were observed in the Real-Time PCR reactions when the *E. coli* GAPA primer set was used (iii).

ATG start codons and differ from the *E. coli* consensus RBS by only 1 and 2 bases, respectively. BLAST analysis of the two ORFs revealed that they are closely related to pairs of proteins that are transcribed from genes similar in size and organization, which have been identified during genome sequencing studies of bacteria such as *Xanthomonas campestris, Xanthomonas axonopodis, Aeromatoleum aromaticum* and *Nitrosomonas europaea* (Fig. 2.7). Although none of these proteins has been experimentally verified as an element of a TA system, they are distantly related to either the plasmid-encoded PemIK or chromosomal MazEF TA systems.

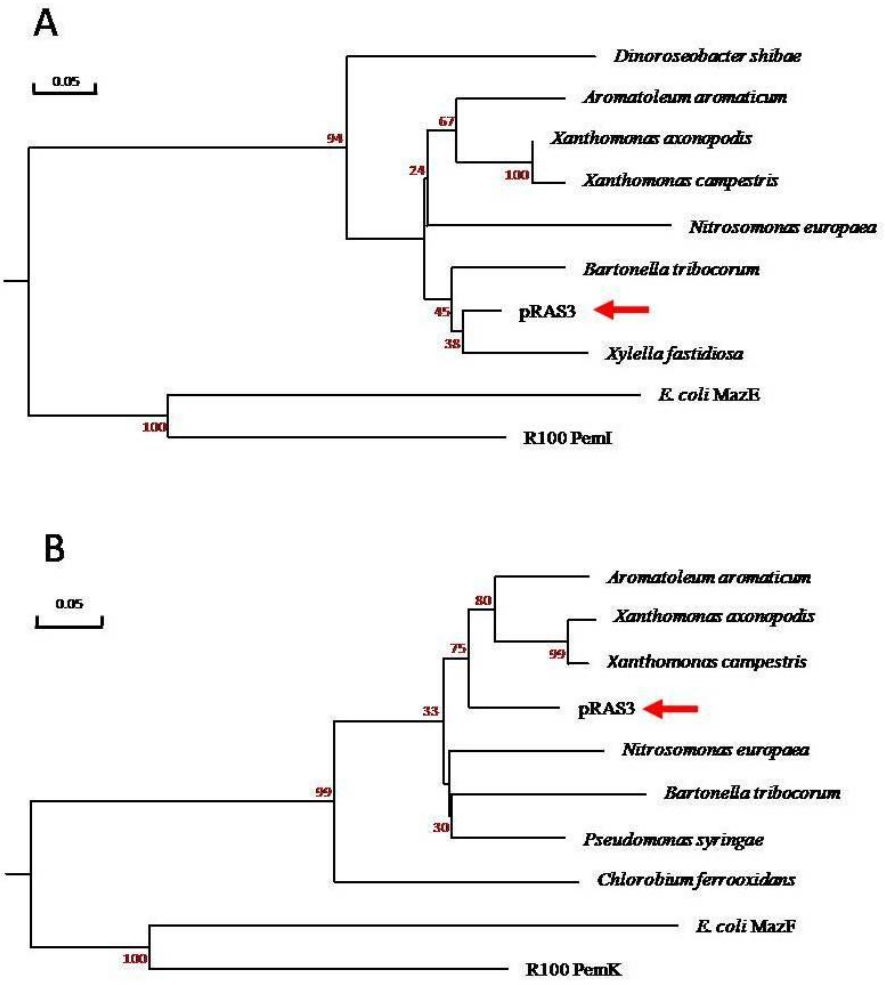


FIG. 2.7. Phylogeny of toxin-antitoxin proteins of pRAS3 and comparison with closely related proteins as well as the more distantly related PemIK and MazEF proteins. A. antitoxins; *Aromatoleum aromatium*, CAI08016; *Bartonella tribocorum*, CAK00897; *Dinoroseobacter shibae* YP_001541878; *Nitrosomonas europaea* ATCC19718, CAD85218; *Xanthomonas axonopolis pv. citri str*. 306, NP_644761; *Xanthomonas campestris pv. vesicatoria str.* 85-10, CAJ19793; *Xylella fastidiosa* Ann-1, ZP_00682677; *E. coli* MG1165 MazE, AAA69293; plasmid R100 PemI, P13975. B. toxins; *Aromatoleum aromatium*, CAI08015; *Bartonella tribocorum*, CAK00896; *Chlorobium ferrooxidans* EAT59633; *Nitrosomonas europaea* ATCC19718, CAD85217; *Pseudomonas syringae pv. phaseolicola*, AAZ37969; *Xanthomonas axonopolis pv. citri str*. 306, NP_644760; *Xanthomonas campestris pv. vesicatoria str.* 85-10, CAJ19792; *E. coli* MG1165 MazF, AAA69292; plasmid R100 PemK, P13976.

To verify whether the two ORFs on the pRAS3 plasmids are able to enhance plasmid stability, a 731-bp PCR fragment containing the two genes and their promoter region, was cloned onto the segregationally unstable pOU82 test vector. The pOU82 vector contains a β-galactosidase gene which allows for the differentiation of plasmid-containing (blue) and plasmid-free (white) colonies on LA containing X-gal. *E. coli* DH5α cultures containing either the pOU82-TA construct or the pOU82 vector as a control were grown in serial batch cultures in the absence of antibiotic selection and monitored at ~18-generation intervals for plasmid retention on differential LA plates. After ~72 generations of growth approximately 98% of the *E. coli* DH5α cells retained the pOU82-TA plasmid compared to 35% of *E. coli* DH5α cells that retained the pOU82 control plasmid (Fig. 2.8). The presence of the *pemIK*-like genes on an unstable heterologous replicon significantly enhanced the stability of the plasmid within the bacterial population, thus confirming that the *pemIK*-like genes could function as a plasmid stability system.

2.3.7 The Two pRAS3 Plasmids are Mobilized at Similar Frequencies

The two pRAS3 plasmids were isolated by L'Abée-Lund and Sørum (2002) by allowing their respective *A. salmonicida* hosts to mate with an *E. coli* recipient host. No mobilization frequency was, however, reported. The RP4 conjugative system was used to determine the mating frequencies of pRAS3.1 and pRAS3.2 in *E. coli*. *E. coli* S17.1, which carries the *Mpf* genes of the IncPα plasmid RP4 as well as a streptomycin resistance gene on its chromosome was used as a donor host, and the naladixic acid-resistant *E. coli* CSH56 as a recipient host. The donor and recipient hosts were mixed at a ratio of 1:100 donors per recipient and incubated for 30 min. The mating frequency was found to be similar for both plasmids at 0.032 ± 0.014 and 0.021 ± 0.013 transconjugants per donor, respectively. The mobilization frequencies were equal to or exceeded the donor to recipient ratio (reached saturation) when a smaller donor to recipient ratio (1:10) and or a longer incubation time (60 min) was used. The *E. coli* S17.1 (pRAS3.1 or pRAS3.2) donor was not able to grow on the recipient-selective LA plates and the *E. coli* CSH56 recipient was not able to grow on the donor-selective plates.

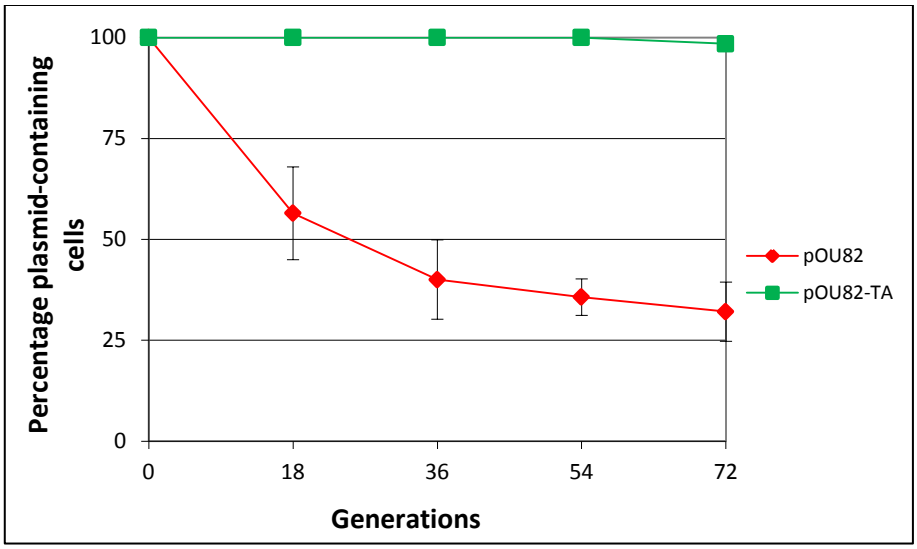


FIG. 2.8. Loss of the low-copy-number test plasmid, pOU82, with and without the PemIK-like genes from pRAS3 in the absence of antibiotic selection.

2.3.8 Functional Relatedness of IncQ-2 Mobilization Regions

The *oriT* of pRAS3 differs by only 5-bp over a 48-bp region from the *oriT* of pTC-F14 and by 17-bp from that of pTF-FC2 (Fig. 2.9). The 6-bp region comprising the putative *nic* sites of all three plasmids is, however, perfectly conserved. In contrast to the *oriT*, the mobilization genes of pRAS3 are more similar to those of pTF-FC2 than they are to the mobilization genes of pTC-F14. The percentage amino acid identity between the pRAS3 and pTF-FC2 *mob* genes varies from 89 to 97%, and from 14 to 72% between the *mob* genes of pRAS3 and pTC-F14 (Fig. 2.1). Interestingly, the percentage identity between the 410-aa of the MobA relaxase of pRAS3 and pTC-F14 is only 69%, however the two proteins share 79% similar amino acids. The difference between the number of identical (92%) and similar (94%) amino acids between the relaxase domains of pRAS3 and pTF-FC2 on the other hand is not as pronounced.

Although the *oriT*s of pTF-FC2 and pTC-F14 share only 60% nucleotide identity and the MobA proteins 71% amino acid identity, the pTF-FC2 mobilization system was shown to be able to mobilize the *oriT* of pTC-F14 (when cloned into a pUC-vector) at a frequency of 3.48×10^{-2} transconjugants per donor per hour using the RP4 conjugative system (Van Zyl *et al.*, 2003). The pTC-F14 mobilization system was, however, unable to mobilize the pTF-FC2 *oriT*.

Considering that the pRAS3 *oriT* is more similar to that of pTC-F14 while the mobilization proteins are more similar to the mobilization proteins of pTF-FC2, we wished to determine whether the pRAS3 mobilization system is able to mobilize the *oriT*s of pTF-FC2 or pTC-F14 and *vice versa*. A 196-bp region spanning the pRAS3 *oriT* was amplified by PCR and cloned into a non-mobilizable pUC19 vector to give pOriT-pRAS3. The cloned *oriT* (pOriT-pRAS3) was mobilized at saturating frequencies within 60 min when its own mobilization system was provided *in trans* by either pRAS3.1 or pRAS3.2 in *E. coli* S17.1, and was not mobilized at all when the mobilization components were absent. When either pTF-FC2Cm or pTC-F14Cm was present instead of one of the pRAS3 plasmids, no transconjugants were obtained. In the reverse experiment, however, pRAS3.1 was able to mobilize the *oriT*s of pTF-FC2 and pTC-F14, which were previously cloned into pUC19 (Van Zyl *et al.*, 2003), at a frequency of 1.20 (±0.44) × 10^{-1} and 3.42 (±1.64) × 10^{-4}, respectively. Unidirectional complementation of the pTF-FC2 and pTC-F14 *oriT*s by the mobilization proteins of pRAS3 is therefore possible. Oddly, although the *oriT* of pTC-F14 bears more similarity to the *oriT* of pRAS3 than does the *oriT* of pTF-FC2, it was mobilized at a lower frequency than the pTF-FC2 *oriT* by pRAS3, thus suggesting that the conservation of specific bases is more important than overall similarity.

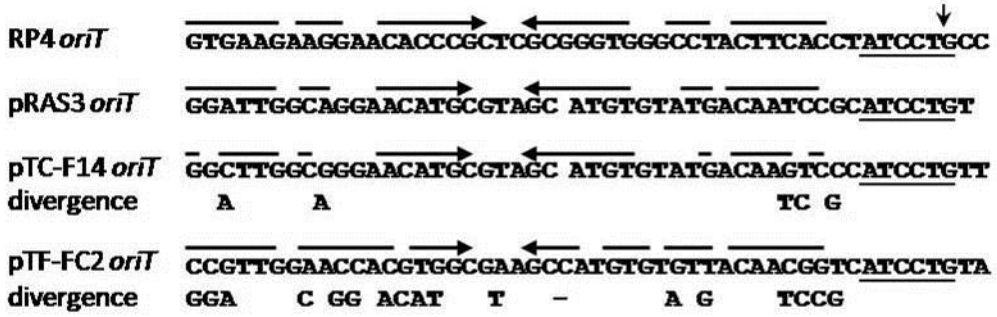


FIG. 2.9. Alignment of *oriT* regions of IncQ-2 plasmids and the IncPα plasmid RP4 showing the sequence divergence that could be tolerated by the Mob proteins of plasmid pRAS3 while still able to mobilize DNA from an *oriT*. A vertical arrow indicates the relaxase *nic* site at which single strand cleavage takes place as determined for plasmid RK2/RP4 (Pansegrau *et al.,* 1987).

2.3.9 Incompatibility of the pRAS3 Plasmids With Other IncQ-Like Plasmids

In order to group pRAS3 into the IncQ classification scheme (Gardner *et al*., 2001) we wished to determine the incompatibility relationship of pRAS3 with other IncQ-like plasmids. The IncQ-1 plasmids RSF1010 (Guerry *et al.,* 1974), pIE1108 (Smalla *et al*., 2000) and pIE1130 (Tietze *et al.,* 1989) are the prototype plasmids for the α, β and γ IncQ-1 incompatibility groups and were thus selected for incompatibility testing. *E. coli* DH5α cells containing either pRAS3.1 or pRAS3.2 coresident with RSF1010K, pIE1108Cm or pIE1130 were cultured in LB media overnight while selecting for both plasmids before the cultures were taken through two cycles of growth (~34 generations) in non-selective LB media. The IncQ-1 plasmids RSF1010K, pIE1108Cm and pIE1130 were slightly less stable on their own (94 to 98%) than was pRAS3.1 or pRAS3.2 (100% for both plasmids; Table 2.2). When co-resident with either pRAS3.1 or pRAS3.2, retention of the IncQ-1 plasmids was no different from their respective controls, while retention of either pRAS3 plasmid remained at 100%. No incompatibility was therefore evident between the two pRAS3 plasmids and the IncQ-1 group plasmids.

Attempts to co-transform pRAS3.1 or pRAS3.2 with either the IncQ-2α plasmid pTF-FC2 or IncQ-2β plasmid pTC-F14 were unsuccessful and no *E. coli* DH5α host cells containing both plasmids were isolated. This suggested that the pRAS3 plasmids and the two IncQ-2 plasmids pTF-FC2 and pTC-F14 are unable to coexist even when both plasmids are selected for by using antibiotics. We therefore determined which plasmid was responsible for the antagonistic phenotype by means of reciprocal transformation experiments wherein antibiotic selection was initially provided for only the incoming plasmid. Chemically competent *E. coli* DH5α cells containing a resident pTF-FC2Cm or pTC-F14Cm plasmid were transformed with either pRAS3.1 or pRAS3.2 as an incoming plasmid and spread onto LA plates containing selection for the incoming plasmid only. Sixteen colonies from the incoming-selective plates were picked and replica-plated onto differential LA plates containing selection for the resident plasmid, the incoming plasmid or both plasmids together, as well as onto non-selective LA plates to account for cell viability. None of the tested colonies contained the resident pTF-FC2 or pTC-F14

TABLE 2.2. Plasmid segregation patterns observed during incompatibility assays

Plasmid Combinations	Plasmid Segregation Pattern (Percentage Retention)				
	pRAS3.1	pRAS3.2	RSF1010K	pIE1108Cm	pIE1130
pRAS3.1	100	-	-	-	-
pRAS3.2	-	100	-	-	-
RSF1010K	-	-	94 ± 0	-	-
pIE1108Cm	-	-	-	97 ± 4	-
pIE1130	-	-	-	-	99 ± 1
pRAS3.1 + RSF1010K	100	-	95 ± 0	-	-
pRAS3.2 + RSF1010K	-	100	96 ± 0	-	-
pRAS3.1 + pIE1108Cm	100	-	-	100	-
pRAS3.2 + pIE1108Cm	-	100	-	96 ± 3	-
pRAS3.1 + pIE1130	100	-	-	-	95 ± 8
pRAS3.2 + pIE1130	-	100	-	-	100

plasmids after transformation with either of the pRAS3 plasmids. The loss of pTF-FC2Cm or pTC-F14Cm was not as a result of poor plasmid stability as either plasmid remained perfectly stable in *E. coli* DH5α host cells in the absence of antibiotic selection for the duration of the assay. No colonies were obtained in the reverse transformation experiment wherein either pTF-FC2Cm or pTC-F14Cm served as the incoming selected plasmid and pRAS3.1 or pRAS3.2 as the resident plasmid and for which selection was not provided. The inability to transform *E. coli* DH5α host cells containing the pRAS3 plasmids with a second plasmid was not due to poor transformation competency as compatible vectors such as pUC19 or pACYC177 could readily be transformed into these hosts. The pRAS3 plasmids are thus highly incompatible with pTF-FC2 and pTC-F14 as these plasmids are unable to coexist even in the presence of selection. Furthermore, the incompatibility phenotype is unidirectional as the pRAS3 plasmids displaced the resident IncQ-2 plasmids from the host and also prevented these plasmids from becoming established in a host already containing either of the pRAS3 plasmids.

2.3.10 Displacement of pTF-FC2 and pTC-F14 Is Mediated by ORF3

The strong unidirectional nature of the incompatibility between the pRAS3 plasmids and pTF-FC2 or pTC-F14 suggested that the incompatibility phenotype was not mediated by the iterons of the pRAS3 plasmids. To verify that the iterons were not responsible for the incompatibility phenotype, a 751-bp region containing the *oriV* of pRAS3.1 was cloned by means of PCR into the pGEM-T vector to give pGEM-OriV3.1. The ColEI-based replicon of pGEM-T is not able to replicate in a *polA E. coli* mutant such as *E. coli* GW125a. Therefore, as the cloned *oriV* in pGEM-OriV3.1 was able to sustain replication of the pGEM-T vector in *E. coli* GW125a when pRAS3.1 provided the replication proteins *in trans*, it demonstrated that the cloned *oriV* was functional. No displacement of pTF-FC2 or pTC-F14 from *E. coli* DH5α was, however, observed when pGEM-oriV3.1 was transformed as an incoming plasmid similar to the experiments described above (Fig. 2.10A).

The *repC* gene of pRAS3.1 was previously cloned into the pBAD28 expression vector and the resulting construct, pBAD28-repC, was shown to express a functional protein by its ability to compliment a *repC* deletion mutant (see section 2.3.2). To test whether the RepC protein of pRAS3 was responsible for the displacement of the two IncQ-2 plasmids, the pBAD28-repC plasmid was transformed as an incoming plasmid into *E. coli* DH5α containing resident pTF-FC2 or pTC-F14 plasmids. No displacement of the two plasmids after a round of non-selective growth was evident. Next the *repAC* and *repB* genes of pRAS3.1 were also cloned behind the arabinose inducible P_{BAD} promoter of pBAD28, to give pBAD28-repAC and pBAD28-repB, and placed *in trans* of pTF-FC2 and pTC-F14 to test for a displacement phenotype. Both constructs expressed functional proteins as pBAD28-repAC and pBAD28-repB were able to compliment the *repC* deletion mutant R6K.1.RepC$^{\Delta}$ and interrupted *repB* mutant pRAS3.1::RepB, respectively. Neither the RepA nor the RepB, however, mediated a displacement phenotype.

As the individual replication proteins did not mediate a displacement phenotype I wished to determine whether all three proteins function in concert in order to displace pTF-FC2 and pTC-F14. A 4.3-kb region from pRAS3.1 which includes the *repB*, *pemIK* and *repAC* genes was cloned behind the P_{BAD} promoter to give pBAD28-repBAC and expression of the replication genes was confirmed by its ability to support the replication of R6K-OriV3.1 in *E. coli* DH5α. As before, no displacement was observed when pBAD28-repBAC was transformed as an incoming plasmid into a host cell containing resident pTF-FC2Tet or pTC-F14Km plasmids. These results

demonstrated that the active displacement of the two IncQ-2 plasmids by pRAS3 was not mediated by the replicon components of pRAS3 and suggested that some other unknown protein product or region on pRAS3 was responsible for the displacement phenotype.

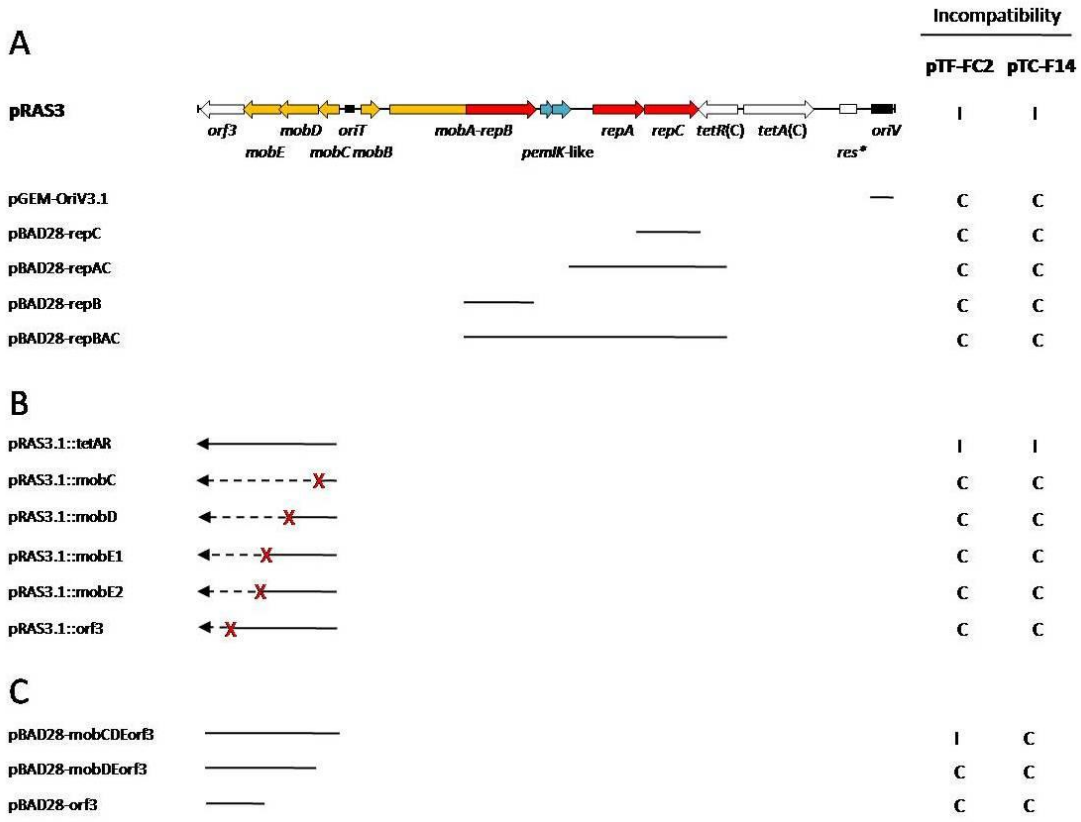


FIG. 2.10. Regions of the pRAS3 plasmids that were tested for incompatibility against pTF-FC2 and pTC-F14. (A) The regions indicated by a solid line were cloned and tested for incompatibility against pTF-FC2 and pTC-F14. All of the constructs displayed a compatible phenotype (indicated by a C). (B) Expression of the genes within the *mobCDE-orf3* operon (indicated by a solid arrow) resulted in an incompatible phenotype (indicated by an I) with both pTF-FC2 and pCT-F14. Interruption of the genes within the operon by insertion of EZ-Tn5 (indicated by a red X) resulted in a compatible phenotype. (C) When the region containing the *mobCDE-orf3* operon (indicated by a solid line) was cloned behind the inducible promoter of pBAD28 it resulted in an incompatible phenotype when coresident with pTF-FC2, but not pTC-F14. Regions containing the *mobDE-orf3* genes or only the *orf3* gene did not result in an incompatible phenotype with pTF-FC2 or pTC-F14, thus suggesting that either the *mobDE-orf3* genes were not translated properly when the *mobC* gene was not present in its upstream position, or that the Orf3 protein required the activity of the MobC protein for displacement.

As the incompatibility determinant remained cryptic, the EZ-Tn5 transposon mutagenesis system was used to generate a bank of random knockouts on pRAS3.1 in an attempt to inactivate and thereby identify the unknown determinant. The bank of plasmid DNA with the randomly inserted EZ-Tn5 transposon was prepared from 3266 colonies that were pooled together. Restriction analysis of extracted plasmid DNA from 24 of the 3266 colonies revealed that only one third of the bank contained pRAS3.1 plasmids with EZ-Tn5 insertions.

The bank was screened for pRAS3.1 mutants that are able to coexist with pTF-FC2 by transforming the plasmid DNA into *E. coli* EC100D electrocompetent cells containing a resident pTF-FC2Cm plasmid. Dilutions of the transformation mixture were spread onto plates containing antibiotic selection for both plasmids, as well as plates containing selection for pRAS3.1 mutants. Approximately 30% (1 653 out of 5.6×10^3 transformants) of the random knockouts displayed a compatible phenotype. The control plasmid pRAS3.1::tetAR (EZ-Tn5 cloned into the *tet*A gene of pRAS3.1) could not be co-transformed with pTF-FC2Cm, while R6K (EZ-Tn5) did not interfere with the replication of pTF-FC2. Two hundred random colonies were transferred onto LA plates containing selection for only the pRAS3.1 mutants. After two rounds of growth on the pRAS3.1 mutant-selective media the colonies were replica-plated onto pTF-FC2-selective media as well as non-selective media to account for cell viability. All 200 of the tested colonies retained pTF-FC2 along with the pRAS3.1 mutant plasmids. Restriction endonuclease analysis using BamHI and SalI of 48 of the 200 mutants indicated that the mutants fell into five groups, with all of the EZ-Tn5 insertions grouped in a ±1.75-kb region. One representative of each of the five groups was isolated and purified as a single plasmid after re-transformation of a crude extract of the mixed plasmid DNA. Nucleotide sequence analysis of the purified plasmid DNA revealed that the EZ-Tn5 transposon insertions were evenly spaced throughout a region containing the *mobCDE* genes as well as the previously unidentified *orf3* gene, with one insertion in each of the genes except for *mobE* which had two (Fig. 2.10B). As *orf3* is the most distal gene in this operon that was inactivated it is the most probable displacement determinant.

DNA fragments containing the *orf3*, the *mobDE-orf3* genes or the entire *mobCDE-orf3* operon were cloned into pBAD28 using a combination of endonuclease restriction sites present on the pRAS3.1 plasmid and sites that were introduced by the EZ-Tn5 transposon to give pBAD28-orf3, pBAD28-mobDEorf3 and pBAD28-mobCDEorf3, respectively. These pBAD28-based

plasmids containing the three different clones as well as a pBAD28 control plasmid were transformed into *E. coli* DH5α host cells containing either a resident pTF-FC2Tet or pTC-F14Km plasmid and screened for a displacement phenotype as before. Both pTF-FC2Tet and pTC-F14Km were stably maintained in the presence of the pBAD28 vector for the duration of the assay. The pBAD28-based clones containing the *orf3* gene or the *mobDE-orf3* genes had no effect on the stability of the IncQ-2 plasmids (Fig. 2.10C). RT-PCR using primers specific to the *mobE* and *orf3* genes verified that the *orf3* in pBAD28-orf3, and both the *mobE* and *orf3* genes on pBAD28-mobDE-orf3, were actively transcribed from behind the P_{BAD} promoter. However, as the function of Orf3 is unknown, no assay was available to check if the *orf3* gene is translated properly when taken out of the *mobCDE* operon. Plasmid pTF-FC2Tet, but not pTC-F14Km, was displaced from the host when a clone containing the entire *mobCDE-orf3* was co-resident. Increased induction of the P_{BAD} promoter by adding 0.1% L-arabinose to the media during a displacement assay did not result in the displacement of pTC-F14Km. As displacement of pTF-FC2 was noticed only when the entire *mobCDEorf3* operon was cloned behind the P_{BAD} promoter, it was possible that either the genes were not translated properly when the *mobC* gene was not present in its upstream position, or that Orf3 required the activity of MobC for it to cause a displacement phenotype.

2.3.11 The Role of ORF3

The 753-bp *orf3* gene is the most promoter-distal gene in an operon containing the *mobCDE* mobilization genes. It does not have a recognizable promoter of its own, its putative start codon overlaps the stop codon of the *mobE* and a possible RBS (GGAGG) sequence is located 5-bp upstream of the ATG codon. Plasmid pTF-FC2 also has a gene, called *orf4*, located in a similar position downstream of *mobE*. The 270-aa product does not however bear any similarity to Orf3. BLAST analysis of the Orf3 sequence (250-aa) returned two strong hits, one of 98% amino acid identity across 229-aa of a hypothetical protein (240-aa) from an uncultured bacterium and one of 97% amino acid identity across 152-aa of a truncated protein, namely OrfX (163-aa), from plasmid Rms149. Plasmid Rms149, which was isolated from *Pseudomonas aeruginosa*, is a 57-kb IncP-6 plasmid (Haines *et al.,* 2005). It thus has a replicon different from that of the IncQ plasmids. Its mobilization system, however, is similar to mobilization systems of pTF-FC2 and pRAS3 (Table 2.3). The five genes (*mobA* through *mobE*) of the Rms149

mobilization system are organized around the *oriT* similarly to the five-component mobilization systems of the IncQ-2 plasmids, and the protein products share between 86 and 100% amino acid identity to the corresponding proteins of pRAS3 and pTF-FC2.

TABLE 2.3. Percentage identity of the mobilization proteins of Rms149 with those of pRAS3 and pTF-FC2

Rms149 Proteins	Percentage amino acid identity	
	pRAS3	**pTF-FC2**
MobA	86	92
MobB	89	92
MobC	98	96
MobD	100	96
MobE	87	92
ORFX	97[A]	14[B]

[A] Comparison is to Orf3

[B] Comparison is to Orf4

As *orfX* of Rms149 was interrupted by a natural insertion of *Tn1012* no function could be assigned to the protein (Haines *et al.,* 2005). Therefore we attempted to determine whether its uninterrupted equivalent on pRAS3 is expressed and whether it has any influence on the mobilization frequency. Total RNA was isolated from *E. coli* DH5α cells containing pRAS3.1 and analyzed by RT-PCR in order to verify whether ORF3 is expressed, as well as whether the protein is translated from the same mRNA transcript as the *mobCDE* genes. Genomic DNA isolated from *E. coli* (pRAS3.1) was used as a positive control for the different primer sets, and expression of the *gapA* gene of *E. coli* was monitored as a positive control for RNA extraction and cDNA conversion. RNA controls were also included to show that there was no DNA contamination in the RNA samples. Amplification products were obtained in the cDNA

samples that were similar in size to products obtained when the gDNA was amplified using the same primer combinations (Fig. 2.11). The results thus show that *orf*3 is indeed expressed, and is done so on the same mRNA transcript as *mobCDE*.

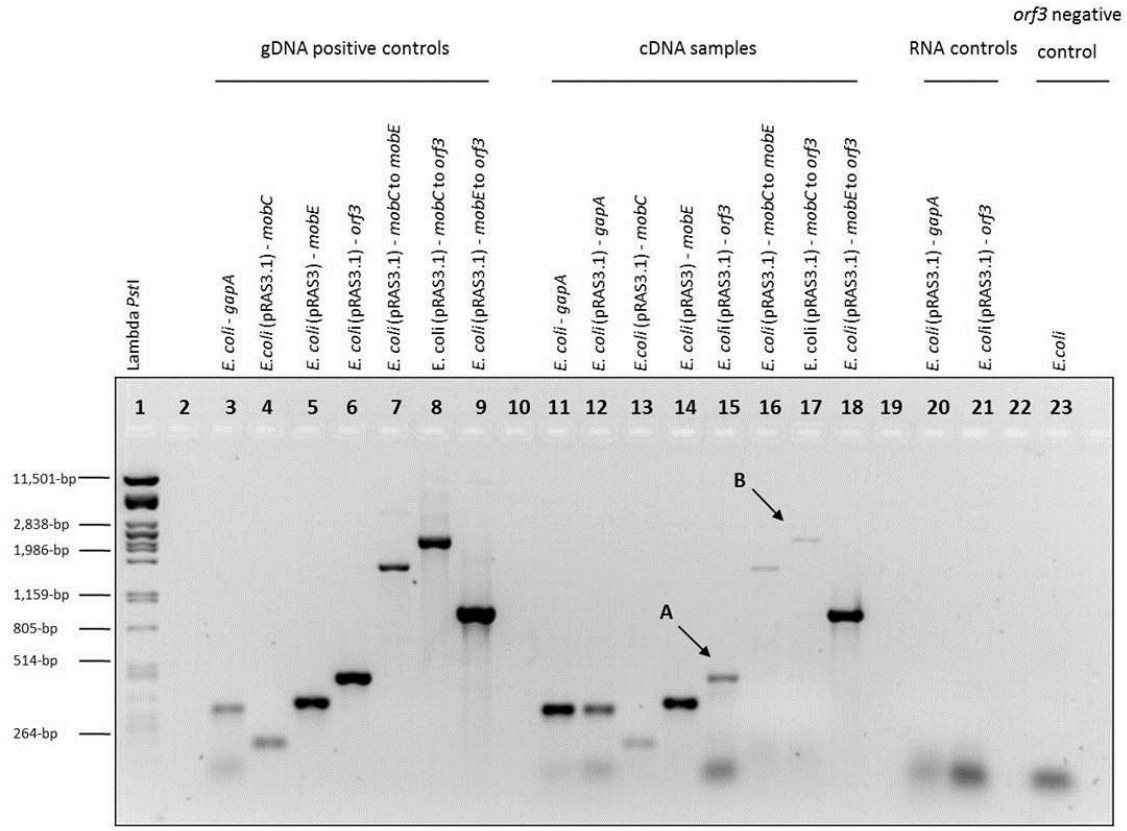


FIG. 2.11. A photo of an agarose gel showing the amplification products of various RT-PCR reactions. The arrows at A and B indicate amplification products from cDNA that show that *orf3* is expressed, and that it is located on a polycistronic mRNA transcript also containing *mobCDE* genes, respectively. Genomic DNA extracted from *E. coli* (pRAS3.1) was used as a positive control for the different primer pairs. The *E. coli* glyceraldehydes phosphatase A (*gapA*) gene was used as a positive control for gene expression and cDNA synthesis. Samples containing RNA were amplified to show that there was no DNA contamination in the samples. An *E. coli* gDNA sample that does not contain pRAS3.1 was included to show that the pRAS3ORF3 primers were specific to the *orf3* gene only. The *E. coli* GAPA, pRAS3MOBC, pRAS3MOBE and pRAS3ORF3 forward and reverse primers were used in different combinations to amplify single genes or regions consisting of multiple genes.

As expression of *orf3* is linked to expression of the mobilization genes we wished to determine whether it influences the mobilization frequency of pRAS3. The pRAS3.1::orf3 and pRAS3.1::mobC random knockouts that were generated while looking for the element that mediates displacement of the two IncQ-2 plasmids, along with pRAS3.1::tetAR (a control to account for possible polar effects caused by the EZ-Tn5 transposon), were used to determine whether inactivation of the *mobC* (and possibly the *mobDE-orf3* downstream genes) and *orf3* genes altered the mobilization frequency. The *E. coli* S17.1 donor cells containing the respective plasmids and *E. coli* CSH56 recipient cells were mixed at a ratio of 1:100 donors per recipient and allowed to mate for 60 min at 37°C. The incubation time was 30 min longer than when WT pRAS3.1 and pRAS3.2 were mated, in anticipation of a lower mobilization frequency. Construct pRAS3.1::tetAR mobilized at a frequency of 12.44 ± 3.95 transconjugants per donor while pRAS3.1::mobC and pRAS3.1::orf3 mobilized at a frequency of 0.08 ± 0.10 and 6.04 ± 4.33 transconjugants per donor, respectively. Inactivation of the *mobC* caused a significant 155-fold ($p = 0.005$) decrease in mobilization frequency compared to the control, thus confirming the deleterious effect of the inactivation by insertion of EZ-Tn5. Inactivation of *orf3*, however, resulted in a non-significant two-fold ($p = 0.132$) decrease in mobilization frequency, thus suggesting that the protein product of *orf3* has a very minor role, if any, in mobilization when the RP4 conjugative system is used.

2.4 DISCUSSION

Careful sequence analysis showed that the *repC* gene, which was previously not annotated, is in fact present in a similar location on the pRAS3 plasmids as other IncQ-like *repC* genes and encodes a protein which is also within the size range of other IncQ-like RepC proteins (283 to 314-aa based on the sequence of the RSF1010 and pIE1130 RepC proteins, respectively). The sequence of the RepC protein, however, bears only 25% sequence identity to the RepC of pTF-FC2 and lacks a conserved sequence pattern (A-L-N-P-x-[IL]-A-[DEQ]-A-[IV]-[IM]-G-[AEG]-x-[PQ]-H-[TV]-R-I-[ENS]-M-[AD]-E-V-R-x-L-[DEQ]-[ST]-[DE]-[APT]-A-R-L-[IL]-H-Q-R-L-C-G-W-I-D-P-G-K-[ST]-G) which is present in the RepC proteins of other plasmids that are representative of the different IncQ subgroups. Due to the extent of the differences, it is more likely that the

repC gene of the ancestral pRAS3 plasmid was acquired by gene exchange with a different, as yet, unidentified plasmid rather than through the accumulation of point mutations.

The RepC protein, as has been shown by others (Kim and Meyer, 1985; Scherzinger *et al*., 1984) and verified here, is essential for plasmid maintenance. In addition, RepC proteins are highly specific for their cognate iterons and the iteron sequences are therefore highly conserved (Lin and Meyer, 1986; Rawlings and Tietze, 2001). As a result, both the *repC* gene and the iterons are under strong selective pressure. Exchange of the *repC* gene would therefore require that the *oriV* also be exchanged during the recombinatorial event. An alternative scenario could be that the *repC* gene was exchanged while the ancestral pRAS3 plasmid existed as a cointegrate with a different unknown plasmid. Cointegrate formation has been suggested as an evolutionary mechanism which allows for mutation or exchange of essential plasmid genes as the selective pressure for maintenance is transferred to the alternate replicon (Sýkora, 1992). During this time both the newly acquired *repC* gene and the iterons could accumulate point mutations until such time that productive binding of the RepC to the iterons restores replication and the cointegrate is resolved.

Integration of an IncQ-like plasmid into a host chromosome or another larger plasmid is not an unlikely event. An approximately 6-kb region of RSF1010 containing the *repA*, *repC, sul2, strA* and *strB* genes has integrated into the large IncI plasmid of *S. enteritidis,* and during a second unrelated event a similar region of RSF1010 but with the *repA* truncated at the 5'-end, integrated into the chromosome of a geographically unrelated *S. typhimurium* ST193 (Daly *et al.,* 2005). In another example, a region of pTC-F14 containing the *oriT* and all the mobilization genes, the *pasAB* TA system as well as partial remnants of the *repA* and *repC* genes have integrated into the large (65-kb) pTcM1 plasmid of *At. caldus* MNG (Van Zyl *et al.,* 2008). Interestingly, a putative *repC*-like ORF encoding a 362-aa protein to which the RepC of pRAS3 bears more similarity (45% identity and 55% similarity) than it does to any of the IncQ-1 or IncQ-2 RepC proteins is also present on this plasmid. The origin of this *repC*-like ORF and its function in pTcM1 is, however, unknown.

A 10.1-kb region of pRAS3.2 was found integrated on the chromosome of seven strains of *Chlamydia suis* by Dugan *et al.* (2004) during a search for genomic islands carrying tetracycline resistance genes within these obligatory intracellular pig pathogens. The size difference between the integrated plasmid and pRAS3.2 is the result of one large ±1.7-kb and two small,

44 and 8-bp, deletions, respectively. The deleted ±1.7-kb region includes 256-bp of DNA coding for the carboxy-terminus of the *mobA* gene, the entire *repB* gene, the promoter of the *pemIK*-like systems as well as 161-bp of the *pemI*-like antitoxin gene. A study on the diversity of TA-systems revealed that genes encoding TA systems are mostly absent from obligatory intracellular prokaryotes and it was proposed that such systems are superfluous in these organisms as a result of their lifestyle (Pandey and Gerdes, 2005). In light of this, it is interesting that the promoter and antitoxin region of the *pemIK*-like genes on the integrated plasmid has been deleted, thus probably rendering the TA system non-functional. Deletion of the smaller 44-bp fragment from the integrated plasmid is equally as interesting as exactly 2 of the 3 × 22-bp iterons were deleted, thus rendering the *oriV* non-functional. Large sections of single stranded DNA are generated during replication of IncQ plasmids (Kok *et al*., 1989; Scherzinger *et al*., 1991) and, therefore, a functional *oriV* integrated onto the chromosome would possibly have resulted in instability of the chromosomal DNA. The 8-bp deletion is located within the operator region of the divergently transcribed *tetAR* genes and resulted in the *tetA* gene being constitutively expressed (Dugan *et al.*, 2004). As *Chlamydia suis* is an obligatory intracellular pathogen of pigs, it occurs in an environment with a constant supply of nutrients and a stable temperature. Thus unregulated expression of the *tetAR* genes, which is presumed to be energy intensive, in an environment where antibiotic use is common practice (Sarmah *et al.,* 2006) would possibly not be detrimental to the cell compared with a cell in a less stable environment.

Early attempts to clone and express the *repC* gene of RSF1010 were unsuccessful unless the *repC* gene was cloned as part of the *repAC* operon, and it was suggested that the *repC* is not translated unless the *repA* gene is translated (Scholz *et al.*, 1989). Similarly, our initial attempts to clone and express the *repC* (without also cloning the *repA*) of pRAS3 were repeatedly unsuccessful (data not shown). Close inspection of the pRAS3 sequence revealed that the *repC* gene does not have its own RBS. An artificial AGGAGG RBS was, therefore, provided 6-bp upstream of the AUG start codon by incorporating it into the forward primer used to amplify the *repC* gene. Cloning of the resulting DNA fragment behind an L-arabinose-inducible promoter resulted in functional expression and translation of the *repC* gene, and thus it seems that co-translation of the *repA* and *repC* may be a general regulatory feature of IncQ-like plasmids.

Real-Time qPCR was used to determine the copy number of pRAS3.1 and pRAS3.2 in *E. coli* and it was estimated to be 45 ± 13 and 30 ± 5 plasmids per chromosome, respectively. This is higher than the estimated copy numbers in *E. coli* of 10 to 12 plasmids per chromosome for RSF1010 (Sakai and Komano, 1996), 10 to 14 copies for pTF-FC2 (Dorrington and Rawlings, 1989) and 12 to 16 copies for pTC-F14 (Gardner *et al*., 2001). The finding that pRAS3.1 has a copy number that is 1.5-fold higher than that of pRAS3.2 was unexpected. The number of iterons within a plasmid has been shown to affect the copy number of the plasmid, and the general trend that has been observed corresponds to a decrease in PCN with an increase in iteron copy number (Chattoraj, 2000; Tsutsui *et al*., 1983). The same trend has also been verified for R1162 as iterons cloned into a second location on the R1162 plasmid resulted in a decrease in copy number (Kim and Meyer, 1985). Therefore, the reasons for pRAS3.1 and pRAS3.2 having such different and unique copy numbers were further investigated and the results are reported in Chapter 3.

The TA system that was discovered on pRAS3 greatly enhanced the stability of a highly unstable R1-based vector (pOU82) from 35% to 98% within an *E. coli* DH5α population over a period of ±72 generations. Sequence analysis revealed that the TA system bears similarity to a number of chromosomally located TA systems, identified during genome sequencing projects, that are distantly related to the MazEF and PemIK TA systems. The MazEF and PemIK systems are homologous and are encoded on the chromosome of *E. coli* and plasmids R1 and R100, respectively (Gerdes *et al*., 2005). The MazE and PemI antitoxins are proteins and thus these systems are classified as Type II TA systems (Van Melderen *et al*., 2009). Chromosomally-encoded TA systems are thought to be involved in diverse functions such as programmed cell death, growth modulation and persistence to name a few (Van Melderen *et al*., 2009). Plasmid-encoded TA systems generally serve to enhance the stability of a plasmid within a population of cells by inhibiting the growth of plasmid-free segregants (Gerdes *et al*., 1986).

The *pemIK* genes of plasmid R100 have been shown to eliminate plasmid-free segregates in order to increase plasmid stability within the population (Tsuchimoto *et al*., 1988). The PemI antitoxin of plasmid R100 has a short half-life compared to the PemK toxin (Tsuchimoto *et al.,* 1992). Approximately 60 min after induction of the pemIK genes, which were cloned behind a heterologous thermostable promoter, the PemI protein had completely disappeared while the PemK toxin was still present. The short half-life of the PemI antitoxin was found to be due to

degradation by Lon protease. The antitoxin protein was unstable in a *lon*-proficient host and stable in a *lon*-deficient host. Also, growth of the host cells was arrested in the *lon*-proficient host after repression of the heterologous promoter while growth remained uninhibited in the *lon*-deficient host. Most host cells could not be revived after the promoter from which the *pemIK* genes were being expressed was repressed, suggesting that the PemK toxin is bacteriocidal rather than bacteriostatic. Killing of the host cell occurs as a result of the sequence-specific endoribonuclease activity of the PemK toxin which leads to inhibition of protein and DNA synthesis (Zhang *et al*., 2004). It was demonstrated, both *in vivo* and *in vitro*, that the PemK protein cleaves single stranded RNA preferentially at the 5' or 3' side of the A residue in the UAH sequences (where H is C, A or U). The MazF toxin is also a sequence-specific endoribonuclease that cleaves specifically single stranded RNA (Zhang *et al.*, 2003) in order to induce cell death in response to stressful conditions such as amino acid starvation, DNA damage and oxidative stress (Amitai *et al*., 2004). The MazF toxin, however, recognizes an ACA sequence and cleaves preferentially at the 3' end of the first A residue (Zhang *et al.*, 2003). As indicated by its sequence similarity to the R100 PemIK and *E. coli* MazEF systems, the TA system of pRAS3 may also be bacteriocidal as plasmid-free segregates were almost never detected during stability assays using the pOU82 system.

The PasABC and PasAB TA systems of pTF-FC2 and pTC-F14 have also been shown to function as effective stability systems as they increased the stability of the pOU82 vector from 33 to 92 and 60%, respectively (Deane and Rawlings, 2004). Although the PasABC system of pTF-FC2 is unusual in that it has an additional PasC protein, it is closely related to the pTC-F14 PasAB system (81 and 72% similarity between the PasA and PasB proteins) and both are deep branching members of the RelBE family of TA systems (Rawlings, 2005). The RelBE toxins are also endoribonucleases, however, they cleave mRNA at the ribosomal A-site in a manner dependent upon the ribosomes and are thus unrelated to the MazEF and PemIK-like systems (Gerdes *et al*., 2005). Therefore, as the TA system of pRAS3 is unrelated to the TA system of pTF-FC2 and pTC-F14 but the flanking replication (*repB*) and mobilization (*mobABCDE*) genes are clearly related, it was probably also acquired during a recombinatorial event. Whether the ancestral plasmid had a TA system related to either the Pas or the PemIK systems, or whether the PemIK and Pas systems were acquired independently of each other is not known.

Both pRAS3.1 and pRAS3.2 were efficiently mobilized by the RP4 conjugative system located on the chromosome of *E. coli* S17.1. The mobilization frequency was in fact so high that the duration of the mating had to be decreased from 60 to 30 min and the ratio of donor to recipient had to be increased from 1:10 to 1:100 to prevent mating saturation and permit the calculation of a mobilization frequency. Although the pRAS3 mobilization frequency was not determined in parallel to pTF-FC2 and pTC-F14, it seems to be higher than that of pTF-FC2 and pTC-F14 (Van Zyl *et al.*, 2003). Van Zyl *et al*. (2003) showed that pTF-FC2 was mobilized by the RP4 conjugative system at a frequency 3,500-fold higher than pTC-F14. They also demonstrated that the mobilization frequency of pTC-F14 could be increased to nearly that of pTF-FC2 by providing it with the MobD and MobE proteins of pTF-FC2. It was suggested the IncQ-2-like mobilization systems are able to adapt to a different conjugative system by changing the accessory mobilization proteins (Van Zyl *et al.*, 2003). Therefore, the higher frequency at which pRAS3 was mobilized compared to pTF-FC2 is may be as a result of point mutations that allow the pRAS3 mobilization proteins to better interact with the RP4 conjugative system.

Incompatibility assays to determine the functional relationship of the two pRAS3 replicons to representative members of the IncQ-1 and IncQ-2 plasmid groups revealed that both pRAS3 plasmids were compatible with the IncQ-1-like plasmids and highly incompatible with the IncQ-2-like plasmids. The nature of the incompatibility was such that the two pRAS3 plasmids were unable to co-exist with either pTF-FC2 or pTC-F14 even in the presence of antibiotic selection. Transformants were only ever obtained when the pRAS3 plasmids were transformed and selected for in host cells already containing a resident unselected pTF-FC2 or pTC-F14 plasmid. Furthermore, when we tested for the presence of pTF-FC2 and pTC-F14 in these transformants we found that they had been completely displaced by the pRAS3 plasmids. As the iterons of IncQ-like plasmids have been shown to be the major incompatibility determinants (Lin and Meyer, 1986; Rawlings and Tietze, 2001; Smalla *et al*., 2000) the pRAS3 *oriV*s were cloned and placed *in trans* of pTF-FC2 and pTC-F14. No displacement was, however, observed and it was evident that some other pRAS3 region or protein product was responsible for the displacement of pTF-FC2 and pTC-F14.

Gardner and Rawlings (2004) found that pTC-F14 and RSF1010 were incompatible and that RSF1010, as an incoming selected plasmid, was able to displace a resident unselected pTC-F14

plasmid. Placement of the RSF1010 iterons *in trans* of pTC-F14 showed that displacement was also not iteron-mediated. A further search for the incompatibility determinant revealed that displacement of pTC-F14 was as a result of what seemed to be non-productive binding of the RSF1010 RepC protein to the *oriV* of pTC-F14, thereby inhibiting replication of the competing plasmid. Plasmids pTF-FC2 and pTC-F14 were also found to be incompatible (Deane and Rawlings, 2004). Co-transformed cells were grown for ~100 generations in the absence of selection for both plasmids before being tested for the retention of either plasmid. After this period less than 5% of the tested colonies retained pTC-F14 whereas more than 95% retained pTF-FC2. A pTF-FC2 plasmid with the *pasABC* system deleted was less able to displace pTC-F14 (75% pTC-F14 retention). Experiments wherein the pTF-FC2 PasABC proteins were placed *in trans* of the *pasAB* promoter of pTC-F14 fused to a β-galactosidase reporter gene system confirmed that the PasABC proteins of pTF-FC2 are able to repress the pTC-F14 *pasAB* promoter. It was therefore suggested that the displacement phenotype was as a result of repression of the pTC-F14 *pasAB* promoter which lead to decreased transcription of the *repAC* genes. These incompatibility experiments demonstrated that non-productive binding of replicon components or cross-regulation of promoter regions can also confer an incompatible phenotype.

In light of the work by these researchers, I wished to determine whether displacement of pTF-FC2 and pTC-F14 was perhaps mediated by non-productive binding of the replication proteins of pRAS3 to the *oriV* or other regulatory regions of the displaced plasmids. Various constructs containing the *repC, repA* or *repB* genes cloned individually or in combination behind an inducible promoter on pBAD28 were tested for an ability to displace resident unselected pTF-FC2 or pTC-F14 plasmids from the host. No displacement phenotype was however observed even though expression of the genes was confirmed using various deletion mutant controls.

As the determinant on pRAS3 responsible for the displacement remained elusive, the EZ-Tn5 mutagenesis system was used to create a bank of random gene knockouts. A ±1.7-kb region containing the *mobC, mobD, mobE* and an ORF with an unknown function (*orf3*) was identified that contained all the EZ-Tn5 insertions which rendered pRAS3 compatible with pTF-FC2 and pTC-F14. The *orf3* gene was the most promoter-distal gene in the operon to contain an EZ-Tn5 insertion, thus suggesting that the protein product of *orf3* was responsible for the displacement of pTF-FC2 and pTC-F14. To verify that it was *orf3* that was responsible for the

displacement phenotype, the corresponding unmutated 1.7-kb region containing all four genes was cloned behind the P_{BAD} promoter and tested as an incoming plasmid *in trans* of pTF-FC2 and pTC-F14. Only pTF-FC2 was, however, displaced. When smaller subclones of the 1.7-kb region containing either the *mobDE-orf3* genes or only the *orf3* gene were tested for a displacement phenotype, the phenotype was no longer present. RT-PCR experiments verified that the genes in all the constructs were transcribed from behind the P_{BAD} promoter even without addition of L-arabinose to the LB media. However, as the function of the Orf3 protein is unknown we were unable to determine whether the cloned *orf3* gene was properly translated from each of the constructs. Only pTF-FC2 and not pTC-F14 was displaced when pBAD28-mobCDE-orf3 was placed *in trans*. This could mean that, if Orf3 was only produced in small quantities, pTF-FC2 is more susceptible to the inhibitory effects of Orf3 than pTC-F14. That displacement of pTF-FC2 was observed when the entire *mobCDE-orf3* operon was placed *in trans* and not when only *mobDE-orf3* or *orf3* was placed *in trans*, suggests that Orf3 might require the activity of MoBC to cause a displacement phenotype. One possible means by which the Orf3 protein of pRAS3 could have mediated the displacement of pTF-FC2 was through non-productive binding to the *oriT* region. The promoter for transcription of the genes involved in replication as well as mobilization for both pTF-FC2 and pTC-F14 is located between the *oriT* and the *mobB* (Gardner and Rawlings, 2004; Frey *et al*., 1992), and therefore non-productive binding in this region could have inhibited transcription and thus replication. However, Orf3-*oriT* binding experiments would be needed to establish this.

Sequence comparisons of the replication and mobilization genes indicate that pRAS3 belongs to the IncQ-2 subgroup, which until the discovery of the pRAS3 plasmids consisted of only pTF-FC2 (IncQ-2α) and pTC-F14 (IncQ-2β) (Gardner *et al*., 2001; Rawlings and Tietze, 2001). As the pRAS3 plasmids did not exhibit replicon-based incompatibility towards any of the IncQ-1 and IncQ-2 plasmids against which they were tested, we propose that they be classified into a new IncQ-2γ subgroup. The need for a separate group for the pRAS3 plasmids is also made evident by the lack of similarity of its RepC to other IncQ-like RepC proteins as well as its copy number which is different from that determined for other IncQ-like plasmids. Finally, the mosaic nature of the pRAS3 plasmids as a result of possible gene exchange events, along with their global distribution (L'Abee-Lund and Sørum, 2002) demonstrates their active participation in the horizontal gene pool and divergence from other IncQ-like plasmids.

Chapter 3

Reason for Copy Number Differences Between pRAS3.1 and pRAS3.2

Table of Contents

3.1 INTRODUCTION

The iterons within the *oriV*s of almost all theta-replicating plasmids (e.g. F, R6K, RK2, prophage P1, etc) including the *oriV*s of the IncQ plasmids, are the binding sites for the respective replication initiator proteins of these plasmids (Germino and Bastia, 1983; Kittell and Helinski, 1991; Lin and Meyer, 1986; Pal *et al*., 1986; Rawlings and Tietze, 2001; Stalker *et al.,* 1979; Tsutsui *et al*., 1983). Apart from the role of iterons in initiation of replication, iterons have also been implicated in the control of the upper limit of plasmid copy number (Chattoraj, 2000; Chattoraj *et al.,* 1984; Chattoraj *et al.,* 1985; Kittell and Helinski, 1991; Paulsson and Chattoraj, 2006). By inserting additional iterons elsewhere on these plasmids, or by inactivating some of the iterons within their *oriV*s, the PCN could be decreased or increased, respectively (Paulsson and Chattoraj, 2006; Park *et al.,* 2001; Tsutsui *et al*., 1983). Cloning of a second *oriV* containing a full set of iterons elsewhere on the IncQ-1 plasmid R1162 resulted in a decrease in the copy number of the plasmid (Kim and Meyer, 1985). Furthermore, when 1, 2 or 3 × 22-bp iterons were cloned onto a vector and placed *in trans* of the parent R1162 plasmid, it had a reducing effect on the copy number of the 1162 plasmid (Lin and Meyer, 1986). One iteron on a multicopy vector *in trans* resulted in a slight but visible decrease in copy number while 3 × 22-bp iterons *in trans* had the most drastic effect. The concentration of the iterons within a cell, therefore, serves as a means for the plasmid to sense its copy number.

Most of the IncQ-like plasmids sequenced and characterized to date have only 3 perfectly conserved iterons and an average PCN between 10 to 16 plasmids per chromosome in *E. coli* (Rawlings and Tietze, 2001). In the event that a functional *oriV* contains more than 3 × 22-bp iterons the additional iterons are partly deleted, contain point mutations or are incorrectly spaced as illustrated in Figure 3.1. The IncQ-2-like plasmid pRAS3.1 is an exception as, unlike the other IncQ-like plasmids including its isogenic counterpart pRAS3.2, it contains 4 fully conserved and correctly spaced iterons within its *oriV*. As a contribution to our studies on the evolution of IncQ-like plasmids, our long term aim is to address the question of why two natural versions of pRAS3 exist. To do so I needed to determine whether the number of iterons within the *oriV* have a similar influence on PCN as has been demonstrated for other iteron-containing plasmids, and if so, why pRAS3.1 has a copy number that is ~1.5-fold higher than that of pRAS3.2 even though it has 4 × 22-bp iterons within its *oriV* compared to 3 for pRAS3.2, and thus theoretically, may be expected to have a lower copy number.

FIG. 3.1. Comparison of the number of pRAS3 iterons to the number of functional iterons present in plasmid representatives of each of the IncQ-subgroups; IncQ-1α (RSF1010), -β (pIE1107), -γ (pIE1130) and IncQ-2α (pTF-FC2) and -β (pTC-F14). Conserved iterons are indicated by arrows above the iterons. The 2-bp spacers are blue and underlined. Partial or mutated iterons are red and point mutations are indicated by an asterisk (*) above the mutated nucleotide.

Various methods are available to accurately estimate the average number of plasmids per host cell (absolute copy number) or to measure the difference in copy number between two or more isogenic plasmids (relative copy number). Techniques for determining the absolute copy number require that both the amount of plasmid and chromosome in a sample be quantified and the results are expressed as the number of plasmids per chromosome. In relative copy number experiments the amount of plasmid in a sample is compared to a reference sample and thus the results are expressed as a fold-difference relative to the reference sample. To be able to accurately compare two different samples a control is included that allows the samples in an experiment to be normalized. Thus as differences between samples can be more easily accounted for in relative quantification experiments than in absolute quantification experiments, the relative quantification techniques tends to be more robust. Although variations exist, the two general techniques for absolute and relative quantification of PCN are densitometric analysis and quantitative PCR (qPCR).

Densitometric copy number analysis involves electrophoresis of DNA samples of known concentrations on an agarose gel to separate the DNA bands of interest and estimate their relative contributions to the total amount of DNA by comparing the intensity of their

respective fluorescence signals (Projan *et al.,* 1983; Pushnova *et al.,* 2000). If the experiment is set up to determine the absolute copy number of a plasmid, the ratio of plasmid DNA to chromosomal DNA in a total genomic DNA (gDNA) preparation is determined based on the respective sizes of the plasmid and the chromosome and the relative amount (out of the total) of each present on the agarose gel. If the experiment is set up to determine relative differences in the copy numbers of isogenic plasmids the intensities of linearized plasmid DNA bands on an agarose gel are compared to each other (Park *et al.,* 2001). In such an experiment a second plasmid with a copy number that remains unchanged in each of the samples is included to normalize the samples. Variations of this technique include capillary electrophoresis (Schmidt *et al.*, 1996) as opposed to gel electrophoresis, quantification of [^{3}H]thymine-labeled plasmid and chromosomal DNA after separation by ethidium bromide CsCl equilibrium centrifugation (Barth and Grinter *et al*., 1974) and analysis of the relative amounts of plasmid and chromosomal DNA in total DNA samples by comparing signal intensities in Southern hybridization experiments (Dorrington and Rawlings, 1989; Gardner *et al*., 2001). The latter two techniques were used to determine the absolute copy numbers of the IncQ-1 plasmid R300B (similar to RSF1010) (Barth and Grinter *et al*., 1974), and the IncQ-2 plasmids pTF-FC2 (Dorrington and Rawlings, 1989) and pTC-F14 (Dorrington and Rawlings, 1990) respectively.

Real-Time qPCR presents an alternate method for both absolute and relative quantification of plasmid copy number (Gerdes *et al*., 2005; Lee *et al.,* 2006a; Lee *et al*., 2006b). Specific targets on the plasmid and chromosomal DNA within a sample are amplified by means of PCR and the increase in target amplification product is monitored in real-time by measuring fluorescence signals at the end of each amplification cycle (Logan *et al.,* 2009). The fluorescence signal is emitted either during the hydrolysis of fluorescent TaqMan probes or by the excitation of intercalated SYBR Green dye (PfaffI, 2004). TaqMan probes are specific to each amplicon and thus allow quantification of the plasmid and chromosomal DNA in the same reaction, but are more expensive and less informative with regards to the formation of primer dimers and non-specific PCR products. Primer dimers and non-specific PCR products sequester primers away from the gene target and thereby reduce the amplification efficiency. SYBR Green dye intercalates with any double stranded DNA and therefore requires that the plasmid and chromosomal DNA be amplified in separate reactions, but it does permit monitoring of the formation of primer dimers and non-specific PCR products which would, in this instance,

possibly result in an artificially high amplification efficiency. As the amplification process is monitored in real-time the relative amounts of starting material in each of the samples can be inferred by comparing the number of cycles required for the amplification product to reach a concentration threshold well before the reaction reached saturation as during conventional PCR.

If the Real-Time qPCR experiment is designed to determine absolute copy number, a standard curve is included for each of the targets. A standard curve is a plot of the threshold cycle (C_T) values obtained from amplification of serial dilutions of template DNA with known concentrations against the Log of the number of molecules (calculated from the molecular weight and concentration) in each reaction. The number of each of the plasmid and chromosomal DNA molecules in the total DNA samples is then extrapolated from the respective standard curves using the C_T values obtained for each target, and the PCN is determined by calculating the ratio of the number of plasmid molecules to the number of chromosome molecules (Lee *et al.,* 2006a). If the goal is to determine differences in the copy numbers of isogenic plasmids the amount of plasmid DNA (represented by its C_T value) present in each of the samples is compared to that of a reference and the difference is expressed as a fold-difference (PfaffI, 2004). Trends can be readily identified using this method but it requires careful selection of the reference plasmid as well as normalization of the samples so that they can be readily compared. To compensate for variations in gDNA extraction and input in each of the reactions, a second target, referred to as the normalizer, which remains unchanged in all the samples is amplified. The difference in the C_T values of each of the normalizer reactions is subtracted from the difference in the C_T values between the reference and the target samples (referred to as the $\Delta\Delta C_T$ method) in order to adjust all the samples so that they are comparable (PfaffI *et al.,* 2002). Targets such the 16S rDNA or *gapA* (encodes the D-glyceraldehyde-3-phosphate dehydrogenase) genes of *E. coli* are often chosen as the normalizer in gene expression assays (Huggett *et al.,* 2005; Tao *et al.*, 2005), however, in essence any region of the chromosome which can be readily amplified by PCR will suffice. As PCR reactions using different primers do not always attain the theoretical maximum amplification efficiency, i.e. a doubling in the amount of product after every cycle, the amplification efficiencies (calculated using standard curves) of each of the primer sets used in an experiment are accounted for during the normalization process (Logan *et al.,* 2009).

Absolute quantification Real-Time PCR combined with TaqMan chemistry has been applied to determine the copy number of a cryptic plasmid present in the infectious elementary bodies (EBs) of *Chlamydia trachomatis* (Pickett *et al.,* 2005). The average PCN of the plasmid in the EBs was found to be 4.0 ± 0.8 plasmids per chromosome (± 95% confidence interval) and it was found to increase up to 7.6 plasmids per chromosome during the developmental cycle. Additionally a plasmid similar to that of *Chlamydia trachomatis* but with a deletion in one of the ORFs thought to be involved in, but not essential for, replication was also present in *Chlamydophila pneumonia*. Its copy number, however, was found to be 1.3 ± 0.2 plasmids per chromosome (mean ± 95% confidence interval), thus supporting their hypothesis regarding the involvement of the unknown ORF in replication. Matcher and Rawlings (2009) applied relative quantification Real-Time PCR combined with SYBR Green chemistry to monitor changes in the PCN of pTF-FC2 in response to specific deletions in its *pasABC* operon. This operon, which represents a plasmid toxin-antitoxin system (see Chapter 1 section 1.2.6.2) was also thought to influence the level of transcription of the downstream *repAC* genes and thus, the availability of the replication proteins. They showed that deletion of the entire *pasABC* operon did not significantly affect the plasmid copy number ($p \geq 0.001$) as a result of read-through transcription from a promoter even further upstream. If, however, the genes of the *pas* operon but not the promoter were deleted the PCN decreased by 2-fold ($p \leq 0.001$) relative to the control as a result of unregulated expression of the *pas* promoter. Furthermore they demonstrated the effect of the availability of the RepA and RepC proteins on plasmid copy number by showing that the copy number increased rapidly in response to small increases in transcription of the *repAC* genes until a plateau was reached, whereafter higher *repAC* transcription levels resulted in a rapid decrease in PCN, similar as to when the *pasABC* promoter was unregulated. With these copy number experiments, as well as other gene transcriptional regulation experiments, they demonstrated that the *pasABC* operon of pTF-FC2 is strategically positioned so as to be integrated into the copy number regulatory system of the plasmid in addition to its role as a plasmid stability system (Matcher and Rawlings, 2009).

These and other experiments (Lee *et al.,* 2006a; Skulj *et al.,* 2008; Tao *et al.*, 2005; Turgeon *et al.,* 2008) demonstrate the versatility of Real-Time PCR to measure with a high degree of sensitivity either the absolute copy number of plasmids or differences in the relative copy numbers of similar plasmids. Furthermore, the experimental design allows for statistical validation of the results in order to increase confidence levels. Having said that, as a result of

the high sensitivity of the assay, differences in experimental design and total DNA extraction methods Real-Time PCR has been accused of poor reproducibility and thus strict guidelines such as those presented in the recently published MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines need to be followed to ensure accuracy and reproducibility (Bustin *et al.,* 2009).

The copy numbers of pRAS3.1 and pRAS3.2 were estimated to be 45 ± 13 and 30 ± 5 plasmids per chromosome, respectively, by means of Real-Time qPCR (see Chapter 2 section 2.3.5). These two plasmids differ only in the number of 22-bp iterons within their *oriV*s and the number of 6-bp repeats in an intergenic region between the *oriT* and the *mobB*. In this chapter a set of derivative plasmids were created from the pRAS3.1 backbone to isolate the contribution of each of these loci on the copy number of the two WT pRAS3 plasmids. Copy number experiments were done by relative Real-Time qPCR and a pBR322 cloning vector with a known copy number was included in the assays to verify the absolute copy numbers that were previously estimated for the WT pRAS3.1 and pRAS3.2 plasmids. By including the pBR322 control it was also possible to extrapolate, from the relative copy number data, absolute copy numbers for each of the derivative plasmids that were created. The results confirmed that plasmid copy number decreases when the number of iterons within the *oriV* of a pRAS3 is increased. It was also shown that the additional 6-bp repeat within the *oriT-mobB* intergenic region of pRAS3.1 was responsible for increased transcription of the downstream *repB* gene, and that this was the reason for pRAS3.1 having a higher copy number than pRAS3.2

3.2 MATERIALS AND METHODS

3.2.1 Bacterial Strains, Plasmids, Media and Growth Conditions. For information on the bacterial strains, plasmids, media and additives as well as the growth conditions used refer to Chapter 2 section 2.2.1.

3.2.2 General DNA Techniques. Plasmid preparation, restriction endonuclease digestions, gel electrophoresis and cloning were carried out using standard techniques (Ausubel *et al.,* 1993; Sambrook *et al.,* 1989). Plasmid, total DNA and cDNA samples were divided into aliquots and

stored at -20°C when kept for extended periods. Aliquoted RNA samples were stored at -80°C. DNA and RNA concentrations were determined on a NanoDrop™ spectrophotometer (Thermo Scientific) at 260 and 280 nm, respectively. The degree of RNA and protein contamination in the total DNA samples was determined by measuring the ratios of the 260/280 and 260/230 nm wavelengths respectively. The integrity of the DNA in the total DNA samples was assayed on a 0.8% (w/v) agarose gel in 1 × TBE buffer while the integrity of RNA samples was visualized on a 1% (w/v) agarose/18% (v/v) formaldehyde denaturing gel in 1 × MOPS electrophoresis buffer (Appendix A).

3.2.3 Relative plasmid Copy Number Determinations Using Real-Time qPCR. Total gDNA was extracted from cells growing in the mid-logarithmic phase using a QIAamp® DNA Mini Kit (QIAGEN), and the Real-time qPCR amplification was performed using a LightCycler™ (version 2.0) instrument with the LightCycler FastStart DNA Master SYBR Green I chemistry (Roche Diagnostics) as described in Chapter 2 section 2.2.4.

Plasmid pRAS3.1.35 (Appendix B) was chosen as a reference plasmid for the relative quantification of pRAS3.1 and its derivatives, pRAS3.2 and pBR322 according to the method of Lee *et al*. (2006a) using the Relative Expression Software Tool – 384 (REST – 384 ©) (PfaffI *et al.,* 2002). The REST – 384 software uses a Pair Wise Fixed Reallocation Randomization Test© to determine statistically significant relative differences between samples. The single copy *gapA* gene on the chromosome of *E. coli* was chosen as a target for the normalization of samples using the *E. coli* GAPA primer set (Appendix C). The plasmid DNA fragment was amplified using the pRAS3A primer set (Appendix C). The amplification efficiencies, required for sample normalization, for each of the *E. coli* GAPA and pRAS3A primer sets was calculated (using the REST – 384© software tool) from the standard curves that were generated during the absolute quantification experiments (see Chapter 2 section 2.2.4).

3.2.4 Relative Plasmid Copy Number Determinations Using Densitometric Analysis. Changes in relative plasmid copy numbers were also visualized and compared by means of densitometric analysis (Park *et al.,* 2001). Plasmid-containing *E. coli* DH5α cultures were grown in antibiotic-containing LB media at 37°C overnight, whereafter the cultures were diluted 1.0 × 10^3-fold into fresh LB without antibiotics and grown till they reached mid exponential phase (OD_{600} ± 0.8). The cells (4 ml) were harvested by centrifugation and resuspended in an equal

volume PBS (pH 7.4). The OD_{600} of each of the pRAS3-containing cultures were repeatedly measured and adjusted to an absorbance value of between 0.5 and 0.7 using PBS until the cell densities of all the cultures were close to equal. 1.5 ml of each of the pRAS3-containing cultures was mixed with 0.3 ml of a pUC19-containing culture and the plasmid DNA was extracted from the mixed cultures using a Wizard SV Plasmid Purification Kit (Promega) as per manufacturer's protocol. Each of the samples were split into three equal volumes and the purified plasmid DNA in each of the samples was linearized using HindIII restriction endonuclease. The fragments were separated in a 0.8% (w/v) agarose/TBE gel.

3.2.5 Relative Gene Expression. Total RNA was extracted using a RiboPure™ Bacteria Kit (Ambion Inc.) from *E. coli* EC100D (pR6K.3.1.repC$^{\Delta}$) and *E. coli* EC100D (pR6K.3.2.repC$^{\Delta}$) plasmid-containing *E. coli* cultures and converted to cDNA using a Transcriptor First Strand cDNA Synthesis Kit (Roche) as described in Chapter 2 section 2.2.9. The cDNA samples were diluted 2-fold in ddd.H_2O prior to use. Differences in the cDNA levels of the target genes were measured by Real-Time qPCR using a LightCycler™ (version 2.0) instrument with the LightCycler FastStart DNA Master SYBR Green I kit (Roche Diagnostics) using the reaction conditions described in Chapter 2 section 2.2.4. The Real-Time qPCR data was analyzed using the Relative Expression Software Tool – 384 (REST – 384 ©) (PfaffI *et al.,* 2002).

The R6KKANR primer set was used to amplify the cDNA originating from the kanamycin resistance gene of the EZ-Tn5 transposon in pR6K.3.1.repC$^{\Delta}$ and pR6K.3.2.repC$^{\Delta}$ and served as the normalizer. The pRAS3REPB2 and pRAS3ORF3 primer sets were used to amplify the *repB* and *orf3* gene targets respectively. The standard curves used to determine the amplification efficiencies of the R6KKANR and pRAS3REPB primer sets consisted of pR6K.3.1.repC$^{\Delta}$ cDNA that was serially diluted from 10^{0} to 10^{-4}. As *orf3* cDNA levels were not detected in dilutions greater than the 10^{-2} dilution its amplification efficiency was determined using a standard curve that was made up of serially diluted pRAS3.1 plasmid DNA (4.0^{-1} to 4.0^{-5} ng per reaction).

3.3 RESULTS

3.3.1 Construction of pRAS3-Derivatives with Increased Iteron Copy Numbers.

The presence of fortuitous BstEII endonuclease restriction sites at the end of each iteron, but absent from the rest of the plasmid, made it possible to create a set of plasmid derivatives containing different iteron copy numbers which could be used to investigate changes in PCN as a result of having varying numbers of iteron repeats within a functional *oriV* (Fig. 3.2). Plasmid pRAS3.1 DNA was cut with BstEII for 1 hour at 37°C where after the restriction endonuclease enzyme was heat inactivated at 65°C for 20 min. The reaction was then split and volumes of 2, 4 and 8 μl of the digested DNA reaction mix were used to carry out a concentrated self-ligation overnight at 15°C using T4 DNA Ligase. The ligation reaction mixtures were transformed into chemically competent *E. coli* DH5α cells and 100 μl of each transformation was spread onto LA media containing tetracycline. A total of 451 transformants were obtained after 2 days of incubation at 37°C, and the number of transformants obtained for each transformation varied proportionally according to the amount of digested DNA that was included in the ligation reaction. No transformants were obtained in the control reaction in which the T4 ligase enzyme was substituted with ddd.H_2O, thus indicating that the plasmid DNA was digested to completion and all the transformed colonies were as a result of the ability of re-ligated plasmid DNA molecules to replicate.

The plasmid DNA from 80 randomly selected transformants was extracted and screened by PCR for fragment length polymorphisms, as a result of insertion or deletion of 22-bp iteron units, using single strand DNA primers (pRAS3oriV Fwd and pRAS3oriV Rev, Appendix C) homologous to either side of the *oriV*. The majority of the bands appeared to have sizes similar to the PCR amplified *oriV*s of pRAS3.1 (751-bp) and pRAS3.2 (729-bp) containing 4 and 3 × 22-bp iterons respectively. Four pRAS3.1 derivatives that appeared to have different sized amplicons were identified and DNA sequencing of the *oriV* regions (using the pRAS3OriV2 Fwd primer, Appendix C) revealed that plasmids containing 3-, 4-, 5- and 7 × 22-bp iterons had been created (Fig. 3.2). The plasmids are numbered pRAS3.1.35, pRAS3.1.45, pRAS3.1.55 and pRAS3.1.75 with the last two digits indicating the number of 22-bp iterons and 6-bp repeats respectively. Using this terminology, derivative plasmid pRAS3.1.45 is equivalent to the WT pRAS3.1 plasmid.

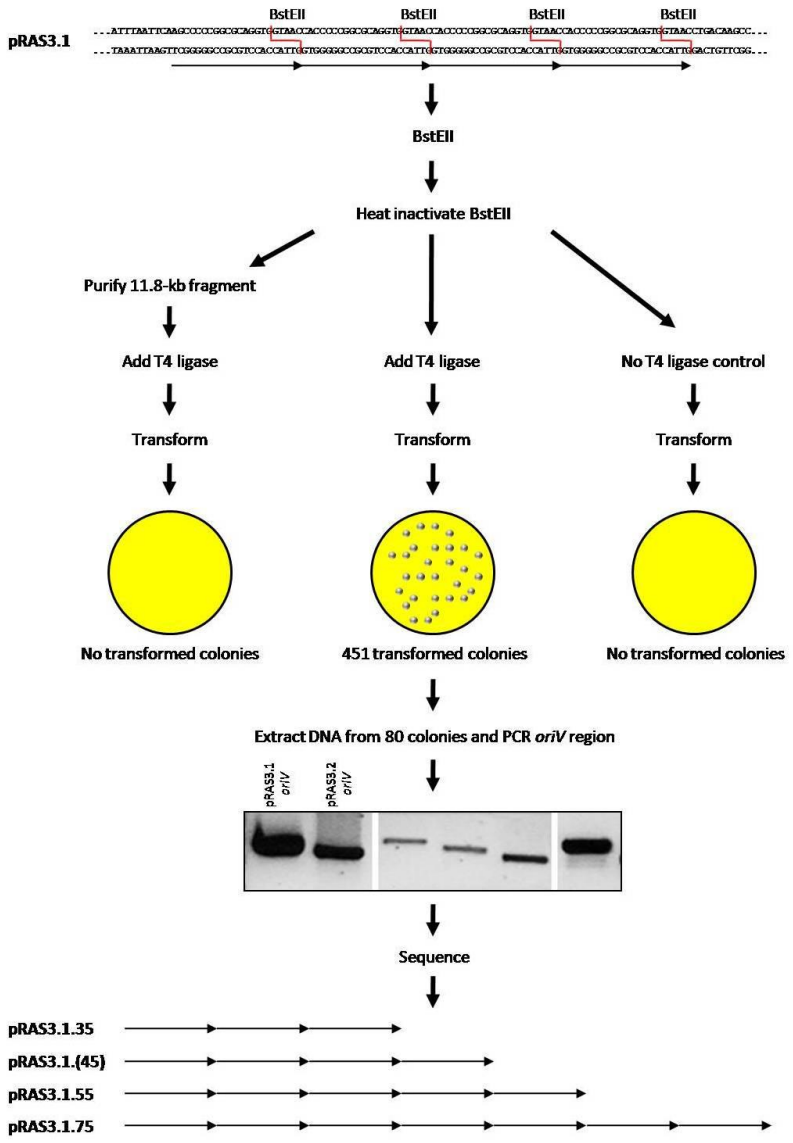


FIG. 3.2. Construction of pRAS3.1 derivatives with 3-, 4-, 5- and 7 × 22-bp iterons. The iterons within the *oriV* of pRAS3.1 (top) and its derivatives (bottom) are indicated by repeated arrows with the number of arrows indicative of the number of iterons. The position at which the BstEII restriction endonuclease enzymes cleavage occurred within each of the iterons are indicated by red lines. The selective LA plates on which the transformed cultures were plated out are indicated by yellow-filled circles, and the presence of transformed *E. coli* colonies are represented by small off-white circles on the plates. The pRAS3.1 and pRAS3.2 *oriV* agarose gel labels indicate the 742- and 720-bp PCR amplification products of pRAS3.1 and pRAS3.2, respectively, using the pRAS3OriV primers. The other four bands on the photo of the agarose gel are the PCR products of the *oriV* regions of the clones that were sequenced using the pRAS3OriV2 primer (Appendix C).

In a parallel experiment, the large 11,785-bp fragment containing a single remaining iteron after digestion with BstEII was separated from the smaller 22-bp fragments by electrophoresis on a 0.8% agarose gel prior to self-ligation in an attempt to generate a plasmid containing only 1 × 22-bp iteron (Fig. 3.2). No colonies were obtained after the ligation reaction was transformed (despite a pRAS3.1 transformation control showing that the *E. coli* DH5α cells were competent). Thus, as no plasmids containing either 1 or 2 × 22-bp iterons were identified, plasmids with less than 3 × 22-bp iterons appeared to be non-viable.

3.3.2 Plasmid Copy Number Decreases with Increasing Iteron Copy Number.

The general trend for iteron-containing plasmids is to have a lower plasmid copy number in response to having more iterons (Kim and Meyer, 1985; Park *et al.,* 2001; Paulsson and Chattoraj, 2006; Tsutsui *et al*., 1983). To determine whether the same trend is valid for the pRAS3 plasmids, the WT pRAS3.1 plasmid and its derivatives with different iteron copy numbers were subjected to relative quantification Real-Time PCR experiments. Total DNA was extracted from actively growing *E. coli* DH5α cells containing pRAS3.1.35, pRAS3.1, pRAS3.1.55 and pRAS3.1.75. The DNA concentrations in the samples were standardized and equal amounts were subjected to qPCR amplification using two sets of forward and reverse primers (*viz*. *E. coli* GAPA and pRAS3A) specific to the *E. coli* chromosome and the pRAS3 plasmids, respectively (see Chapter 2 section 2.3.5 for amplification efficiencies). The *E. coli* GAPA amplicon served to normalize concentration variations between the samples within the reactions. The sample containing the pRAS3.1.35 plasmid, which has the fewest iterons and thus was predicted to have the highest copy number, was used as the reference sample against which all the other plasmid copy numbers were compared. The *tetAR* tetracycline resistance genes of the well known cloning vector pBR322 is identical to that of pRAS3.2 and, in addition, its PCN was estimated by Lee *et al*. (2006a) to be 18 plasmids per chromosome using both absolute and relative Real-Time PCR quantification methods. For these reasons pBR322 was included as a standard from which to extrapolate the PCN of each pRAS3.1 variant in the relative qPCR experiments.

The PCN of pRAS3.1 (4 × 22-bp iterons) and its derivatives with 5- and 7 × 22-bp iterons were 0.69 ± 0.062, 0.48 ± 0.092 and 0.32 ± 0.070 relative to that of the 3 × 22-bp iteron (pRAS3.1.35)

plasmid ($p \leq 0.001$). The control plasmid pBR322 had a copy number of 0.30 ± 0.007 relative to that of pRAS3.1.35 ($p \leq 0.001$). From the relative copy number data and assuming a PCN of 18 plasmids per chromosome for the pBR322 standard, the copy number of the WT plasmid pRAS3.1 was determined to be 41 ± 4 plasmids per chromosome. It was thus in good agreement with the PCN of 45 ± 13 plasmids per chromosome that was estimated for pRAS3.1 in a previous experiment using absolute quantification techniques. Accordingly, a reduction in the number of 22-bp iterons from four, present in the WT pRAS3.1, to three (pRAS3.1.35) resulted in a ~44% increase in PCN (PCN 59) while an increase in the number of iterons from the 4 present in pRAS3.1 to 5 (pRAS3.1.55) or 7 (pRAS3.1.75) resulted in a decrease in copy number of ~32% (PCN 28) and ~54% (PCN 19) respectively.

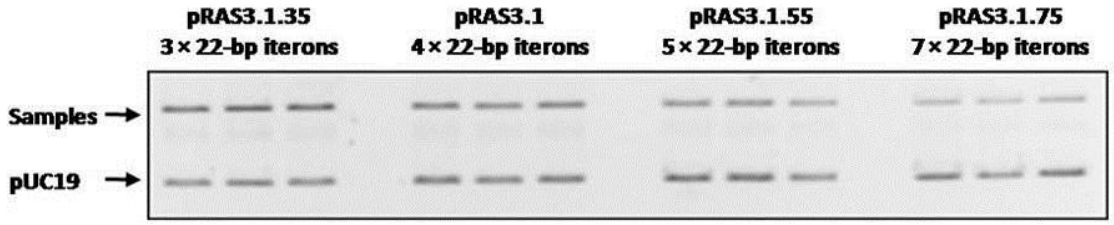


FIG. 3.3. An increase in iteron copy number brings about a corresponding decrease in PCN. The plasmid DNA was extracted from *E. coli* DH5α cultures with equal cell densities, and thus the decrease in fluorescence of the linearized pRAS3.1.35, pRAS3.1, pRAS3.1.55 and pRAS3.1.75 DNA in response to an increase in iteron copy number was indicative of a corresponding decrease in PCN respectively. The fluorescent intensity of the pUC19 band was similar for each sample and thus indicated that the extraction efficiencies were similar.

The relative changes in PCN for pRAS3.1 and its three derivatives were verified by a second method in which the amount of plasmid DNA extracted from approximately equal cell numbers was visually compared on an agarose gel. Each of the cell suspensions were spiked with an equal volume of an *E. coli* DH5α culture containing the pUC19 vector prior to plasmid DNA extraction to account for differences in extraction efficiency between the samples. The fluorescent intensity of the linearized pUC19 band was approximately equal in all the extractions, indicating that the extraction efficiencies between the different samples were

similar. A clear trend indicating a decrease in PCN in response to an increase in iteron copy number was evident (Fig. 3.3) and thus, although only qualitative and not quantitative, the copy number differences seemed to be in agreement with the results obtained during the relative quantification PCR experiments.

3.3.3 Intergenic 6 bp Repeats are Responsible for the Difference in Plasmid Copy Number Between pRAS3.1 and pRAS3.2.

Apart from pRAS3.1 having one more 22-bp iteron than pRAS3.2, it also has 5 × 6-bp CCCCCG repeats whereas pRAS3.2 has only 4 of these repeats. These repeats are located in a putative promoter area between the σ^{70}-like promoter and ribosomal binding sites upstream of the *mobB* (Fig. 3.4). This raised the possibility of whether a difference in the number of these 6-bp repeats possibly has an effect on *mobB* gene expression with a polar effect on downstream *repB* gene expression and hence PCN. A 2.9-kb HindIII-PvuI region containing the 4 × 6-bp repeats of pRAS3.2 was used to replace the corresponding regions on pRAS3.1 and its derivative plasmids to create pairs of plasmids where both plasmids in a pair have the same number of 22-bp iterons but with either 4- or 5 × 6-bp repeats, thus allowing us to isolate the effects of these repeats from the effect of the iterons. In keeping with the naming system previously described where the two last digits in the name of the pRAS3.1-derivatives described the number of 22-bp iterons and 6-bp repeats, respectively, the last digit in the name of each of the newly created pRAS3.1-derivatives was changed in order to reflect the number of 6-bp repeats. Thus after exchange of the 5 × 6-bp repeats of pRAS3.1.35, pRAS3.1, pRAS3.1.55 and pRAS3.1.75 with the 4 × 6-bp repeats of from pRAS3.2, the plasmids were labeled pRAS3.1.34, pRAS3.1.44, pRAS3.1.54 and pRAS3.1.74, respectively. The pRAS3.1 derived plasmid pRAS3.1.34 is essentially similar to pRAS3.2 as it too has 3 × 22-bp iterons and 4 × 6-bp repeats. The only difference remaining between the two plasmids is the point mutations within the *tetAR* genes from pRAS3.1 which is highly unlikely to affect PCN. Therefore, if the 6-bp repeats have an influence on PCN, the copy numbers of WT pRAS3.2 and the derivative plasmid pRAS3.1.34 would be expected to be the same.

The PCN of the pRAS3.1-derivative plasmids with 4 × 6-bp repeats as well as WT pRAS3.2 was determined relative to pRAS3.1.35 by means of Real-Time qPCR as in the previous experiment.

As before, a decrease in PCN was associated with an increase in iteron copy number, however, the respective copy numbers for the derivative plasmids with 4 × 6-bp repeats was in all cases lower than the copy numbers of the plasmids with 5 × 6-bp repeats ($p \leq 0.001$) (Table 3.1). Most importantly, when the number of 6-bp repeats of the 3 × 22-bp iteron derivative plasmid (pRAS3.1.35) was reduced from 5 to 4 (pRAS3.1.34) to be similar to pRAS3.2 (3 × 22-bp, 4 × 6-bp), its relative copy number decreased by ~47% to 31 ± 1. Therefore by beginning with pRAS3.1 and reducing the number of 22-bp iterons from 4 to 3 and the number of 6-bp repeats from 5 to 4, the PCN (31 ± 1) of this altered plasmid was similar to that of pRAS3.2 (30 ± 5). Furthermore, when the 6-bp repeats of WT pRAS3.1, which has a copy number that is 1.5-fold higher than pRAS3.2 despite having 4 × 22-bp iterons as opposed to 3, was reduced from 5 to 4 the PCN of the resulting construct (pRAS3.1.44) decreased by 43% to 23 ± 2 plasmids per chromosome. This is ~25% lower than the copy numbers of the equivalent plasmids pRAS3.2 and pRAS3.1.34 with 3 × 22-bp iterons. Thus the reason for pRAS3.1 having an unexpectedly higher copy number than pRAS3.2 can be attributed to the effects of an additional CCCCCG repeat upstream of its *mobB* gene.

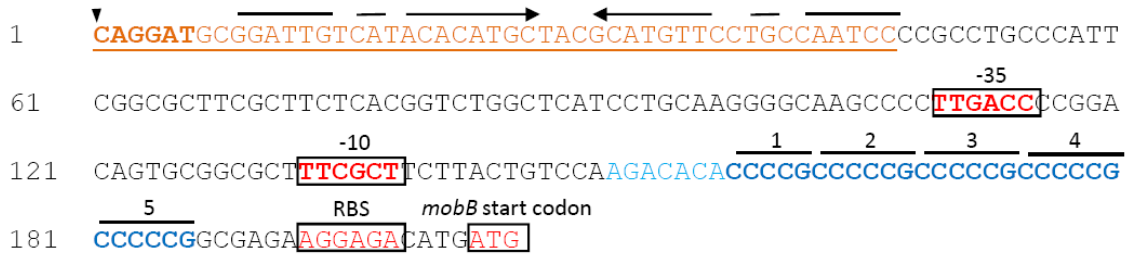


FIG. 3.4. The intergenic sequence between the *oriT* and *mobB* of pRAS3.1, showing the location of the 6-bp repeats. The *oriT* is underlined and indicated in brown. The imperfect inverted repeat of the *oriT* is indicated by broken inverted arrows. The conserved hexameric *nick* site is indicated in bold with a vertical arrow indicating the putative nick position. A putative promoter with a near-consensus -35 region and weak -10 region is shown in bold red within a box, and is separated by a 17-bp spacer. The region containing the 6-bp CCCCCG repeats (labeled 1 to 5) and the AGACACA-tail is indicated in dark and light blue respectively. The first repeat consists of only 5 bp as it lacks a cytosine base. The putative RBS and *mobB* start codon is indicated by red text within a box.

TABLE 3.1. Effect of the number of 6-bp repeats and 22-bp iterons on plasmid copy number

Plasmid construct	No. of 6-bp repeats	No. of 22-bp iterons	Relative plasmid copy number[B]	Calculated plasmid copy number[C]
pRAS3.2	4	3	0.51 ± 0.09	30 ± 5[D]
pRAS3.1	5	4	0.69 ± 0.062	41 ± 4[D]
pRAS3.1.35[A]	5	3	1.0	59
pRAS3.1.34	4	3	0.53 ± 0.002	31 ± 1
pRAS3.1.44	4	4	0.39 ± 0.037	23 ± 2
pRAS3.1.55	5	5	0.48 ± 0.092	28 ± 5
pRAS3.1.54	4	5	0.27 ± 0.012	16 ± 1
pRAS3.1.75	5	7	0.32 ± 0.070	19 ± 4
pRAS3.1.74	4	7	0.25 ± 0.005	15 ± 1
pBR322	-	-	0.30 ± 0.007	18

[A] pRAS3.1.35 served as the reference for the determination of relative copy numbers and standard deviations.

[B] The number of replicates for relative copy number determinations was 4 to 8 and a p value of 0.001 was obtained for each qPCR experiment, except for pRAS3.1.54 where an unexplained p value of 0.081 was obtained.

[C] Plasmid numbers have been calculated to the nearest whole plasmid based on a PCN of 18 plasmids per chromosome for pBR322.

[D] The PCN of pRAS3.1 and pRAS3.2 in this table was extrapolated from their relative copy number to pBR322.

When the 6-bp repeats of pRAS3.1 was reduced from 5 to 4 (pRAS3.1.44), its PCN decreased from 41 ± 4 to 23 ± 2. The resulting copy number was thus ~1.8-fold lower than that of pRAS3.2 and in agreement with the trend that was initially expected based on the number of iterons in their *oriV*s. A reduction in the number of 6-bp repeats in pRAS3.1.55 and pRAS3.1.75 from 5 to 4 resulted in a PCN decrease of ~1.8-fold and ~1.3-fold, respectively. The difference in PCN for the 7 × 22-bp iteron plasmid as a result of having 4- or 5 × 6-bp repeats was thus not

as great as for plasmids with 5 or less iterons. Thus, the effect of the 6-bp repeats on PCN could not overcome the decreasing effect the iterons have on PCN.

3.3.4 The Number of 6-bp Repeats Influences the *mobB-mobA/repB* Operon Transcription Levels.

As the 6-bp repeats are located in a putative promoter area, it is likely that these repeats affected the level of expression of the downstream *mobB* and *mobA/repB* genes. To test this, the expression levels of the *mobB-mobA/repB*, and the divergently transcribed *mobCDE* and *orf3* operon of pRAS3.1 were determined and compared relative to that of pRAS3.2 using Real-Time qPCR. In order to ensure that the possible difference in gene expression is not as a result of differences in PCN, the native replicons of pRAS3.1 and pRAS3.2 were inactivated by truncation of the two respective *repC* genes at the NheI sites. Replication of these two deletion mutants, pR6K.3.1.repC$^{\Delta}$ and pR6K.3.2.repC$^{\Delta}$, was supported by an R6K origin (Ez-Tn5) which was previously cloned into the *tetAR* genes (see Chapter 2 section 2.3.3). Total RNA was extracted from logarithmically growing *E. coli* EC100D cells and converted to cDNA. The relative concentrations of the *repB* and *orf3* cDNA were measured using the pRAS3REPB2 and pRAS3ORF3 primer sets. The relative cDNA concentrations of the kanamycin resistance gene of Ez-Tn5 were monitored using the R6KKANR primer set and was used to normalize the samples. An amplification efficiency of 1.92 was obtained for both the R6KKANR and pRAS3REPB2 primer sets and 2.10 for the pRAS3ORF3 primer set. The correlation coefficient (R^2) values for the R6KKANR, pRAS3REPB2 and pRAS3ORF3 standard curves were 0.999, 0.989 and 0.992 respectively (Fig. 3.5). Although the amplification efficiency of the pRAS3ORF3 primer set is at the maximum theoretical amplification efficiency (*E* = 2.0), no non-specific amplification products, which could have resulted in an abnormally high amplification efficiency, were observed during the melting curve analysis and thus the standard curve was not discarded.

The additional 6 bp repeat in the pRAS3.1 equivalent (pR6K.3.1.repC$^{\Delta}$, 5 × 6-bp repeats) caused a small but significant 2-fold (1.989 ± 0.897; *n* = 12 and p = 0.048) increase in expression of *repB* compared to the pRAS3.2 equivalent (pR6K.3.2.repC$^{\Delta}$, 4 × 6-bp repeats). In contrast, expression in the opposite direction (*orf3*) was not different between the two plasmids (1.097

± 0.54; *n* = 4 and p = 0.8745). These results suggest that the reason for the increase in PCN of pRAS3.1 compared to pRAS3.2 was that the additional 6-bp repeat caused an increase in expression of the *mobB-mobA/repB* operon, and that increased expression of *repB* was the actual cause.

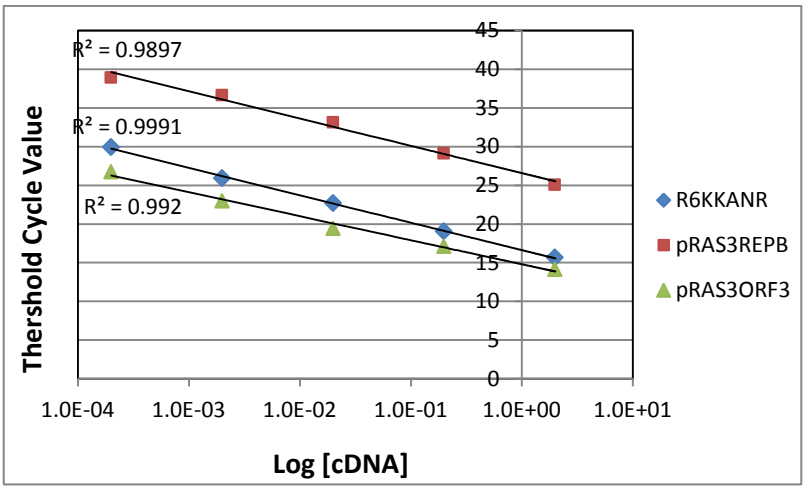


FIG. 2.5. Standard curves for the R6KKANR, pRAS3REPB2 and pRAS3ORF3 primer sets. The correlation value (R^2) for each standard curve is indicated next to the trend line. The actual amplification values were determined using the REST – 384 © software tool. As the cDNA concentration used to construct the standard curves were not known the concentration values are arbitrary.

3.3.5 Effect of Increased *repBAC* Expression on Plasmid Copy Number

To confirm that the increased *repB* expression resulted in a concomitant increase in PCN, the pBAD28-repB construct (previously described in Chapter 2 section 2.3.11) was used to supply additional RepB proteins during copy number determination experiments. The arabinose-inducible *repB* construct was transformed into *E. coli* (pRAS3.1.34) cells and the copy number of the pRAS3-like plasmid was compared by means of Real-Time qPCR relative to pRAS3.1.34 with the pBAD28 vector *in trans*. The presence of additional RepB increased the PCN of pRAS3.1.34 approximately 2.2 ± 0.53-fold (*n* = 4, p ≤ 0.001) relative to the vector control. This is approximately equal to the 1.96-fold difference in PCN between plasmids differing only in their number of 6-bp repeats, such as pRAS3.2 (3 × 22-bp iterons and 4 × 6-bp repeats) and

pRAS3.1.35 (3 × 22-bp iterons and 5 × 6-bp repeats), as shown in Table 3.1. Therefore, the additional RepB provided by pBAD28-repB *in trans* raised the copy number of a plasmid which contained 4 CCCCCG repeats to approximately the same as that of a plasmid with 5 of the repeats.

To examine whether the availability of RepA or RepC proteins is also a rate-limiting factor in plasmid replication, as demonstrated for RepB, the pBAD28 constructs containing the *repAC* or *repC* genes (previously described in Chapter 2 section 2.3.11) were placed into *E.coli* DH5α cells containing pRAS3.1.35. The reason for using pRAS3.1.35 (as opposed to pRAS3.1.34 that was used in the above experiment) was that it already had higher levels of *repB* expression and it was possible that the *repA* and *repC* gene products were limiting. Provision of additional RepC by placing pBAD28-repC *in trans* of pRAS3.1.35 did not result in a change in PCN (1.02 ± 0.22; n = 4; p = 0.874). Additional RepA and RepC, provided by pBAD28-repAC, however, resulted in an approximately 30% reduction (0.67 ± 0.13; n = 4; p = 0.008) in the PCN of pRAS3.1.35. The availability of neither the RepC nor the RepA was thus a limiting factor for plasmid replication, and in fact, an excess supply of both the RepA and RepC proteins had a negative influence on PCN.

3.4 DISCUSSION

To investigate how the number of 22-bp iterons within the *oriV* of the pRAS3 plasmids influences the biology of these plasmids we needed to create derivative pRAS3-like plasmids that differ only with regards to the number of iterons within their *oriV*s were created. This was achieved by digesting pRAS3.1 plasmid DNA with BstEII, a restriction endonuclease that cuts specifically at the end of each iteron, and then allowing the large 11.8-kb and the small 22-bp iteron fragments to anneal randomly in a concentrated self-ligation reaction. In such a way, a range of pRAS3.1 derivatives were obtained that contained 3 (pRAS3.1.35), 4 (pRAS3.1.34, identical to WT pRAS3.1), 5 (pRAS3.1.55) and 7 × 22-bp iterons (pRAS3.1.75). No plasmids with fewer than 3 × 22-bp iterons were obtained from the self-ligation reaction. In a parallel experiment we attempted to create a pRAS3.1 derivative containing only 1 × 22-bp iteron by separating the small 22-bp iterons from the 11,875-bp fragment prior to self-ligation. No plasmid-containing *E. coli* DH5α cells were obtained when the ligation mixture was

transformed into competent cells. The inability to generate pRAS3.1 derivatives containing fewer than 3 × 22-bp iterons within the *oriV* thus suggests that 1 or 2 × 22-bp iterons are insufficient for initiation of replication. These results are in agreement with the experiments of Lin and Meyer (1986) wherein they showed that the replication proteins of R1162 are not able to sustain the replication of a cloned *oriV* containing fewer than 3 × 22-bp iterons. It has been demonstrated the degree of bending at the iterons due to the binding of RepC proteins is dependent on the number of iterons (Miao *et al.*, 1995). A possible reason for the inability of plasmids with only 1 or 2 × 22-bp iterons to replicate could be thus be due to insufficient bending in the *oriV*, thus resulting in a failure of the dsDNA to open up at the A + T-rich region which is needed for entry of the RepA helicase and RepB primase.

Relative plasmid copy number experiments demonstrated that the PCN of pRAS3.1 increased by ~31% when the number of iterons in its *oriV* was decreased from 4 to 3, and decreased by ~30 and ~53% when the number of iterons within the *oriV*s were increased from 4 to 5 and 7 copies respectively. The copy numbers of the two WT plasmids, pRAS3.1 and pRAS3.2, however, did not follow this trend. The estimated PCN for pRAS3.1, which has 4 × 22-bp iterons, was ~50% higher (PCN ~45) than the PCN of pRAS3.2 (PCN ~30), which has 3 × 22-bp iterons and by convention would be expected to have the higher copy number. As pRAS3.1 has 5 × 6-bp CCCCCG repeats upstream of its *mobB* gene while pRAS3.2 has only 4 of these repeats, I wished to determine whether the additional 6-bp repeat of pRAS3.1 was responsible for it having a higher PCN than pRAS3.2. When the region containing the 5 × 6-bp repeats of the pRAS3.1 derivatives was exchanged with a corresponding 2.9-kb region from pRAS3.2 containing 4 × 6-bp repeats, the copy numbers of the resulting plasmids decreased proportionately relative to their 5 × 22-bp counterparts. The PCN of the pRAS3.1-equivalent (pRAS3.1.44) with four, instead of five, of the 6-bp repeats was now ~24% lower than that of pRAS3.2 and in agreement with the trend that was initially predicted. By exchanging the 5 × 6-bp repeats of the three-iteron pRAS3.1 derivative (pRAS3.1.35) for 4 × 6-bp repeats to give the pRAS3.2 equivalent plasmid pRAS3.1.34, its copy number was reduced from ~59 to ~31 plasmids per chromosome, which is similar to the ~30 plasmids per chromosome that was estimated for pRAS3.2. These experiments thus confirmed that the reason for pRAS3.1 having a higher PCN than pRAS3.2 is due to the presence of an extra 6-bp repeat.

The region between the *oriT* and *mobB* gene of IncQ plasmids such as RSF1010 and pTC-14 is known to contain a cluster of divergently transcribed promoters which are responsible for expression of the *mobB, mobA/repB* genes in one direction and the *mobC* gene (including the *mobDE* genes for the IncQ-2 plasmid) in the opposite direction (Derbyshire and Willetts, 1987; Gardner and Rawlings, 2004; Scholz *et al.*, 1989). Northern blot analysis of the mRNA transcripts from, pTC-F14, a different IncQ-2 plasmid, revealed that a polycistronic mRNA transcript is produced from the promoter region between the *oriT* and *mobB* gene which includes not only the *mobB* and *mobA/repB* genes, but also the *repAC* genes (Gardner and Rawlings, 2004). Also, when the promoter region of the *pasABC* system (located between the *repB* and *repA*) of pTF-FC2 was carefully deleted, the mutant plasmid was still maintained at a copy number similar to that of the WT plasmid (Matcher and Rawlings, 2009). It thus seems that the promoter region upstream of the *mobB* gene plays an important role in plasmid replication and maintenance. To determine whether the number of CCCCCG repeats in the *mobB* promoter region of the pRAS3 plasmids has an influence on promoter activity, the expression levels of the divergently transcribed *repB* and *orf3* genes from a pRAS3.1 *mobB* promoter with 5 × 6-bp repeats were compared relative to that of a pRAS3.2 *mobB* promoter with 4 × 6-bp repeats. It was found that the transcription level of the *mobB-mobA/repB* operon was 2-fold higher for the promoter with 5 × 6-bp repeats, and remained unchanged in the opposite direction. Increased transcription of the *repB* of pRAS3.1 compared to pRAS3.2 thus probably resulted in higher concentrations of the RepB protein and would be the underlying reason for the iteron-independent copy number difference between these two plasmids.

The presence of the additional 6-bp repeat could mediate the increase in transcription either by altering the ability of the promoter, within which it is located, to initiate transcription of the *mobB-mobA/repB* operon or by decreasing the efficiency at which the promoter is autorepressed. The 6-bp repeats are located 19-bp downstream of the -10 position of the σ^{70}-like promoter region and 6-bp upstream of the putative RBS. A difference in the number of 6-bp repeats would therefore alter the spacing between these elements (Figure 3.4) and the side of the DNA helix on which the -10 region and RBS elements occur. The optimal spacing between the -10 and RBS for *E. coli* δ^{70} promoters is ~30 – 40-bp (Mendoza-Vargas *et al.,* 2009; Shultzaberger *et al.,* 2007). As an extra 6-bp repeat increases the spacing between these two

putative elements from 48 to 54-bp (which is further away from the 35-bp optimal spacing) and results in increased transcription (rather than a decrease in transcription), it is unlikely that the difference in spacing between the -10 and RBS loci is responsible for the difference in promoter activity. In addition to the presence of 4 or 5 of these 6-bp repeats and the AGACACA-tail between the *oriT* and *mob*, these 6-bp repeats are also located within the *mobC* and *mobB* genes as well as in the *oriV* (L'Abee-Lund and Sørum, 2002). There are, however, only three of the repeats in each of the alternate locations and the number of cytosine residues in each of these repeat units varies between 4 and 5 to give unit lengths of either 5 or 6-bp. Nonetheless, the repeat units seem to be conserved in multiple *loci.* No function has yet been assigned to these repeat units, however, when located in the *mobB* promoter area, they were found to influence promoter activity. It is thus likely that the 6-bp repeats are protein binding sites, and when the cognate protein is bound to the 6-bp repeats within the *mobB* promoter region it could possibly act as a negative regulator. Furthermore, as the level of transcription was influenced by the number of 6-bp repeats present in the promoter area, the ability of the negative regulator to bind at the correct position for maximal repression is probably decreased when more than four repeats are present.

The negative regulator of the promoter upstream of the *mobB* of pTC-F14 was identified by measuring transcription levels from the promoter (fused to a β-galactosidase reporter system) while the mobilization and replication genes were expressed from a vector *in trans* (Gardner and Rawlings, 2004). When a pGL10-based clone containing the *mobE* to *repB* genes was placed *in trans* of the fused promoter, the transcription levels of the reporter system was 77% lower than that of the vector-only control. When the *repB* domain of the *mobA/repB* fusion gene in the pGL10-based clone was deleted, the expression levels of the reporter system increased from 33% to 92% relative to the vector-only control. In an additional experiment the minimal replicon of pTC-F14, consisting of the *repBAC* replication genes and the *oriV* fused to an antibiotic resistance marker, was placed *in trans* of the reporter system, and it was found that transcription from the promoter was reduced by 57% relative to the vector-only control. These experiments thus implicated the RepB primase as the negative regulator of the promoter upstream of the *mobB* gene of pTC-F14. Considering that pRAS3 and pTC-F14 are closely related (see Chapter 2, section 2.3.1), it is possible that the RepB of pRAS3, like that of pTC-F14, is the autoregulator of the promoter containing the 6-bp repeats. Whether the 6-bp

repeats and the AGACACA-tail serve as binding sites of the RepB for autoregulatory purposes, however, remains to be confirmed by DNA foot printing assays.

The genes of the three replication proteins, RepB, RepA and RepC, were previously cloned into the pBAD28 expression vector to give pBAD28-repB, pBAD28-repAC and pBAD28-repC, and functional expression of the genes was validated by their ability to compliment *repB* and *repC*-deletion mutants (see Chapter 2 section 2.3.11). By measuring the copy number of a plasmid containing 4 × 6-bp repeats (pRAS3.1.34) while being provided with additional RepB from pBAD28-repB *in trans* we found that the PCN of the plasmid increased ~2.2-fold relative to the copy number of a vector-only control (pRAS3.1.34; pBAD28). This correlates with the increase in PCN that was suggested to be as a result of increased transcription of the *repB* gene from a plasmid with 5 × 6-bp repeats (pRAS3.1.35) compared to a plasmid with 4 × 6-bp repeats (pRAS3.1.34). Provision of additional RepA and RepC by placing pBAD28-repAC or pBAD28-repC *in trans* of pRAS3.1.5 did not have an increasing effect on PCN. This confirmed that the presence of an additional 6-bp repeat increased the transcription levels of the *repB* gene, and that the resulting increase in RepB concentrations was responsible for pRAS3.1 having a 1.5-fold higher PCN than pRAS3.2.

The RepC protein of RSF1010/R1162 has been demonstrated to be involved in copy number control by two independent studies. Expression of the *repAC* genes from an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible *taq* promoter (cloned on a pACYC184-based vector) *in trans* of an R1162 derivative resulted in a ~3-fold increase in PCN compared to the uninduced strain (Kim and Meyer, 1985). In a different study the RSF1010 *repAC* genes were cloned behind a *taq* promoter on a pKT101-based vector and expressed *in trans* of a RSF1010 derivative. In this instance, induction of the *taq* promoter by addition of IPTG resulted in a ~6-fold increase in the PCN of the RSF1010 derivative (Haring *et al*., 1985). The amount by which the PCN increased for these two near identical plasmids differs by 2-fold. This could be due to a difference in the level of induction of the *taq* promoters on the pACYC184-based vector (p15A replicon) and the pKT101-based vector (ColD replicon), a difference in growth media (TYE broth *vs* M9 minimal media) or a difference in the method by which the relative copy numbers were determined. The change in copy number of RSF1010 was estimated by comparing the difference in band intensity on an agarose gel between two plasmids extracted from approximately equal amounts of induced and uninduced cells (Kim and Meyer, 1985).

For R1162 the change in copy number was determined by comparing the ratio of [^{3}H]-labeled plasmid and chromosomal DNA after hybridization on filters containing probes for the plasmid and chromosomal DNA respectively (Haring *et al*., 1985). The difference in the ratio of plasmid to chromosome was then compared between induced and uninduced samples in order to obtain the difference in PCN between the two samples. Overexpression of only the respective *repA* genes *in trans* of either RSF1010 or R1162 did not result in a change in PCN (Haring *et al*., 1985; Kim and Meyer, 1985). Thus, despite the 2-fold difference in the two experiments, both experiments demonstrated that the PCN of the IncQ-1 plasmids RSF1010/R1162 was dependent upon the amount of available RepC initiator as in both instances.

This is in contrast to the two IncQ-2 plasmids. No change in the PCN of pRAS3.1.35 was evident when the pRAS3 *repC* gene was overexpressed *in trans*. When the *repAC* genes, on the other hand, were provided in excess a ~33% decrease in the copy number of pRAS3.1.35 was observed. The *repB* gene of the pRAS3.1.35 plasmid was already expressed at a higher level due to it having 5 × 6-bp repeats, and thus could not have been limiting. The relative copy number of a pTF-FC2 *repAC* deletion derivative also did not increase beyond that of the WT pTF-FC2 copy number when its cognate *repAC* genes were provided in excess (Matcher and Rawlings, 2009). The basal level transcription (in M63 minimal media) of the *repAC* genes from the P_{BAD} promoter in pBAD28 was sufficient to maintain the pTF-FC2 *repAC* deletion derivative at a relative copy number (measured by Real-Time qPCR) that was similar to the copy number of the WT pTF-FC2 control plasmid. Induction of the P_{BAD} promoter by addition of 2.0×10^{-6} and 2.0×10^{-5}% L-arabinose resulted in small non-significant increases in PCN. Further induction with 2.0×10^{-4}% L-arabinose, however, resulted in a decrease of 42% of the PCN of the pTF-FC2 deletion derivative relative to the WT. Thus, that no increase in PCN was observed when the RepA and RepC proteins of both pRAS3 and pTF-FC2 (work carried out by different researchers) were present in excess suggests that the availability of the RepA and RepC proteins for replication at the physiological concentrations was not rate-limiting for these two IncQ-2 plasmids. Such naturally high concentrations of the pRAS3 and pTF-FC2 RepA and RepC proteins may have to do with the strongly autoregulated TA system present immediately upstream of the *repAC* genes in IncQ-2 plasmids but not IncQ-1 plasmids. As RepC dimerization has previously been observed for an IncQ-like RepC protein the possibility of

RepC dimerization to dampen the effect of high RepC concentrations is discussed together with suggestions on how to experimentally verify such a phenomenon in Chapter 5.

The copy numbers of both pRAS3.1.35 and the pTF-FC2 *repAC* deletion derivative decreased by 33 and 42% (Matcher and Rawlings, 2009), respectively, when their respective *repA* and *repC* genes were transcribed from an alternate promoter *in trans*. As this phenomenon did not occur when only the RepC protein of pRAS3 was overexpressed in a host with a coresident pRAS3.1.35 plasmid, it was probably caused by the increased concentration of RepA. Attempts to clone and express the pRAS3 *repAC* genes in vectors other than pBAD28 were largely unsuccessful. Furthermore, the growth rate of cells containing the pBAD28-repAC construct was noticeably slower than that of cells containing the pBAD28 vector only (data not shown). Overexpression of the pTF-FC2 *repAC* genes from behind the P_{BAD} promoter had a similar effect on growth rate of the host cells (G. Matcher, pers. comm.). The RepA protein of IncQ plasmids is a DNA helicase (Scherzinger *et al*., 1997), however, it has not been established whether its activity is highly plasmid specific. The RepA of pTF-FC2 is able to substitute for the closely related RepA of pTC-F14 during replication of pTC-F14 (Gardner and Rawlings, 2004) and thus it is possible that the RepA specificity might be relaxed to some extent. Overexpression of a helicase with relaxed specificity could therefore interfere with chromosome replication and thus would have a detrimental effect on the growth rate of the host and possibly plasmid replication.

The copy number of pBR322 was previously determined in *E. coli* DH5α by means of both absolute and relative Real-Time PCR quantification methods in a LightCycler instrument (Lee *et al.,* 2006a). By including pBR322 as a control in the relative copy number experiments performed on the two WT pRAS3 plasmids and the pRAS3.1 derivatives, it was possible to validate the copy numbers that were obtained for pRAS3.1 and pRAS3.2 during the absolute quantification experiments. The reason for these two plasmids having a PCN that is approximately 1.8- (pRAS3.2) and 2.6-fold (pRAS3.1) higher than that of other IncQ-like plasmids (10 – 16 plasmids per chromosome), however, remains speculative. A likely explanation could be that the RepC protein of the pRAS3 plasmids is very different from that of other IncQ plasmids (see Chapter 2 section 2.3.3) and could therefore have altered copy number regulatory properties.

It was shown both quantitatively using Real-Time qPCR and qualitatively by means of densitometric analysis that an increase in the number of iterons in the *oriV* of plasmids derived from pRAS3.1 resulted in a decrease in PCN. By looking at differences in the expression levels of the *repB* genes of pRAS3.1 and pRAS3.2, as well as the effect of increased *repB* transcription on PCN, it was demonstrated that increased transcription of the *repB* gene was able to bring about an increase PCN. Consequently, the reason for pRAS3.1 having a higher PCN than pRAS3.2, despite having one more iteron (which would otherwise have decreased the PCN), was as a result of increased transcription of the *repB* gene due to pRAS3.1 having one more 6-bp repeat in the *mobB* promoter region than pRAS3.2. How the differences in the number of iterons and 6-bp repeats in the backbones of pRAS3.1 and pRAS3.2 affects overall inter- and intracellular plasmid fitness was investigated and the results are reported in Chapter 4.

Chapter 4

Competitive Fitness of pRAS3.1 and pRAS3.2

Table of Contents

4.1 INTRODUCTION

The IncQ-like plasmid family has been subdivided into IncQ-1 and IncQ-2 subgroups on account of their mobilization systems which are of two types (Rawlings and Tietze, 2001). The IncQ-1 subgroup has a mobilization system comprising of three genes, while that of the IncQ-2 plasmids consists of five genes in addition to the *oriT* respectively. The *repBAC* genes, encoding the RepB primase, RepA helicase and RepC initiator proteins, respectively, of IncQ-like the replication systems are however related and demonstrates a common evolutionary lineage for both the IncQ-1 and IncQ-2 groups (Rawlings and Tietze, 2001). Nonetheless, the replication systems of the IncQ plasmids have diverged sufficiently for the plasmids to have been divided into even smaller IncQ-1- and IncQ-2α, β and γ subgroups based on their incompatibility (same subgroup) or compatibility (different subgroup) with other IncQ-like plasmids (see Chapter 1 section 1.2.5.1), and thus demonstrates that the evolution of new incompatibility groups from existing groups is ever ongoing.

The evolution of plasmid replicons into new incompatibility groups occurs as a result of the accumulation of mutations within the replicon and is driven by the need for plasmids to ensure their own replication (Jiang *et al*., 2002; Thomas, 2004). Plasmid replicons are, however, sensitive to mutations as it often leads to over replication or a decrease in replication efficiency and stability (Krakauer and Plotkin, 2002). Thus the evolution of replication systems to increase its competitive fitness faces selection at two levels (Paulsson, 2002). Within the host closely related plasmids need to compete for the available plasmid component of the replication machinery. If one of the plasmids has a replication advantage it will tend to displace the competing plasmid from the host. This form of intracellular competition bears analogy to plasmid incompatibility (Novick, 1987) and illustrates, with respect to competitive fitness, the drive for the formation of new incompatibility groups so that the plasmids do not compete with other closely related plasmids. Plasmids also have to compete at the host population or intercellular level to persist within an environment. Plasmid carriage imposes an additional metabolic load on the host (Andersson and Levin, 1999; Enne *et al.,* 2004; Lenski and Bouma, 1987; Lenski *et al*., 1994), and therefore a population of host cells containing a plasmid that imposes a lower metabolic load might be more competitive than an isogenic host population that carries a similar plasmid but which imposes a higher metabolic load due to having a higher replication frequency or the accessory genes it carries (Becker and Meyer,

1997; Lenski, 1992). If the PCN is too low, however, then the plasmid risks segregational loss (Allen and Blaschek, 1990; Becker and Meyer, 1997; Nordstrom and Austin, 1989). It is thus the interplay between these factors which keep the plasmid in the population or see it lost from the population.

The competitive fitness of two bacterial populations, plasmid-containing or otherwise, can be inferred by monitoring the relative change in abundances of two competing populations within an environment (Lenski, 1992). The log ratio of the number of cells of each population is plotted against time, and the gradient of the resulting regression, referred to as the selection coefficient, is indicative of the rate at which one population is displaced by another due to differences in growth rate. By dividing the selection coefficient by the average number of generations per day the relative difference in fitness can be expressed as a per generation fitness impact (Lenski *et al*., 1994; Reynolds, 2000). The value assigned to the fitness of a population however is relative to the fitness of the competing population under the given experimental circumstances, and thus there are several assumptions, and hence limitations, associated with measuring competitive fitness in this manner (Lenski, 1992). The first assumption is that the physiological states of the competing populations are comparable. If both populations are not equally acclimated to the environmental conditions then the result will be biased. It is also assumed that the change in the frequency of occurrence of individuals in each population is due to differences in growth rate. Plasmid loss or conjugative transfer between plasmid-containing and plasmid-free competing populations would subtract from or add to the plasmid-containing population cell numbers, thereby affecting the ratio of the competing populations and would thus result in an artificially high or low metabolic burden, respectively. Finally it is assumed that the selection coefficient is constant. Periodical changes in the growth conditions or genetic drift may increase the selection for one of the populations and thus also change the outcome of the experiments. Nonetheless, such relative fitness assays combined with long-term evolutionary experiments have been successfully applied to study the evolution of the plasmid-bacterial relationships (Bouma and Lenski, 1988; Dahlberg and Chao, 2003; Dionisio *et al.,* 2005; Lenski *et al*., 1994; Modi and Adams, 1991)

Plasmids pRAS3.1 and pRAS3.2 are two almost identical, naturally occurring IncQ-like plasmids. Apart from a number of point mutations within their respective tetracycline resistance genes, they differ only with regards to the number of 22-bp iterons and 6-bp repeats within their

respective *oriV* and *mobB* promoter regions. Based on the small number of differences between these two plasmids, they probably evolved from a common ancestor relatively recently. As both plasmids have however managed to persist in the environment (L'Abee-Lund and Sørum, 2002), it raised the question of whether the number of iterons and 6-bp repeats, which have compensatory effects on PCN (see Chapter 3 sections 3.3.2 and 3.3.3), has contributed to the overall competitive fitness of these two plasmids. By monitoring the segregation patterns of coresident plasmids with different iteron and or 6-bp repeat copy numbers it was found that pRAS3.1 has an intracellular competitive advantage over pRAS3.2 due to it having 4 × 22-bp iterons as opposed to 3. Relative fitness assays to determine the intercellular competitive fitness of these two plasmids showed that pRAS3.1 has a metabolic load that is ~2.8% higher than that of pRAS3.2 due to it having a copy number that was estimated to be ~1.5-fold higher than that of pRAS3.2. By similarly investigating the competitive fitness of pRAS3 derivative plasmids it was demonstrated that the evolution of the pRAS3 plasmids are guided by the antiredundant (increased sensitivity to mutation) properties of the pRAS3 replicon.

4.2 MATERIAL AND METHODS

4.2.1 Bacterial Strains, Plasmids, Media and Growth Conditions. For information on the bacterial strains, plasmids, media and additives as well as the growth conditions used refer to Chapter 2 section 2.2.1. Plasmid preparation, restriction endonuclease digestions, gel electrophoresis, and cloning were carried out using standard methods (Ausubel *et al.,* 1993; Sambrook *et al.,* 1989).

4.2.2 Displacement Assays. Plasmid-containing *E. coli* DH5α cultures were made chemically competent and transformed with a second plasmid and spread onto LA plates containing antibiotic selection for both plasmids. Single colonies were picked into 5 ml LB media containing the respective antibiotics for both plasmids and incubated overnight at 37°C. Thereafter, the cultures were serially diluted in PBS and a range of dilutions (one of which contained ~1.0×10^3 cells) were re-inoculated into 5 ml fresh non-selective LB media and grown for 24 h so as to obtain ~20 generations. The PBS-diluted cultures from the selective growth cycle, as well as the PBS-diluted cultures from the non-selective growth cycle were

spread onto non-selective LA plates and incubated at 37°C. Fifty colonies originating from the selective growth cycle, and 50 colonies originating from the non-selective growth cycle that was inoculated with ~1.0×10^3 cells (determined from the plate counts) were replica-plated from the non-selective plates onto sets of LA plates containing selection for either plasmid individually or both plasmids together, as well as plates without antibiotic selection. The percentage plasmid retention for each sample was determined by scoring the number of colonies that grew on the respective selective LA plates compared to the non-selective plates. As a control, cells containing the individual plasmids were grown and replica-plated similarly in order to monitor plasmid stability.

4.2.3 Relative Plasmid Copy Number Determinations Using Densitometric Analysis. Plasmid copy numbers were compared by means of densitometric analysis (Park *et al.,* 2001) as described in Chapter 3 section 3.2.4. The only differences to the previous assay were that the plasmid DNA was extracted from the mixed cultures using a PureYield™ Plasmid Purification Kit (Promega) as per manufacturer's protocol, and the purified DNA in each sample was eluted in 80 μl ddd.H_2O and digested in a 100 μl volume using SalI restriction endonuclease. The 100 μl restriction digest reaction was split into 3 prior to loading on a 0.8% agarose/TBE gel for visualization.

4.2.4 Plasmid Stability Assay. Single *E. coli* DH5α colonies transformed with the respective plasmids were inoculated into 5 ml LB or M9 (Appendix A) media containing the appropriate antibiotic selection and incubated at 30 or 37°C as required. The cultures were then diluted 1.0×10^6 into fresh media without antibiotic selection and incubated as before. This was repeated daily for 5 days. Each day serial dilutions of the cultures were also spread onto non-selective LA agar plates and incubated at 37°C. The percentage plasmid retention was determined by replica plating 50 colonies from each of the plates onto pRAS3-selective plates as well as non-selective plates. When a coresident pACYC177 vector was included the vector was maintained under constant antibiotic selection for the duration of the assay.

4.2.5 Relative Fitness Assays. Plasmids were transformed into chemically competent *E. coli* JM109, plated onto Luria agar and incubated O/N at 37 °C. Single plasmid-containing (P+) as well as plasmid-free (P-) colonies were inoculated into 10 ml DM25 minimal media (Appendix A) and grown overnight at 37°C while shaking. The cultures were adapted to the minimal

media prior to the assay by diluting the cultures 1×10^{-2} into fresh media every day for 3 days without antibiotic selection. On the third day the P+ and P- cultures were split and inoculated into 6 Erlenmeyer flasks containing fresh media and incubated overnight at 37°C while shaking. Each of the 6 P+ cultures were then mixed with one of the 6 P- cultures in a 60:50 P+:P- volumetric ratio, diluted 1×10^{-2} into fresh media and incubated again at 37°C. The cultures were mixed in a 60:50 P+:P- ratio as the P+ culture was expected to be rapidly outcompeted and were thus inoculated at a slightly higher concentration to increase the duration of the assay. The cultures were grown in serial batch cultures for a total of 6 days by diluting the cultures 1×10^{-2} into fresh media every 24 hours. Each day, starting from when the cultures were mixed, serial dilutions were spread in triplicate onto selective antibiotic-containing and non-selective LA plates and incubated overnight at 37°C for colony counts. The number of P- colonies was obtained by subtracting the number of P+ colonies obtained on the selective media from the total number of colonies obtained on the non-selective media. The log ratio of P+ to P- cells was plotted over time and the slope of the linear regression, referred to as the selection rate constant, was used to calculate the relative fitness as described by R. E. Lenski (1992). The results were expressed as a percentage fitness impact per generation (Appendix A), by dividing the selection coefficient by the average number of generations per day.

Monocultures of the P+ samples were maintained and sampled in a similar manner as a control for plasmid stability. Fifty colonies from each of the non-selective plates were also replicated onto selective and non-selective Luria agar plates in order to verify as indicated by the plate counts. The relative fitness data was analyzed by means of a single sample *t*-test with a constant of 100% (selection rate constant of 0) for each of the samples, and the significance for comparisons between different plasmid samples was determined by means of Fischer LSD test using the Statistica version 9 (StatSoft Inc., Oklahoma USA) software.

4.2.6 Determination of Mobilization Frequency. *E. coli* S17.1 donor cultures containing the respective plasmids and *E. coli* CSH56 recipient cultures were inoculated into 5 ml LB media containing the appropriate antibiotics and incubated overnight at 37°C. Two milliliters of the cells were harvested by centrifugation at 8,000 rpm for 2 min and washed by three successive rounds of resuspension in 2 ml PBS followed by centrifugation before finally being resuspended in 1 ml PBS. The absorbance of each sample was determined at OD_{600} and standardized to an absorbance value of 1. The cultures were then mixed in a donor-to-

recipient ratio of 1:100 and 100 μl of each mixture was spotted onto a LA plate and incubated at 37°C for 30 minutes. The agar plugs were excised, resuspended in 10 ml PBS and vigorously shaken for 30 min after which cells from 8 ml of each sample were collected by centrifugation and resuspended in 1 ml PBS. Serial dilutions were spread onto donor- and transconjugant-selective LA plates and the number of transconjugants per donor calculated after incubation at 37°C. Donor and recipient cultures that were not mated were also spread onto the respective antibiotic-containing plates as a control.

4.3 RESULTS

4.3.1 The Ability of pRAS3.1, pRAS3.2 and Derivatives to Displace Each Other Within a Host Cell

It has been demonstrated that IncQ-like plasmids with similar or identical replicons segregate in a symmetrical pattern when coresident in the same host (Smalla *et al*., 2000). Given that pRAS3.1 has 1 more iteron than pRAS3.2, we wished to determine how this additional iteron would influence the segregation pattern when these two plasmids are coresident. For the purpose of being able to co-transform and select for both pRAS3 plasmids in a single cell, the kanamycin resistance gene from pSKm2 was cloned into the unique BamHI-EcoRI sites of either pRAS3 plasmid, thereby also inactivating the *tetAR* tetracycline resistance genes, to produce pRAS3.1.Km and pRAS3.2.Km.

The co-transformed *E. coli* DH5α cultures, containing combinations of either pRAS3.1.Km and pRAS3.2, or pRAS3.2.Km and pRAS3.1, were grown for 12 hours (~20 generations) without antibiotic selection before the plasmid segregation patterns were determined by replica-plating 50 single colonies onto plasmid-specific selective plates. Plasmid pRAS3.1 (or pRAS3.1.Km) displaced pRAS3.2 (or pRAS3.2.Km) in both reciprocal experiments as ~98% of the tested colonies contained only pRAS3.1 while the remaining ~2% in all cases contained both plasmids (Table 4.1). No colonies consisting of cells containing only pRAS3.2 (pRAS3.2.Km) were observed. No loss of either plasmid was observed when hosts containing only a single plasmid were grown in parallel and treated in a similar manner. The 4 × 22-bp iteron plasmid pRAS3.1 thus rapidly displaced the 3× 22-bp iteron pRAS3.2 plasmid

(irrespective of the type of antibiotic resistance gene on the plasmid), while displacement of pRAS3.1 by pRAS3.2 was never observed in any of the colonies tested. When the band intensities of pRAS3.1 and pRAS3.2.Km (extracted from *E. coli* DH5α containing these two plasmids as coresident plasmids under antibiotic selection) was compared on an agarose gel it was observed that the copy number of pRAS3.2 was suppressed to such an extent that the pRAS3.2 band was barely visible (Fig. 4.1). Displacement of pRAS3.2 by pRAS3.1 thus suggests that pRAS3.2 is unable to replicate effectively when it is coresident with pRAS3.1.

TABLE 4.1. Segregational patterns of coresident pRAS3 plasmids and derivatives in an *E. coli* host

	pRAS3.2.km	pRAS3.1.35.Km	pRAS3.1.Km	Percentage colonies with resistance to respective antibiotics			Direction and strength of segregation bias as a result of iteron and or 6-bp repeat copy number, as well as the influence of the antibiotic resistance genes
				Tet only	Km only	Tet and Km	
pRAS3.2			●[A]	0 ± 0	98 ± 3	2 ± 3	3 iterons; 4 × 6-bp repeats <<< 4 iterons; 5 × 6-bp repeats
pRAS3.1	●			98 ± 0	0 ± 0	2 ± 0	4 iterons; 5 × 6-bp repeats >>> 3 iterons; 4 × 6-bp Repeats
		●		79 ± 4	0 ± 0	21 ± 4	4 iterons >>> 3 iterons
pRAS3.1.35			●	0 ± 0	87 ± 10	12 ± 8	3 iterons <<< 4 iterons

[A] Dots indicate the plasmid combination used in each experiment. Each plasmid was completely stable on its own for the duration of the assay in the absence of antibiotic selection.

As the two natural pRAS3 plasmids differ both in the number of 6-bp repeats present within their *mob* promoter region as well as the number of 22-bp *oriV*-associated iterons, I wished to isolate the effect of the number of 22-bp iterons alone on plasmid displacement. To do so, the assay was repeated by competing the pRAS3.1-derivative pRAS3.1.35, which has 3 × 22-bp iterons (same as pRAS3.2) and 5 × 6-bp repeats (same as pRAS3.1), against pRAS3.1.Km. After ~20 generations of growth without antibiotic selection in *E. coli* DH5α, the plasmid with 4 × 22-bp iterons (pRAS3.1.Km, PCN ~41) displaced the plasmid with 3 × 22-bp iterons (pRAS3.1.35, PCN ~59) in 87 ± 10% of the cells while 12 ± 8% retained both plasmids. A reciprocal

experiment wherein a version of pRAS3.1.35 that bears resistance to kanamycin was competed against pRAS3.1 gave largely the same result with the 4 × 22-bp iteron plasmid (pRAS3.1) displacing the 3 × 22-bp iteron plasmid (pRAS3.1.35.Km) in 79 ± 4% of the host cells while 21 ± 4% retained both plasmids. In both experiments no host cells containing only the 3 × 22-bp iteron plasmid were detected. The natural 3 × 22-bp iteron plasmid pRAS3.2 on its own in a cell was shown to have a copy number of ~30 plasmids per chromosome while the pRAS3.1 derivative plasmid with 3 × 22-bp iterons (pRAS3.1.35) was shown to have a copy number of ~59 plasmids per chromosome (see Chapter 3 section 3.3.3). Thus in both instances the plasmid with 4 × 22-bp iterons (pRAS3.1) displaced a plasmid with 3 × 22-bp iterons irrespective of whether its copy number was lower or higher than that of the 4 × 22-bp iteron plasmid.

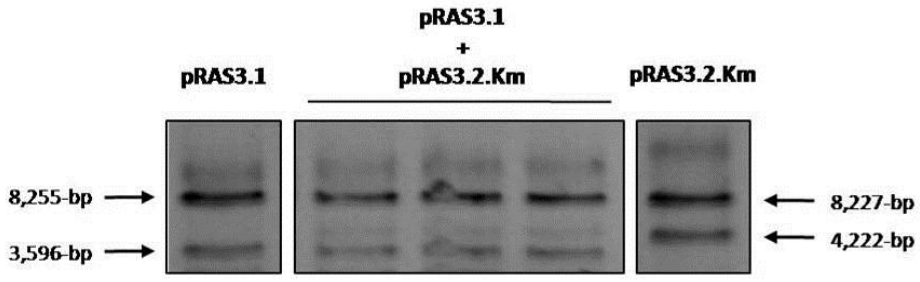


FIG. 4.1. The copy number of pRAS3.2 appears to be suppressed when it is coresident with pRAS3.1. When plasmid DNA extracted from *E. coli* DH5α cells harbouring coresident pRAS3.1 and pRAS3.2.Km plasmids was cut with SalI restriction endonuclease to differentiate between the two plasmids, the intensity of the 4,222-bp DNA fragment originating from pRAS3.2 was so faint compared to the 3,596-bp fragment of pRAS3.1 that it was barely visible. As the DNA fragments which are compared in this experiment were derived from two plasmids extracted from the same culture, an external control as described in the material and methods was not included.

4.3.2 Effect of iteron copy number and the availability of RepC on the displacement of related plasmids.

It was demonstrated that a 3 × 22-bp iteron plasmid was readily displaced from a host by a co-resident 4 × 22-bp iteron plasmid, irrespective of whether the 3 × 22-bp iteron plasmid had a

copy number that was higher (pRAS3.1.35) or lower (pRAS3.1.34) than that of the 4 × 22-bp iteron (pRAS3.1) plasmid when on its own in a host cell. This suggests that replication of the 4 × 22-bp iteron plasmid was favoured over that of a 3 × 22-bp iteron plasmid when both plasmids were coresident. Replication of IncQ-like plasmids is initiated by binding of the RepC monomers to its cognate iterons (Kim and Meyer, 1985; Sakai and Komano, 1996). Thus, the results suggested that the pool of available RepC monomers may be more effectively sequestrated by a plasmid with 4 tandem iterons than a plasmid with 3 × 22-bp iterons. To gain further evidence whether a 4 × 22-bp iteron plasmid was able to sequester the RepC to such an extent that it became limiting to the 3 × 22-bp iteron plasmid the intracellular competition assay (described in the previous section) was repeated using pRAS3.1 (4 × 22-bp iterons, 5 × 6-bp repeats) and pRAS3.1.35.Km (3 × 22-bp iterons, 5 × 6-bp repeats) this time with additional RepC supplied by placing pBAD28-RepC (see Chapter 2 section 2.3.4) *in trans* of the two competing plasmids. As a control the two pRAS3 plasmids were also competed with the pBAD28 vector only *in trans*. If displacement of the 3 × 22-bp iteron plasmid was simply as a result of sequestration of the RepC by the 4 × 22-bp iteron plasmid, then, when the availability of RepC in the cell is increased through increased transcription of the *repC* gene in pBAD28-RepC, the 3 × 22-bp iteron plasmid should be able to better withstand displacement or even displace the 4 × 22-bp iteron plasmid. However, the 4 × 22-bp iteron plasmid pRAS3.1 displaced the 3 × 22-bp iteron plasmid pRAS3.1.35.Km from 88 ± 4% of the colonies tested even when pBAD28-RepC was *in trans*, and from 91 ± 6% of the colonies in the vector-only control experiment (Table 4.2). The result was thus similar to the previous experiment during which a pRAS3.1 and pRAS3.1.Km displaced pRAS3.1.35.Km and pRAS3.1.35 from 79 ± 4 and 87 ± 10% of the colonies tested, respectively (Table 4.1). However, unlike before when none of the colonies tested retained only the 3 × 22-bp iteron plasmid and 2% percent retained both, a small percentage (7 ± 3%) of the colonies contained the 3 × 22-bp iteron plasmid and an approximately equal number of colonies (6 ± 0%) contained both plasmids when supplied with excess RepC by pBAD28-RepC. Thus, an increased supply of RepC provided a small amount of protection for the 3 × 22-bp iteron plasmid against displacement by the 4 × 22-bp iteron plasmid and therefore sequestration of the RepC may contribute to the displacement of the 3 × 22-bp iteron plasmid.

TABLE 4.2. Effect of iteron copy number on coresident plasmid segregation patterns in the absence and presence of excess RepC

	pRAS3.1.35.Km + pBAD28	pRAS3.1.35.Km + pBAD28-RepC	Percentage colonies with resistance to respective antibiotics			Direction and strength of segregation bias as a result of iteron copy number in the absence/presence of additional RepC
			Tet only	Km only	Tet and Km	
pRAS3.1.35	●[A]		53 ± 11	21 ± 6	28 ± 10	Tet > Km
		●	55 ± 8	17 ± 1	29 ± 5	Tet > Km
pRAS3.1	●		91 ± 6	1 ± 1	7 ± 5	4 iterons >>> 3 iterons
		●	88 ± 4	7 ± 3	6 ± 0	4 iterons >>> 3 iterons
pRAS3.1.55	●		79 ± 3	6 ± 3	14 ± 6	5 iterons >> 3 iterons
		●	67 ± 2	16 ± 2	21 ± 1	5 iterons >> 3 iterons
pRAS3.1.75	●		63 ± 14	25 ± 13	15 ± 2	7 iterons > 3 iterons
		●	63 ± 11	26 ± 13	13 ± 4	7 iterons > 3 iterons

[A] Dots indicate the plasmid combination used in each experiment. All plasmids were completely stable in the absence of antibiotic selection for the duration of the assay.

The 5 and 7 × 22-bp iteron plasmids, pRAS3.1.55 and pRAS3.1.75, respectively, were also competed against a coresident 3 × 22-bp iteron pRAS3.1.35.Km plasmid with pBAD28 or pBAD28-RepC *in trans* of the two competing plasmids. This was done to determine whether a plasmid with more than 4 × 22-bp iterons was similarly able to displace a 3 × 22-bp iteron plasmid as well as to investigate the effect of additional RepC. After ~20 generations of non-selective growth with the pBAD28 vector *in trans* (no additional RepC) the percentage of colonies containing the 5 × 22-bp iteron plasmid (pRAS3.1.55) was 79 ± 3% while the portion of the colonies containing only the 3 × 22-bp iteron plasmid (pRAS3.1.35.Km) or both plasmids together was 6 ± 3% and 14 ± 6% respectively. In the presence of additional RepC (pBAD28-RepC *in trans*) displacement of pRAS3.1.35.Km by pRAS3.1.55 decreased by ~12%, the presence of colonies containing only the 3 × 22-bp iteron plasmid (pRAS3.1.35.Km) increased by ~10% and the proportion of colonies still containing both plasmids increased by ~7%. A 5 × 22-bp iteron plasmid was thus also able to displace a 3 × 22-bp iteron plasmid, although to a lesser degree than a 4 × 22-bp iteron plasmid was, and the presence of additional RepC seemed to protect the 3 × 22-bp iteron plasmid from displacement slightly more than it did when the 3 and 4 × 22-bp iteron plasmids were competed. When the 7 × 22-bp iteron plasmid (pRAS3.1.75) was competed against a coresident 3 × 22-bp iteron plasmid (pRAS3.1.35.Km) pRAS3.1.75 was retained as the only plasmid in 63 ± 11% of the colonies with pBAD28 *in trans*,

while 26 ± 13% contained only pRAS3.1.35.Km and 13 ± 4% still contained both plasmids. The presence of additional RepC made no difference in this experiment.

In a control experiment two plasmids each with 3 × 22-bp iterons and 5 × 6-bp repeats, but one with the native tetracycline resistance gene and one with a foreign kanamycin resistance gene, pRAS3.1.35 and pRAS3.1.35 respectively, were competed against each other. This was done to establish whether having cloned a kanamycin resistance marker into pRAS3.1.35 influenced the result especially as the displacement phenotype when plasmids with 5 or 7 × 22-bp iterons were competed against a 3 × 22-bp iteron plasmid was not as prominent as when a plasmid with 4 × 22-bp iterons was competed against a 3 × 22-bp iteron plasmid. If having cloned a kanamycin resistance marker into the plasmids did not make a difference a symmetrical segregation pattern as a result of both plasmids having identical replicons would be observed. The results however showed that there was indeed a bias towards the plasmid with the native tetracycline resistance gene as 53 ± 11% contained only the tetracycline resistance plasmid and 21 ± 6% contained only the kanamycin resistance plasmid. The reason for the noticeable bias in favour of the plasmid with the native tetracycline resistance gene in this experiment, compared to the previous experiment where the bias was not quite as evident (see Table 4.1), is not clear. The copy numbers of two similar plasmids (on their own in a cell) with either the native tetracycline resistance gene or the cloned kanamycin resistance gene did not appear to differ when compared on an agarose gel (Fig. 4.2). The plasmid with the cloned kanamycin resistance gene is, however, 626-bp larger than the plasmid with the natural tetracycline resistance gene and thus a possibly is that the smaller plasmid was replicated slightly faster than a coresident plasmid with a kanamycin resistance marker.

In light of the replication bias in favour of plasmids containing the natural tetracycline resistance gene, a reciprocal experiment to the one in Table 4.2 was done wherein the resistance markers were exchanged. In other words, the 3 × 22-bp iteron plasmid, pRAS3.1.35, contained the native tetracycline resistance gene while the competing plasmids, pRAS3.1.Km, pRAS3.1.55.Km and pRAS3.1.75.Km, contained the heterologous kanamycin resistance gene. When the 3 and 4 × 22-bp iteron plasmids were competed (pRAS3.1.3 *vs* pRAS3.1.Km) in the presence of either pBAD28 or pBAD28-RepC, a trend similar to before was obtained wherein the 3 × 22-bp iteron plasmid was displaced (Table 4.3). However, a large number of colonies were detected that contained only the 3 × 22-bp iteron plasmid in both experiments (with the

pBAD28 or pBAD28-RepC vectors *in trans*). Also, with the antibiotic resistance markers swopped around, the displacement result for the experiment involving the 3 and 5 × 22-bp iteron plasmids was more symmetrical compared to before, while displacement of the 3 × 22-bp iteron plasmid by the 7 × 22-bp iteron plasmid largely disappeared. It thus seemed that when the plasmids with more than 3 × 22-bp iterons contained the kanamycin resistance marker, their ability to displace a 3 × 22-bp iteron plasmid was slightly reduced as a result of the replication bias towards the plasmid with the tetracycline resistance genes, and that, for reasons which are difficult to explain, this was exaggerated by the presence of a third (pBAD28) coresident plasmid. Nonetheless, irrespective of the influence of the cloned kanamycin resistance marker and the pBAD28 vector on the segregation pattern, it was still evident that a 3 × 22-bp iteron plasmid was displaced by a 4 × 22-bp iteron plasmid, while displacement of the 3 × 22-bp iteron plasmid by

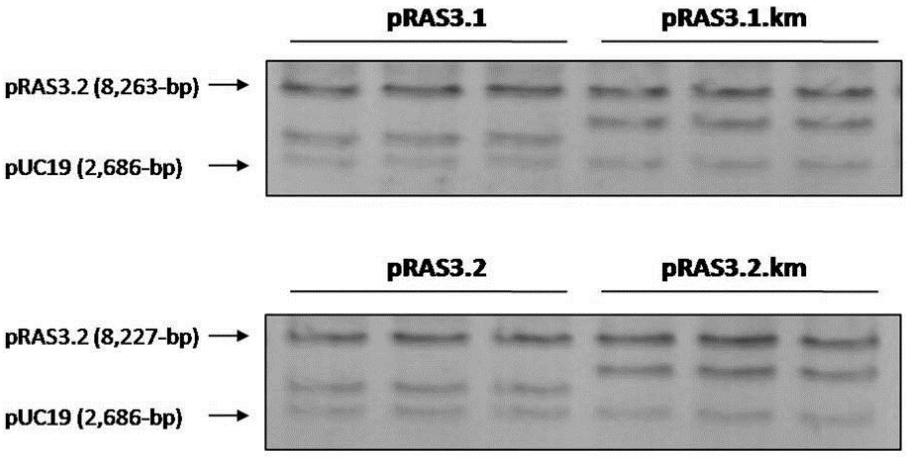


FIG. 4.2. The copy number of pRAS3.1 and pRAS3.2 does not appear to change when their tetracycline resistance genes are replaced by a kanamycin resistance marker. The band intensities of the tetracycline and kanamycin-resistant plasmids, purified from cultures with close to identical absorbance values, appeared to be approximately equal. Equal volumes of an *E. coli* DH5α culture containing pUC19 was added to each pRAS3.1 and pRAS3.2-containig *E. coli* DH5α culture prior to purification to account for differences in extraction efficiencies. The plasmid DNA was digested using SalI to linearize the pUC19 DNA and remove the pRAS3-fragments containing the resistance genes so that the intensity of the larger pRAS3 DNA fragments which are of equal size can be compared. Each sample was loaded in triplicate on the agarose gel.

the 5 and 7 × 22-bp iteron plasmids was weaker and dependent on whether or not the 3 × 22-bp iteron plasmid contained the kanamycin resistance gene. Furthermore, in both reciprocal experiments, the additional RepC provided a small amount of protection for the 3 × 22-bp iteron plasmid against displacement. In fact, in the latter experiment where the ability of plasmids with the 4, 5 and 7 × 22-bp iteron plasmids to displace a plasmid with 3 × 22-bp iterons was slightly reduced as a result of the bias towards the tetracycline resistance gene, the excess RepC provided the 3 × 22-bp iteron plasmid even more protection against displacement. Thus, it seemed that displacement of a 3 × 22-bp iteron plasmid by a 4 × 22-bp iteron plasmid could be interpreted as being the result of sequestration of the RepC by the plasmid with 4 × 22-bp iterons. As the displacement phenotype became progressively weaker when the number of iterons increased from 4 to 5 and 7, it suggested that the ability of these plasmids to sequestrate the RepC was offset by a decreased ability to initiate replication.

TABLE 4.3. Effect of iteron copy number on coresident plasmid segregation patterns in the absence and presence of excess RepC

	pRAS3.1.35 + pBAD28	pRAS3.1.35 + pBAD28-RepC	Percentage colonies with resistance to respective antibiotics			Direction and strength of segregation bias as a result of iteron copy number in the absence/presence of additional RepC
			Tet only	Km only	Tet and Km	
pRAS3.1.Km	●[A]		19 ± 10	64 ± 11	16 ± 0	4 iterons >> 3 iterons
		●	25 ± 13	66 ± 17	9 ± 7	4 iterons >> 3 iterons
pRAS3.1.55.Km	●		37 ± 10	36 ± 6	29 ± 4	5 iterons ≈ 3 iterons
		●	55 ± 27	29 ± 16	16 ± 11	5 iterons < 3 iterons
pRAS3.1.75.Km	●		73 ± 18	20 ± 23	9 ± 1	7 iterons < 3 iterons
		●	59 ± 30	26 ± 25	17 ± 4	7 iterons < 3 iterons

[A] Dots indicate the plasmid combination used in each experiment. All plasmids were completely stable in the absence of antibiotic selection for the duration of the assay.

4.3.3 Stability of the pRAS3 Plasmids and Their Derivatives

The two pRAS3 plasmids were shown to possess *pemIK*-like genes which encode a functional toxin-antitoxin system that was able to greatly enhance the maintenance of an unstable vector within an *E. coli* DH5α population (see Chapter 2 section 2.3.6). In addition, the two plasmids were also shown to have a ~1.5-fold difference in plasmid copy number (see Chapter 2 section

2.3.5). This therefore raised the question of whether the stability of these two plasmids is different and to what extent the TA system contributes to the stability of either plasmid. Furthermore, I wished to determine whether there is a difference in the metabolic burden placed upon host cells by either of these plasmids and therefore wished to know the loss frequency of either plasmid with and without the presence of a functional *pemIK* system. Should one plasmid be lost at a higher frequency than the other, the toxin of the TA system would inhibit a greater proportion of the daughter cells of the host with the less stable plasmid to continue to grow and divide. This inability to grow, due to the loss of the TA-containing plasmid, would then result in an apparently higher metabolic load.

The stability of pRAS3.1 and pRAS3.2 in *E. coli* DH5α, cultured in LB medium without antibiotic selection, was monitored over a period of ~100 generations. No loss of either plasmid was ever detected. The same result (100% stability after ~100 generations) was also found when the assay was repeated in M9 minimal medium. As it is, these plasmids appear to be highly stable, however, the apparent stability might have been due to the effectiveness of the *pemIK*-like TA system. To neutralize the effect of the pRAS3-encoded *pemIK* system upon plasmid loss, the TA system was cloned into pACYC177 to give pACYC177-TA(Ap^R) so that it could be placed *in trans* of the pRAS3 plasmids and thereby neutralize the effect of the pRAS3-encoded TA system. The functionality of the TA system was verified by placing pACYC177-TA(Km^R) (TA system cloned in opposite orientation to inactivate the ampicillin resistance gene) or pACYC177-ΔAmp *in trans* of pOU82-TA or pOU82 and monitoring the stability of the pOU82 test plasmid while maintaining antibiotic selection for the pACYC177-based plasmids. After ~90 generations, 99% of the cells tested in the culture containing pOU82-TA with pACYC177-ΔAmp as a coresident plasmid retained pOU82-TA (Fig. 4.3). This is similar to what was found when pOU82-TA was initially used to identify the *pemIK*-like genes as a TA system. When pACYC177-TA(Km^R) was present *in trans* of pOU82-TA, the stability of the test plasmid drastically decreased as only 12% of the cells tested retained pOU82-TA. This was identical to the high level of instability of the pOU82 plasmid that lacked the TA system.

Knowing that the *pemIK*-like system cloned onto pACYC177 was functional and able to neutralize the same system on a coresident plasmid, pACYC177-TA(Ap^R) was placed *in trans* of either pRAS3.1 or pRAS3.2 and their stability was monitored same as before. No change in the stability of either pRAS3 plasmid was observed. To further investigate the stability of the

pRAS3 plasmids for purposes of the metabolic load experiments to follow, I monitored the stability of the highest and lowest copy number pRAS3.1 derivatives, pRAS3.1.35 (PCN ~59) and pRAS3.1.74 (PCN ~15) respectively, in *E.coli* JM109 while the TA system was neutralized by having pACYC177-TA(Ap^R) present as a coresident plasmid. A vector-only control, in which the pACYC177 vector (pACYC177-ΔKm) without the TA system was present *in trans* of the pRAS3 derivatives, was included in parallel. Both pRAS3.1.35 and pRAS3.1.74 were found to be completely stable (100% plasmid retention) irrespective of whether the *pemIK* system was neutralized or not. The plasmid-containing cultures were grown in both LB and M9 minimal media, as well as at 30 and 37°C, however, no change in stability was evident when these parameters were changed. It thus seemed that the two natural pRAS3 plasmids were highly stable under these conditions, and the high level of stability was not dependent on the addictive nature of the TA system but rather the relatively high copy number of these plasmids. Furthermore, both the high and low copy number derivatives pRAS31.35 and

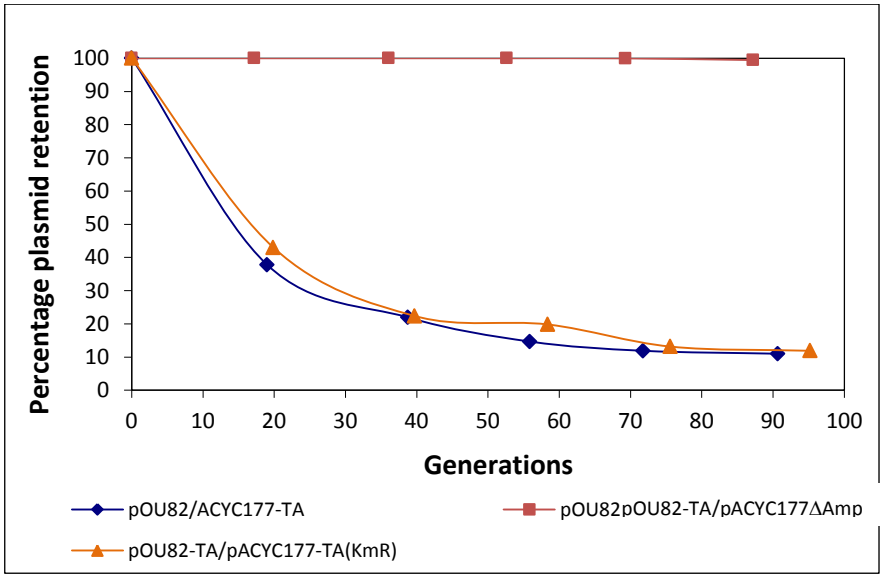


FIG 4.3 The *pemIK*-like toxin-antitoxin genes on the unstable pOU82 test vector are effectively neutralized when the *pemIK*-like genes are also present on a coresident stable pACYC177 vector.

pRAS3.1.74, respectively, were also highly stable even when their TA systems were inactivated, and thus I was able to continue with the metabolic load experiments without having to neutralize the effect of *pemIK* system for any of the pRAS3 plasmids.

4.3.4 Comparative Metabolic Loads of the pRAS3 Plasmids and Their Derivatives

The metabolic burden or fitness impact imposed by pRAS3.1 and pRAS3.2 on a host was investigated by competing plasmid-containing *E. coli* JM109 cells against isogenic plasmid-free cells over a period of 6 days, or ~40 generations, in DM25 minimal media. Each day serial dilutions of the cultures were spread onto selective and non-selective plates and the number of plasmid-containing and plasmid-free colonies obtained was used to calculate the difference in growth rate between the plasmid-containing and plasmid-free cells (Fig. 4.4). Wild-type plasmid pRAS3.2 had a smaller impact on host cell fitness with a cost (difference in growth rate) of ~4.7% than did pRAS3.1, which had a slightly higher cost of ~7.5% (Table 4.3 and Table 4.4). The ~2.8% difference in metabolic load between the two natural plasmids was small but statistically significant ($p < 0.05$). Next I determined the metabolic burden imposed by the derivatives of pRAS3.1 with varying PCNs as a result of the plasmids having different combinations of iterons and 6-bp repeats. As may be expected, plasmid pRAS3.1.34 with a PCN of ~31 plasmids per chromosome, which is approximately equal to that of pRAS3.2 (PCN ~30) had a similar ($p > 0.05$) metabolic load (~5.3%) to that of pRAS3.2 (~4.7%). Plasmid pRAS3.1.35 with the highest PCN at~59 plasmids per chromosome had the highest cost at ~10.9% while plasmid pRAS3.1.74 with the lowest PCN at ~15 plasmids per chromosome also had the lowest cost at ~2.6%. When the percentage relative fitness was plotted against PCN a linear trend, which intercepts the relative fitness-axis at 100.3% (i.e. 0% cost) when the copy number is 0 and which has a correlation value (R^2) of 0.997, was obtained (Fig. 4.5). These results therefore suggest that there is a direct relationship between plasmid copy number and metabolic load for the pRAS3 plasmids.

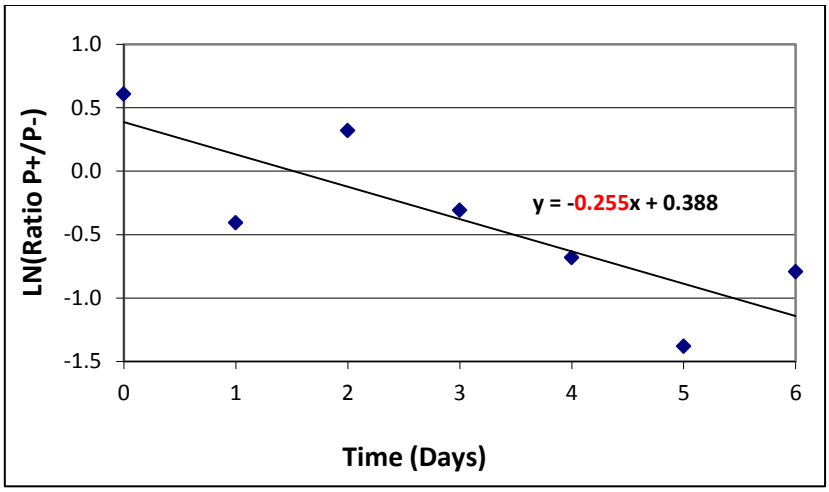


FIG 4.4. Example of a linear regression showing the difference in growth rate of an *E. coli* JM109 culture containing pRAS3.2 (P+) relative to a plasmid-free (P-) culture. The gradient of the linear regression indicates the selection rate constant (red). No statistical significance is attached to a single regression, only to the mean of all the replicates, and thus this graph is just an example of one of six graphs that were consolidated to obtain a single selection rate constant.

TABLE 4.4. Percentage fitness of a plasmid-containing host relative to a plasmid-free host

Plasmid	Approximate PCN	Selection Rate Constant[A]	Relative Fitness (%)[B, C]	N	Std. Error	p
pRAS3.2	30 ± 5	0.313 ± 0.104	95.30 ± 1.56	6	0.636	0.001
pRAS3.1.34	31 ± 1	0.354 ± 0.115	94.70 ± 1.72	5	0.769	0.002
pRAS3.1.35	59	0.725 ± 0.077	89.12 ± 1.16	6	0.475	0.000
pRAS3.1[b]	41 ± 4	0.495 ± 0.181	92.52 ± 2.64	6	1.076	0.001
pRAS3.1.74	15 ± 1	0.177 ± 0.072	97.36 ± 1.07	6	0.440	0.002

[A] The linear regressions used to calculate the selection rate constant consisted of a minimum of 3 data points.

[B] The percentage relative fitness was calculated from the selection rate constant as indicated in Appendix A.

[C] The percentage plasmid cost referred to in the text was calculated as 100 – percentage relative fitness.

TABLE 4.5. Fischer LSD test indicating the probabilities for pairwise comparisons[A] of the percentage relative fitness of each plasmid

Plasmid	pRAS3.2	pRAS3.1.34	pRAS3.1.35	pRAS3.1	pRAS3.1.74
Relative Fitness (%)	95.30 ± 1.56	94.70 ± 1.72	89.12 ± 1.16	92.52 ± 2.64	97.36 ± 1.07
pRAS3.2	-	0.554988	0.000000	0.006870	0.039753
pRAS3.1.34		-	0.000005	0.037714	0.012891
pRAS3.1.35			-	0.001311	0.000000
pRAS3.1				-	0.000021
pRAS3.1.74					-

[A] Post Hoc Tests Error: Between MS = 3.0738, df = 24.000, p values in red are significant ($p < 0.05$).

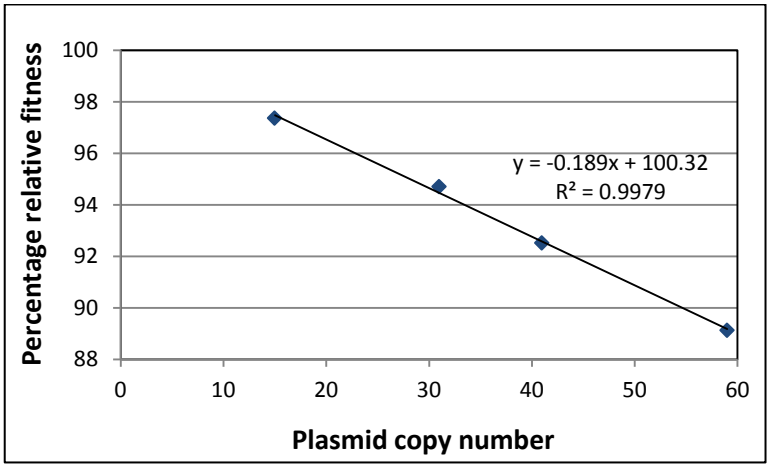


FIG. 4.5. There is a direct relationship between the copy number of a plasmid and its metabolic load on the host. A R^2 value of 0.997 was obtained when a linear trend line was fitted to the data points indicating a good correlation between the trend line and the data points and thus suggests that there was a linear relationship between the PCN of the pRAS3 plasmids and their influence on the relative fitness of their *E. coli* JM109 hosts. This was further verified by the Y-intercept of the trend line which predicts that when the PCN is 0, there will be no difference in relative fitness between the hosts. Only the data points for pRAS3.1 and its derivatives were included in this analysis as there are no mutational differences between these plasmids other than the differences in the number of 22-bp iterons and 6-bp repeats.

4.3.5 Effect of 6-bp Repeats on Mobilization Frequency

The mobilization frequencies of pRAS3.1 and pRAS3.2 were previously found to be similar (see Chapter 2 section 2.3.7). Nonetheless, the mobilization frequency of three of the pRAS3.1 derivatives was also determined to find out whether any of the iterations to the number of 22-bp iterons in the *oriV*, or more importantly, the number of 6-bp repeats in the *mobB* promoter region of the pRAS3.1 plasmid brought about a change in mobilization frequency. No significant difference in the mobilization frequency of the derivative plasmids compared to the WT plasmids was, however, evident (Table 4.5). Thus, the mobilization frequencies of the two pRAS3 plasmids were unaffected by copy number and or changes in the expression levels of the *mobB-mobA/repB* operon as a result of having 4 or 5 × 6-bp repeats.

TABLE 4.5. Mobilization frequency of the pRAS3.1 derivatives

Plasmid	Mobilization Frequency	STDev	n
pRAS3.2[A]	0.021	0.013	4
pRAS3.1.34	0.022	0.004	6
pRAS3.1.35	0.036	0.004	8
pRAS3.1[a]	0.032	0.014	10
pRAS3.1.44	0.037	0.009	6

[A] The mobilization frequencies of pRAS3.1 and pRAS3.2 were determined in Chapter 2 section 2.3.7

4.4 DISCUSSION

The two natural pRAS3.1 and pRAS3.2 plasmids differ in size only by 28-bp. The size difference is due to pRAS3.1 having 4 × 22-bp iterons and 5 × 6-bp CCCCCG repeats while pRAS3.2 has only 3 × 22-bp iterons and 4 × 6-bp repeats. An increasing number of iterons in the *oriV* from 3 to 7 copies was shown to have a reducing effect on PCN (see Chapter 3 section 3.3.2), while an increase in the number of 6-bp repeats in the *mobB-mobA/repB* promoter region from 4 to 5

copies was shown to cause an increase in PCN by increasing the *repB* expression levels (see Chapter 3 sections 3.3.4 and 3.3.5). The contributions of the iterons and the 6-bp repeats to the copy numbers of the two pRAS3 plasmids seems to be compensatory, and thus raised the question of how the individual or combined effects of the iterons and the 6-bp repeats respectively contributes to overall plasmid fitness.

Plasmid copy number needs to be balanced to ensure that the load on the host is not unnecessarily high and at the same time to ensure plasmid stability (Nordstrom and Austin, 1989; Paulsson, 2002). A high PCN would be expected to decrease the chance of plasmid loss on cell division, but too high a PCN would increase the metabolic load imposed on the host by the plasmid. In contrast a low PCN would be expected to reduce the metabolic load imposed by a plasmid on its host. It would also, however, increase the chance of plasmid loss if the plasmid does not have a partitioning system that actively ensures plasmid inheritance during cell division. Additionally, if the plasmid has a TA system such as the *pemIK*-like system that was identified on the pRAS3 plasmids (see Chapter 2 section 2.3.6), loss of the plasmids after cell division would also result in inhibition of growth or even death of the host. Should the PCN be so low that a high rate of plasmid loss occurs, cells failing to inherit the plasmid would cease to grow, thereby reducing the overall population growth rate and hence, the competitiveness of the population. The balance between metabolic load, as a result of carriage of the two pRAS3 plasmids, and plasmid stability, as a result of both the copy number and the *pemIK* TA system of these two plasmids, was therefore investigated to gain insight on the intercellular competitive fitness of cells that contain either of these two plasmids.

Experiments to determine the stability of the two natural plasmids pRAS3.1 and pRAS3.2 in the absence of antibiotic selection revealed that both plasmids were highly stable as no plasmid loss was detected over a period of ~100 generations. The high level of stability was not due to the presence of the TA system on the backbones of these plasmids as neutralization of the TA-system, by having copies of the *pemIK*-like genes present *in trans* on a multicopy vector, did not result in an increase in loss frequency. To determine how changes in the copy numbers of the pRAS3 plasmids would influence stability, the experiments were repeated, with the TA system functional as well as neutralized, for the two pRAS3.1 derivatives pRAS3.1.35 and pRAS3.1.74 with the highest (PCN ~59) and lowest (PCN ~15) copy numbers respectively. No plasmid loss was detected for either of the derivatives irrespective of the presence or absence

of the TA system, suggesting that the high level of stability of the pRAS3 plasmids under the conditions of the assays was purely as a result of their replication efficiency. This suggested that, as was found for pTF-FC2 (Matcher and Rawlings, 2009), the TA system might have a secondary role in regulation of the *repAC* genes as the *repAC* genes do not have a promoter of their own. Such a possibility is further discussed in Chapter 5.

Knowing that the two WT pRAS3 plasmids and the pRAS3.1 derivatives were very stable in *E. coli*, I was also able to determine the growth rate of plasmid-containing cultures relative to that of isogenic plasmid-free cultures, for the purposes of inferring relative fitness, without having to take into account the decreasing effect on population growth rate the *pemIK*-like system would have had upon plasmid loss. There were several advantages in measuring relative fitness by competing plasmid-containing cultures against isogenic plasmid-free cultures as opposed to directly competing two plasmid-containing cultures against each other to see which had the smallest impact on the host. Firstly, plasmid-containing cells were competed against cells that do not have the additional metabolic burden, thus amplifying the growth rate differences and allowing a more accurate estimation in a smaller time-frame. Secondly, it negated the need to have different antibiotic resistance markers on each of the plasmids for determining the numbers of each plasmid-containing culture at intervals during a competition experiment. Different antibiotic resistance genes may contribute differently to the overall metabolic load of the cell and therefore would have resulted in an artificial burden. Lastly, as the different plasmid-containing cultures were all competed against the same reference strain, the percentage relative fitness or plasmid cost inferred for each of the plasmids allowed for comparisons between any of the samples rather than just between those that were paired up in an assay. A disadvantage, however, was that the conclusions derived from comparing the relative fitness between the different plasmids were less direct.

The metabolic load imposed by each of the plasmids compared well with their respective copy numbers and indicated a linear relationship between the PCN of the pRAS3 plasmids and their impact on the relative fitness of the host compared to a plasmid-free host. From the linear regression it could be extrapolated that a pRAS3-like plasmid with a copy number of 1 plasmid per chromosome will have a 0.189% impact on the relative fitness of the host and that the relative fitness will decrease by 0.189% for every additional plasmid copy as the PCN increases from 1 to 59 plasmids per chromosome (Fig. 4.6). It was interesting to note that the regression

line indicated as expected that a cell with zero plasmids would have 100.3% of the growth rate and a plasmid copy number of 529 would reduce the growth rate to 0%. However, whether this relationship remains linear beyond 59 plasmids per chromosome is not known. Although mutations that alter gene expression or protein function could also alter metabolic load (Applebee *et al.,* 2008; Davis and Mingioli, 1950; Fong *et al.,* 2005; Reynolds, 2000; Schrag *et al.,* 1997), the differences (15 amino acid substitutions) in the tetracycline resistance proteins of pRAS3.1 and pRAS3.2 did not influence plasmid cost as there was no significant difference in the relative growth rate of pRAS3.2 and pRAS3.1.34, which have identical replicons but with the latter plasmid having *tetAR* genes that are identical to that of pRAS3.1 (Table 4.3). With this it was thus found that, for the two natural plasmids, pRAS3.1 placed a small, but significantly increased (~2.8% greater), metabolic load on an *E. coli* JM109 host compared to pRAS3.2 and this was as a direct result of pRAS3.1 a result of it having a ~1.5-fold higher PCN than pRAS3.2.

The intracellular competitive fitness of the two pRAS3 plasmids was investigated by allowing cells containing two coresident plasmids to grow in the absence of antibiotic selection for ~20 generations before determining plasmid segregation patterns by replica plating onto differential plasmid selective media. Plasmid pRAS3.1, which has 4 × 22-bp iterons and 5 × 6-bp repeats, displaced pRAS3.2, which has 3 × 22-bp iterons and 4 × 6-bp repeats, from 98% of the colonies that were tested. The remaining 2% contained both plasmids. No colonies in which pRAS3.2 displaced pRAS3.1 were ever observed, and thus pRAS3.1 seemed to have a definite advantage over pRAS3.2 when the two plasmids were coresident. The ability of pRAS3.1 to displace pRAS3.2 was determined to be due to the presence of the additional 22-bp iteron within its *oriV* and not the 6-bp repeats, as when the effect of the additional iteron was isolated by competing plasmids with 3 and 4 × 22-bp iterons, but with similar 6-bp repeats, the 4 × 22-bp iteron plasmid still effectively displaced the 3 × 22-bp iteron plasmid. Furthermore, pRAS3.1 was able to displace both pRAS3.2 and pRAS3.1.35, irrespective of whether they had a lower or higher copy number (when on their own in a cell) than pRAS3.1. This suggested that 4 × 22-bp iteron plasmids were preferentially replicated over the 3 × 22-bp iteron plasmids. A possible reason for this could be that binding of RepC monomers to 4 × 22-bp iterons was enhanced due to the phenomenon of cooperative binding and would therefore have resulted in sequestration of the RepC. Such cooperative binding by the π replication

initiator protein to the 7 × 22-bp iterons of plasmid R6K has been shown (Bowers *et al*., 2007). To test this, plasmids with 4, 5 and 7 × 22-bp iterons were competed against a plasmid with 3 × 22-bp iterons while in the absence and presence of excess RepC. It was argued that if a plasmid with 4 × 22-bp iterons titrated the available RepC monomers away from the plasmid with 3 × 22-bp iterons, then an excess supply of RepC would enable a plasmid with 3 × 22-bp iterons to better withstand displacement, or even displace the 4 × 22-bp iteron plasmid. This is what was observed. The additional RepC did lend a small amount of protection against displacement when the 3 × 22-bp iteron plasmid (pRAS3.1.35) was competed against the 4 and 5 × 22-bp iteron plasmids pRAS3.1 and pRAS3.1.55, respectively, however the effect was less dramatic than one would have expected. As the RepC protein binds to the iterons as monomers, but has also been shown to form dimers (Sakai and Komano, 1996), it is possible that overexpression of the *repC* gene *in trans* only marginally contributed to the amount of available RepC monomers needed to replenish that which had already been sequestered by the 4 tandem iterons. Such an effect wherein the formation of dimers dampened the effect of an oversupply of initiator has been described as a homeostatic copy number control mechanism for prophage P1 (Das *et al.,* 2005).

Plasmids with 5 and 7 × 22-bp iterons were less able to displace a 3 × 22-bp iteron plasmid than was a 4 × 22-bp iteron plasmid. The ability of a 7 × 22-bp iteron plasmid to displace a 3 × 22-bp iteron plasmid was so weak that when the antibiotic resistance genes were switched around such that the 3 × 22-bp iteron plasmid carried the native tetracycline resistance gene, the ability of the 7 × 22-bp iteron plasmid to displace the 3 × 22-bp iteron plasmid had largely disappeared. A possible reason for the weakened ability to displace a 3 × 22-bp iteron plasmid is that although tandem repeats of 5 or 7 × 22-bp iterons are able to sequester RepC monomers, not all binding will result in replication. It has been demonstrated that binding of RepC monomers to iterons of RSF1010 induces bending of the iterons, and the degree of bending is dependent on the number of iterons bound (Miao *et al*., 1995). Binding of a single iteron resulted in a 45° bend, and it increased to 65 and 75° when the number of RepC-bound iterons increased to 2 and 3 respectively. Binding of RepC monomers to all 5 or 7 × 22-bp iterons at the same time as opposed to only the 3 or 4 adjacent to the downstream *oriV* region, could therefore possibly result in the DNA bending to such a degree that it sterically hinders initiation of replication.

Plasmid pRAS3.1 was isolated from *Aeromonas salmonicida* subsp. *salmonicda* strains in Norway and Scotland, while pRAS3.2 was isolated from an atypical *A. salmonicida* strains in Norway (L'Abee-Lund and Sørum, 2002) and an *A. salmonicida* strain in Japan (Aoki and Takahashi, 1986). Unfortunately we were provided with the pRAS3 plasmids but not their respective hosts and the people who isolated the pRAS3 plasmids have since moved on. Although it is not known to what extent the differences between the two WT plasmids affect the plasmid-host relationship in the host populations from which these plasmids were isolated, the overall effect of the mutations on the plasmids in an isogenic *E. coli* host population, as would have been when these two plasmids first diverged from each other, is that pRAS3.2 appears to place a slightly lower metabolic burden on the host but is displaced by pRAS3.1 when both these plasmids occupy the same cell. As the differences on the backbones of pRAS3.1 and pRAS3.2 seemed to have compensatory effects on plasmid copy number, and corresponding compensatory effects on the plasmid-host relationship it gives further support to the idea that the divergence of IncQ-like plasmids such as the pRAS3 plasmids are governed by the sensitivity of their replicons to mutations.

Chapter 5

General Discussion

Aeromonas salmonicida is the causative agent of furunculosis in salmon and is thus a microorganism that has negative economic implications for the salmon aquaculture industry (Reith *et al.,* 2008). Plasmids pRAS3.1 and pRAS3.2 were identified along with three large (45 – 50-kb) conjugative plasmids, *viz.* pRAS1, pRAS2 and pRAS4, as antibiotic resistance plasmids of *A. salmonicida* strains isolated from salmon aquaculture farms in Norway (L'Abee-Lund and Sørum, 2002). Plasmid pRAS3.2 was present in five atypical *A. salmonicida* strains which were isolated between 1980 and 1990, one of which also contained pRAS1. Plasmid pRAS3.1 on the other hand, was present in five strains of *A. salmonicida* subsp. *salmonicida* which were isolated between 1991 and 1992, two of which also contained either pRAS1 or pRAS2. The same researchers also found pRAS3.1 to be present in an *A. salmonicida* subsp. *salmonicida* strain that was isolated in Scotland during 1987, while a different research group isolated a plasmid with an identical restriction profile to pRAS3.2 from an *A. salmonicida* strain in Japan during 1981 (Aoki and Takahashi, 1986).

Initial sequence analysis of pRAS3.1 and pRAS3.2 by L'Abee-Lund and Sørum (2002) indicated that the replication and mobilization systems of these two plasmids are almost identical to each other and are similar to that of the IncQ-2 plasmids pTF-FC2 and pTC-F14. The most important differences between pRAS3.1 and pRAS3.2 are that pRAS3.1 has 4 × 22-bp iterons within its *oriV* and 5 CCCCCG (6-bp) repeats in a putative promoter region upstream of the *mobB* gene, while pRAS3.2 has only 3 × 22-bp iterons and 4 × 6-bp repeats. The only other differences between the two plasmids are a number of point mutations within the *tetAR* genes which has brought about ten amino acid substitutions in the TetA and five in the TetR proteins. Considering the small number of differences between these two plasmids, it is likely that they diverged from a common ancestor relatively recently. As the iterons of IncQ-like plasmids are involved in incompatibility and copy number control (Lin and Meyer, 1986), and the 6-bp repeats of the two pRAS3 plasmids are located in a region which have been demonstrated in both IncQ-1 and IncQ-2 plasmids to contain promoter activity (Derbyshire and Willetts, 1987;

Frey *et al*., 1992; Gardner and Rawlings, 2004; Scholz *et al.*, 1989), it is interesting that both pRAS3.1 and pRAS3.2 have managed to persist in the environment. Thus, as part of an effort to understand the evolution of IncQ-like plasmids, the overall goals of this study were to determine how the two pRAS3 plasmids related to other IncQ-like plasmids and how the differences between pRAS3.1 and pRAS3.2 have contributed to the evolutionary fitness of these two plasmids.

A comparison of the pRAS3 sequences to those of the IncQ-2 plasmids pTF-FC2 and pTC-F14 indicated that their mobilization and replication systems share a common evolutionary ancestor while the *repC*, TA system and *tetAR* genes were probably acquired during separate horizontal gene exchange events. The source of the *repC* gene remains unknown as no closely related homolog could be found during BLAST searches against the non-redundant database at NCBI. The TA system on the other hand was found to be closely related to pairs of putative genes from *Xanthomonas campestris*, *X. axonopodis, Aeromatoleum aromaticum* and *Nitrosomonas europea* and distantly related to the PemIK systems of plasmids R1 (Bravo *et al.,* 1987) and R100 (Tsuchimoto *et al*., 1988) as well as the chromosomally-encoded MazEF system of *E. coli* (Masuda *et al.,* 1993) (see Chapter 2 Figure 2.6). The *tetAR* genes are located on a 2620-bp region, which for pRAS3.1 and pRAS3.2, bears 98 and 99% nucleotide identity, respectively, to a tetracycline- and arsenic-resistance island on R46 (see Chapter 2 Figure 2.2). The mosaic nature of the pRAS3 DNA, as well as the integration of 10.2-kb of pRAS3.2 DNA into the chromosome of the intracellular pig pathogen, *Chlamydia suis* (Dugan *et al.*, 2004), thus demonstrates the active participation of the pRAS3 plasmids in the horizontal gene pool.

Experiments to determine the incompatibility relationship of pRAS3 with other IncQ-like plasmids revealed that the pRAS3 replicon was compatible with the replicons of RSF1010, pIE1107 and pIE1130, which are representative of each of the IncQ-1α, -β and -γ subgroups, respectively, as well as pTF-FC2 and pTC-F14, which are the sole α and β members of the IncQ-2 subgroup. The strong displacement of pTF-FC2 and pTC-F14 by pRAS3 that was observed during initial incompatibility experiments was found to be mediated either by a small 251-aa ORF, encoded by the *orf3* gene on the pRAS3 plasmids, or by a combination of ORF3 and the MobCDE proteins (see Chapter 2 Figure 2.9). As inactivation of ORF3 did not affect the mobilization frequency of pRAS3, its function in pRAS3 and its role in the displacement of the IncQ-2 plasmids remain unknown. Due to strength of the displacement phenotype that was

observed when pRAS3 was cotransformed with either pTF-FC2 or pTC-F14, it is likely that the ORF3 protein, either by itself or in combination with MobCDE, interfered with replication of the two IncQ-2 plasmids through non-productive binding of the proteins to the DNA. Such binding could possibly take place at the *oriT* regions of pTF-FC2 and pTC-F14 as the *oriT* region is not only the binding site for the MobCDE mobilization proteins (Bahassi *et al*., 1999; Rohrer and Rawlings, 1992; Scherzinger *et al.,* 1993; Van Zyl *et al.*, 2003) but also contains a cluster of promoters responsible for transcription of the *mobB-mobA/repB* and *mobCDE* operons (Frey *et al*., 1992; Gardner and Rawlings, 2004; Matcher and Rawlings, 2009). A possible strategy to identify the exact position at which such non-productive binding occurs would be to digest pTF-FC2 and pTC-F14 DNA, which have been incubated in the presence of purified ORF3 and MobCDE proteins, with a DNAse I enzyme (Brenowitz *et al.,* 2001; Hampshire *et al.,* 2007) and then to determine the DNA region which was protected by ORF3 or possibly combinations of ORF3 and the MobCDE proteins. Once such a region is identified, the ability of the pRAS3 proteins to outcompete the native proteins of pTF-FC2 or pTC-F14, thought to bind at that specific position, could be tested by monitoring the binding affinities (Hampshire *et al.,* 2007) of the native pTF-FC2 or pTC-F14 protein in the presence of the pRAS3 ORF3 with or without combinations of the pRAS3 MobCDE proteins. However, this incompatibility phenotype between pRAS3 and pTF-FC2 or pTC-F14 was possibly an artificial lab phenomenon rather than a natural replication or partitioning phenomenon, so this study was persued.

It was observed by L'Abee-Lund and Sørum (2002) that the two pRAS3 plasmids were mobilized at high frequencies by both of the conjugative plasmids, *viz.* pRAS1 and pRAS2, which were coresident with the pRAS3 plasmids in some of the *A. salmonicida* strains. When the Mpf system was provided by the RP4 *Mpf* genes on the chromosome of *E. coli* S17.1 and the donor and recipient cells were allowed to mate for 30 min at 37°C in this study, the mobilization frequencies of pRAS3.1 and pRAS3.2 were found to be 3.2 (±1.4) × 10^{-2} and 2.1 (±1.3) × 10^{-2} transconjugants per donor, respectively. Plasmid RSF1010 was mobilized at high frequencies by IncP plasmids (Cabezon *et al*., 1994), relatively efficiently by IncF, IncFVI, Inclα, IncM and IncX plasmids and less efficiently by IncN and IncW (Cabezon *et al*., 1994; Francia *et al.*, 2004; Guerry *et al.,* 1974; Willits and Crowther, 1981), and the mobilization frequency of pTF-FC2 was found to be 3,500-fold higher than that of pTC-F14 when the RP4 Mpf formation system was provided *in trans* (Van Zyl *et al.*, 2003). IncQ plasmids thus seem to be differently

mobilized by different conjugative plasmids and it was suggested by Van Zyl *et al.* (2003) that a possible reason for such differences could be that the accessory mobilization proteins of IncQ-like plasmids are adapted to function optimally with the conjugative partner plasmid with which they are naturally associated. Whether mobilization of the pRAS3 plasmids by the Mpf systems pRAS1 (IncU) or pRAS2 (unknown incompatibility group) would be more or less effective than mobilization by the Mpf system of RP4 is not known.

Experiments to determine the functional relationship between the mobilization systems of pRAS3, pTF-FC2 and pTC-F14 revealed that the pRAS3 mobilization system was able to mobilize the *oriT*s of pTF-FC2 and pTC-F14 at relatively high frequencies ($\pm 1.2 \times 10^{-1}$ and $\pm 3.42 \times 10^{-4}$ transconjugants per donor per hour, respectively), while the latter two plasmids were not able to able to mobilize the *oriT* of pRAS3. Even though the *oriT* of pTC-F14 is more similar to that of pRAS3 (87% nucleotide identity), it was mobilized by pRAS3 at a 350-fold lower frequency than the *oriT* of pTF-FC2, which shares only 62% identity to that of pRAS3. Meyer (2009) demonstrated that the MobA relaxase of R1162 is able to initiate transfer from a 19-bp degenerate sequence. He also showed that when such sequences occur by chance on the chromosome they can serve as recognition sites for the relaxase and be mobilized at low frequencies. The specificity of the R1162 is thus relaxed, and the implication, as pointed out by the author, is that mobilization from random degenerate sequences by such a MobA could possibly serve as a previously unrecognized potential for horizontal transfer of chromosomal or other DNA. In the context of the findings of Meyer (2009), the results of in study suggest that the mobilization proteins of pRAS3 might be more relaxed than that of the other IncQ-2 plasmids, and that specific nucleotides within the *oriT*-like sequences might increase the probability of transfer.

The similarity of the pRAS3 sequence and its genetic organization compared to that of the IncQ-2α and -β plasmids pTF-FC2 and pTC-F14, respectively, indicated early on that the pRAS3 plasmid are related to these plasmids. The ability of the pRAS3 mobilization system to compliment the *oriT*s of pTF-FC2 and pTC-F14 verified that their mobilization systems are related, however, as no replicon-associated incompatibility was evident, the pRAS3 plasmids should not be classified into either of the α and β IncQ-2 subgroups. It was therefore suggested that the pRAS3 plasmids be classified into a new IncQ-2γ subgroup. Probably the biggest reason for the lack of replicon-associated incompatibility with pTF-FC2 and pTC-F14

was that the RepC of pRAS3 seems to have been acquired by horizontal gene exchange, and thus demonstrates that the evolution of new incompatibility groups within the IncQ plasmid family could, in addition to the accumulation of point mutations, also occurs by exchange of the replication initiator.

The copy numbers of pRAS3.1 and pRAS3.2 were estimated to be 45 ± 13 and 30 ± 5 plasmids per chromosome, respectively, in *E. coli* DH5α using absolute quantification Real-Time PCR techniques. A similar result was obtained using relative quantification Real-Time PCR techniques. By constructing a series of pRAS3.1 derivative plasmids with 3, 4, 5 and 7 × 22-bp iterons it was shown that PCN decreases with increasing iteron copy number (see Chapter 3 Table 3.1). It was also found that having 5 × 6-bp repeats in the *mobB-mobA/repB* promoter region such as in pRAS3.1, as opposed to 4 in pRAS3.2, resulted in a ~2-fold increase in transcription of the *repB* gene, and that the resulting increase in the concentration of RepB mediated an increase in PCN. Overexpression of the *repAC* and *repC* genes *in trans* of pRAS3.1.35 did not have an effect on PCN, suggesting that unlike for the *repB* gene, the physiological levels of transcription of the *repAC* operon was not rate limiting to plasmid replication. The difference in copy numbers of the two pRAS3 plasmids was thus found to be the result of the difference in both the number of iterons within their *oriV*s and the number of 6-bp repeats within their *mobB-mobA/repB* promoter regions. Why the copy numbers of the pRAS3 were higher than the 10 to 16 plasmids per chromosome that has been described for other IncQ-like plasmids (Rawlings and Tietze, 2001), how the number of iterons contribute to the copy number of a pRAS3 plasmid and why an excess supply of RepC initiator *in trans* of a pRAS3 plasmid did not result in an increase in PCN are interesting questions pertaining to PCN control that were raised during the course of this study and will be discussed later in this chapter.

The effect of the differences in iteron copy number on plasmid incompatibility (intracellular competition) and of differences in PCN on plasmid stability and plasmid cost (intercellular competition) was determined to gain insight on how each of these mutations contributed to overall plasmid fitness. It was found that, when coresident, pRAS3.1 was able to displace pRAS3.2 from an *E. coli* host, and thus pRAS3.1 had an intracellular competitive advantage over a coresident pRAS3.2 plasmid. Displacement of pRAS3.2 by pRAS3.1 was found to be as a result of pRAS3.2 having 3- and pRAS3.1 having 4 × 22-bp iterons within their *oriV*s, and not as

a result of pRAS3.2 having a ~1.4-fold lower copy number than pRAS3.1, as pRAS3.1 was also able to displace the pRAS3.1-derivative pRAS3.1.35, a plasmid with 3 × 22-bp iterons (similar to pRAS3.2) but with a ~1.4-fold higher copy number than pRAS3.1 (see Chapter 4 Table 4.2). Experiments to find a possible reason why replication of coresident plasmids with 3 and 4 × 22-bp iterons was biased towards the plasmid with 4 × 22-bp iterons as opposed to the plasmid with 3 × 22-bp iterons were done, however, as the results raised interesting questions pertaining to initiation of plasmid replication and copy number control, it is discussed later in this chapter.

Both pRAS3.1 and pRAS3.2, as well as the pRAS3.1 derivatives pRAS3.1.74 (PCN ~15) and pRAS3.1.35 (PCN ~59) with the lowest and highest copy numbers, respectively, were found to be highly stable as no plasmid loss was detected during a ~100 generations growth period. When the TA system was neutralized by having a cloned copy of the TA system present on a vector *in trans* for the duration of the assay, the plasmids remained highly stable. Thus, the high level of stability in *E. coli* DH5α and JM109 was most likely due to the high PCN of these plasmids and not due to a stabilizing effect from the TA system. It is possible, however, that the pRAS3 plasmids may be less stable in different hosts. De Gelder *et al.* (2007) showed that the stability of pB10, a broad-host-range IncP-1β plasmid, was highly variable in 19 different strains of *Alpha-*, *Beta-* and *Gammaproteobacteria*, ranging from no loss to rapid loss in ~80 generations. They also found that stability was variable even within strains of the same genus and as well as within strains of same species. Plasmid pB10 contains an active partitioning system and thus possible reasons for the variable stability was suggested to be negative effects on the partitioning system as well as differences in replication efficiency due to the requirement of host-encoded factors. Although replication of IncQ-like plasmids is independent of many host encoded factors such as the DnaG primase, DnaB helicase and DNA helicase loader, it does require proteins such as DnaZ, the γ-subunit of DNA polymerase III, and DNA gyrase (Scherzinger *et al*., 1991). The efficiency of the PasABC TA system of pTF-FC2, for example, was shown to be strain dependent (Smith and Rawlings, 1998a). After 100 generations in various *E. coli* hosts, a pOU82 test plasmid containing the *pasABC* genes was stabilized ~2.5-fold in *E. coli* JM105, ~100-fold in *E. coli* CSH50 and ineffective in *E. coli* JM107 and JM109. It was observed that the lethality of the PasB toxin was far higher in strains in which the *pasABC* system was most effective, and that there was a decrease in host cell death

following plasmid loss in *E. coli* mutant strains lacking the Lon protease, which is required for degradation of the PasA antitoxin. Thus, stability of the pRAS3 plasmids in different hosts remains to be tested, and might be dependent on the efficiency of replication due to host-encoded factors, the dependence of the TA system on host-encoded proteases such as Lon (Smith and Rawlings, 1998a) or Clp (Gerdes *et al*., 2005), as well as on the presence of similar host-encoded TA systems which would neutralize the effect of the TA system (Hazan *et al*., 2001).

Seeing as the TA system of the pRAS3 plasmids appears to non-essential, at least in *E. coli*, it does raise the question of why the pRAS3 plasmids encode a TA system. Matcher and Rawlings (2009) demonstrated that besides functioning as a stability system, the PasABC system of pTF-FC2 has a second biological role partly due to its precise location within the replicon. The *pasABC* promoter is a strong autoregulated promoter. As there are no transcription termination signals between the *pas* promoter and the carboxy-terminus of *repC*, it also functions as an additional regulatory system for the *repAC* genes. During periods of low copy number, such as after cell division or upon entry into a new host, autorepression of the *pas* promoter is relieved and results in elevated transcription not only of the *pasABC* genes, but also the *repAC* genes. This momentary burst of *repAC* expression is thought to ensure elevated RepA and RepC protein concentrations which would increase the probability of replication. Like pTF-FC2, the *repAC* operon of pRAS3 does not have its own promoter and there are no transcription termination signals between the promoter of the *pemIK*-like genes and the stop codon of the *repC*. Therefore the TA system of pRAS3 probably has a similar dual role in enhancing plasmid stability and maintenance under low copy number conditions.

The metabolic load imposed on a host by the pRAS3 plasmids and their derivatives was determined by competing plasmid-containing *E. coli* JM109 hosts against isogenic plasmid-free hosts. In doing so it was established that there is a direct correlation between the copy number of the pRAS3 plasmids and the metabolic load that they impose on the host. A plasmid such as pRAS3.1.35, with a PCN of ~59 plasmids per chromosome was found to decrease the relative fitness of its host by ~10.8%, while pRAS3.1.74, with a PCN of ~15 plasmids per chromosome was found to decrease the relative fitness of its host by only ~2.6% (see Chapter 4 Table 4.4). The impact of a plasmid on the relative fitness of its host over evolutionary time is not necessarily constant as plasmids and their hosts have been shown to

co-evolve in order to reduce the fitness impact (Bouma and Lenski, 1988; De Gelder *et al.,* 2008; Modi and Adams, 1991; Lenski *et al*., 1994). The mutations that improved the plasmid-host relationship manifest on either or both the plasmid and host chromosome. As pRAS3.1 and pRAS3.2 were isolated from *A. salmonicida* subsp. *salmonicida* and atypical *A. salmonicida*, respectively, with or without pRAS1 or pRAS2, the fitness impact data obtained during these experiments is not indicative of the impact of these plasmids on their natural hosts. However, when the ancestral pRAS3 plasmid first diverged to bring about pRAS3.1 and pRAS3.2, the mutations occurred in an isogenic host background and therefore, the trend (not the absolute values) is likely to be indicative how the plasmids initially could have impacted their hosts.

Knowing how the additional iteron and 6-bp repeat of pRAS3.1 contributes to its ability to compete with pRAS3.2 within a cell as well as within an isogenic population, I am able to speculate on the evolution of pRAS3.1 and pRAS3.2 and why both plasmids have persisted in the environment. Although one cannot be certain as to which plasmid originated first, or which of the mutations appeared first, a likely scenario is the following. Plasmid pRAS3.2 most closely resembles other IncQ-like plasmids as it has only 3 × 22-bp iterons and a copy number that, although still relatively high, is closer to that of other IncQ-like plasmids than is the copy number (and number of iterons) of pRAS3.1. It is also interesting that the pRAS3-like plasmid that has been integrated into the chromosome of *C. suis* has 4 × 6-bp repeats and *tetAR* genes that are similar to that of pRAS3.2 rather than pRAS3.1 (see Chapter 2 section 2.4) (Dugan *et al.*, 2004). Unlike pRAS3.2 however, the integrated plasmid on the chromosome of the *C. suis* strain has only 1 remaining perfectly conserved iteron. It appears as if the other 2 × 22-bp iterons (if the integrated plasmid initially had 3 × 22-bp iterons similar to pRAS3.2) was precisely deleted, probably as an active *oriV* on the chromosome would have interfered with chromosomal replication. Finally, the *tetAR* genes of pRAS3.2 are identical to that of R46, whereas the *tetAR* genes of pRAS3.1 contain a number of unique point mutations. Thus for these reasons, it is likely that pRAS3.2 was the ancestral plasmid.

Plasmid pRAS3.1 probably arose from pRAS3.2 in two steps, the first being the acquisition of the additional 6-bp repeat. This would have resulted in an approximately doubling of the PCN to ~59 plasmids per chromosome and eventual random segregation of the two sister plasmids. As a doubling in the PCN would have resulted in an increase in the metabolic load on the host (Chapter 4 Table 4.4), this would have provided the evolutionary pressure for a compensating

mutation on the plasmid intermediate to decrease the PCN. Such a mutation would have been the acquisition of an additional 22-bp iteron which would have resulted in a reduction of the PCN to ~41 plasmids per chromosome. At the same time, and as was shown during this study, the additional iteron would have resulted in displacement of the parent plasmid from the host even though the parent plasmid had a higher PCN than the newly formed four-iteron plasmid. Should the first modification in the ancestral plasmid have been the acquisition of a fourth iteron with a corresponding reduction in PCN to approximately ~23 plasmids per chromosome, it would still have resulted in displacement of the ancestral plasmid (Table 4.1). However, since it was shown that such a lower copy number plasmid was still perfectly stable, it seems less likely for there to have been evolutionary pressure for this plasmid to have gained an additional 6-bp repeat, thereby increasing its PCN but also placing an additional load on the host. An alternative scenario for the acquisition of an additional 6-bp repeat could have been that, as the additional 6-bp repeat increases transcription not only of the *repB* gene, but of the entire *mobB-mobA/repB* operon, this could have altered the mobilization frequency of the plasmid to increase the frequency of horizontal transfer and thereby have provided a selective advantage. Horizontal transfer has been demonstrated as an effective means to maintain plasmid stability within a population (Krone *et al.,* 2007; Watve *et al.,* 2010). However, when mobilized by the RP4 conjugative system, the mobilization frequency of plasmids with either 4 or 5 × 6-bp repeats and identical iteron copy numbers remained similar (Chapter 4 Table 4.5). Therefore, a change in the mobilization frequency is unlikely to have provided selective pressure for the acquisition of an additional 6-bp repeat.

During the course of this study on the evolution of pRAS3.1 and pRAS3.2, some interesting questions were raised regarding the initiation of replication and copy number control system of the pRAS3 plasmids, and thus also highlighted general areas in the biology of IncQ-like plasmids that require further investigation. Such a point of interest was that the copy number of pRAS3.1 and pRAS3.2 was estimated to be up to ~2.6- and ~1.9-fold higher, respectively, than the 10 to 16 plasmids per chromosome that was found for other IncQ-like plasmids (Rawlings and Tietze, 2001). There are two possible reasons for this. It was found for RSF1010 and R1162 that overexpression of the *repAC* genes from behind an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible P_{taq} promoter *in trans* of RSF1010/R1162 resulted in a 3 to 6-fold increase in PCN (Haring *et al*., 1985; Kim and Meyer, 1985). The availability of the

RepC protein of RSF1010/R1162 was thus found to be a rate-limiting factor in their replication. Overexpression of the pRAS3 *repC* from behind the L-arabinose-inducible P_{BAD} promoter *in trans* of pRAS3.1.35 did not bring about a change in its copy number and thus suggested that the availability of RepC was not a rate-limiting factor for the pRAS3 plasmid. Such a non-exhaustive supply of RepC protein would therefore result in an increase in the plasmid copy number until the availability of another component becomes rate-limiting. As overexpression of the *repAC* operon *in trans* of pRAS3.1.35, which already has an increased level of *repB* transcription due to it having 5 × 6-bp repeats, also did not result in an increase in copy number it is unlikely that the high copy numbers of pRAS3.1 and pRAS3.2 are due to a high level of *repC* transcription. It was observed that the RepC of RSF1010 binds to the iterons as monomers to initiate replication, but occurs as dimers in solution (Sakai and Komano, 1996). If the concentration at which the pRAS3 RepC proteins dimerize is higher than for other IncQ-like plasmids, it would mean that more monomers are readily available for replication, and thus provide a likely explanation as to why the pRAS3 plasmids have a higher copy number than other IncQ-like plasmids. Dimerization of the RepC protein would also explain why an excess supply of RepC *in trans* of pRAS3.1.35 did not result in an increase in PCN. An excess of RepC protein would only have contributed to the pool of dimers and thus would not have been available for replication.

The two models that have been described for the PCN control system of iteron-containing plasmids (reviewed by Del Solar *et al.* (1998)) are the titration and handcuffing models. The handcuffing model has been widely accepted as a general mechanism for the copy number control of most iteron-containing plasmids as it allows for the maintenance of PCN within narrow limits (Chattoraj, 2000; Del Solar *et al.*, 1998; Paulsson and Chattoraj, 2006). Upon the observation that an increase in the number of iterons within the *oriV* of a pRAS3 plasmid results in a decrease in PCN (see Chapter 3 Figure 3.3), the initial reaction was that such an effect would be due to increased handcuffing and thus a decrease in the initiation of replication. The ability of a four-iteron plasmid, and to a lesser extent a 5- and 7-iteron plasmid, to displace a three-iteron plasmid (see Chapter 4 Table 4.2), however, suggests that handcuffing is an unlikely reason for the decrease in copy number. As the PCN was shown to increase following a reduction in the number of functional iterons in the *oriV*, a pRAS3 plasmid with 3 × 22-bp iterons would probably have initiated replication more frequently than a coresident plasmid with 4 or more iterons, and thus, had the pRAS3 copy number been

regulated by handcuffing, the direction of displacement would have been reversed. These results therefore strengthen the argument that handcuffing is not involved in the regulation of PCN of IncQ-like plasmids (see section Chapter 1, section 1.2.4.1), however, this remains to be confirmed through direct experimentation. There are two general methods which could be used to ascertain whether or not IncQ plasmids and their initiators are capable of forming handcuffing structures. The first would be to directly visualize the presence or absence of handcuffing structures by means of electron microscopy (Gasset-Rosa *et al.,* 2008; Pal and Chattoraj, 1988). As two plasmid DNA molecules coupled by a handcuffing structure migrates differently to uncoupled DNA during 2D gel electrophoresis (Park *et al.,* 2001), the second method would be to infer the presence or absence of handcuffing structures for IncQ-like plasmids from the characteristic pattern that would be formed on a 2D gel by coupled or uncoupled DNA.

It was demonstrated that an excess supply of RepC protected R1162 from displacement by a high copy number vector containing a set of cloned R1162 iterons, which in the control experiment was responsible for titrating the available RepC and thus inhibition of R1162 replication (Kim and Meyer, 1985). Therefore, to determine whether displacement of the 3-iteron plasmid by the 4-iteron plasmid was similarly as a result of titration of the RepC initiator, plasmid pRAS3.1 and its 5- and 7-iteron derivatives, pRAS3.1.55 and pRAS3.1.75, respectively, were competed intracellularly against the 3 × 22-bp iteron plasmid pRAS3.1.35 while an excess of RepC protein was provided by having pBAD28-RepC *in trans*. The additional supply of RepC, however, provided only a small amount of protection against displacement. This result can possibly also be explained by dimerization of the RepC at high concentrations rendering the RepC inactive. As suggested by the copy number experiments in which additional RepC were supplied *in trans*, a portion of the RepC proteins at the physiological concentrations is likely to already occur as dimers. An excess supply of RepC may thus only have contributed to the pool of RepC dimers and may not have been available as monomers for replication.

Control of PCN by limiting the amount of initiator available for replication (titration model) at the transcriptional level has been suggested to be insufficient for maintaining the PCN within narrow limits as small fluctuations in initiator concentrations will result in large fluctuations in PCN (Pal and Chattoraj, 1988). Dimerization of the RepC at high concentrations could possibly limit the amount of monomers available for initiation of replication and thereby buffer

fluctuations in RepC concentrations in order to maintain a stable copy number, especially as transcription of the *repC* gene is indirectly regulated from multiple promoters. Such a role for dimerization of the initiator protein in copy number control has been described for prophage P1 (Das *et al.,* 2005). Whether the pRAS3 RepC protein is indeed capable of forming dimers, and if so, whether such a phenomenon forms part of the copy number regulation system remains to be tested. This could be done by monitoring the changes in PCN of an IncQ-like plasmid for which the *repC* gene has been mutated to produce proteins that have either an increased or decreased affinity for dimerization. In such experiments, the presence of a RepC protein that has an increased affinity for dimerization would result in a decrease in PCN, while a RepC with reduced dimerization abilities would result in an increase in PCN.

One effect not explained by titration and dimerization of the initiator is the reduced ability of plasmids with 5 or 7 × 22-bp iterons to displace a 3-iteron plasmid compared to a 4-iteron plasmid. If displacement of pRAS3.1.35 by pRAS3.1 was as a result of sequestration of the RepC by the 4 tandem 22-bp iterons of pRAS3.1, then why were pRAS3.1.55 and pRAS3.1.75, which have 5 and 7 tandem 22-bp iterons respectively, not able to displace pRAS3.1.35 as or even more effectively than was pRAS3.1? As mentioned previously (see Chapter 4 section 4.4), binding of the RepC to the iterons of RSF1010 was shown to induce bending of the DNA, and the degree of bending was dependent on whether 1, 2 or 3 iterons were bound by the RepC (Miao *et al*., 1995). In the event that an *oriV* contains more than the minimum number of iterons required for initiation of replication, binding of only the minimum number of iterons is needed for initiation provided that the initiator-bound iterons are located adjacent to each other and to the downstream A + T-rich region, as was demonstrated for R6K (McEachern *et al.,* 1985). A minimum of 3 × 22-bp iterons is required for the replication of the pRAS3 plasmids (see Chapter 3 section 3.3.1). Therefore, it is possible that binding of RepC to more than 4 × 22-bp iterons results in bending of the DNA to such an extent that it becomes inhibitory to initiation of replication, possibly by preventing the correct iterons needed for initiation from being bound by the RepC. It is interesting that when the number of 6-bp repeats were kept constant at 4 or 5 copies per plasmid for example, the total number of iterons (PCN × number of iterons) within a cell containing a plasmid with either 3 or 4 × 22-bp iterons remains similar at approximately 90 to 93 iteron copies per cell for pRAS3 plasmids with 3 or 4 × 22-bp iterons, each with 4 × 6-bp repeats, and 164 to 177 copies per cell for the same plasmids but with 5 × 6-bp repeats. If the same calculation was performed on plasmids

with 5 or 7 × 22-bp iterons (and similar 6-bp repeats), the total iteron concentration differed (80 and 105 iterons per cell for 5 and 7-iteron plasmids with 4 × 6-bp repeats, and 140 and 133 iterons per cell for 5 and 7-iteron plasmids with 5 × 6-bp repeats, respectively) from that of plasmids with 3 or 4 × 22-bp iterons. Although it is not conclusive, the (non-linear) change in the relationship of iteron and plasmid copy number for plasmids with more than 4 × 22-bp iterons suggests that the efficiency with which replication is initiated in plasmids with 5 or 7 × 22-bp iterons is different to that of plasmids with 3 or 4 × 22-bp iterons, possibly due to steric hinderance as a result of excessive or incorrect bending of the DNA. To determine whether this would indeed be the case, it would first be required to determine to what degree the pRAS3 DNA bends upon binding of RepC to 4 or more iterons. Next it would be required to determine, when there are more than 3 × 22-bp iterons present within the *oriV,* which of the iterons are required for replication. This could be done by constructing a series of pRAS3 derivative plasmids in which specific iterons have been mutated such that they are unable to bind RepC. Finally, by determining to what degree the DNA containing the various combinations of mutant iterons bends and how it correlates to the frequency of replication (measured by determining the PCN), compared to a plasmid with a similar number of tandem, functional iterons, it would give insight as to how the bending of the DNA affects the frequency of replication.

A limitation of the work presented in this thesis is that most of the experiments were carried out in *E. coli* as opposed to the natural hosts from which pRAS3.1 and pRAS3.2 were isolated. Unfortunately the *A. salmonicida* strains from which the plasmids were isolated were not available to us. However, as naturally occuring IncQ-like plasmids have an extremely broad-host-range which includes *E. coli* (see Chapter 1 Table 1.1) and are highly promiscuous (see Chapter 1 section 1.2.7), it is possible that the *A. salmonicida* host strains from which the pRAS3 plasmids were isolated are not the only hosts in which they are likely to occur. It is possible that results such as the PCN, stability and metabolic load results for example, may differ in different hosts, however, as most of the experiments were comparative between plasmids, any differences in PCN, stability and metabolic load in alternate hosts would probably be relatively similar for the two pRAS3 plasmids and the pRAS3.1 derivatives.

In conclusion, sequence comparisons showed that the two pRAS3 plasmids and the IncQ-2α and IncQ-2β plasmids, pTF-FC2 and pTC-F14, respectively, probably shared a common

ancestor. Divergence of the ancestral pRAS3 plasmid from this common IncQ-2 evolutionary ancestor was aided by the acquisition of a different *repC* gene and led to the establishment of a new IncQ-2γ subgroup. Mutations within the backbone (*oriV* and *mobB-mobA/repB* promoter) of the ancestral pRAS3 plasmid, however, also led to the existence of two similar plasmids with slightly different characteristics, but which have not yet diverged sufficiently enough to have become compatible. There is thus evidence of both macro- and microevolutionary events which have contributed to the evolution of the two pRAS3 plasmids. As pRAS3.1 and pRAS3.2 are at a fairly early stage in their divergence from each other, it would be interesting to see how these two plasmids will continue to evolve in the future. With the exponential increase in genome sequencing projects as a result of advances in cost effective, high-throughput genome sequencing (Tettelin and Feldblyum, 2009) and the expansion of centralized sequence databases, it is not unlikely that future versions of the pRAS3 plasmids might be isolated again.

APPENDIX A: Growth Media, Additives, Buffers, Solutions and Calculations

All media, solutions and buffers were sterilized by autoclaving at 121°C for 20 min. Heat labile substances were filter sterilized using 0.22 μm membrane filters (Micron separations Inc.)

A) Buffers and solutions

10 × PBS

NaCl	80.0 g
KCl	2.0 g
Na_2HPO_4	14.4 g
KH_2PO_4	2.4 g

Make up to 1 liter and autoclave. Dilute 10-fold and set pH to 7.4 using HCl before autoclaving for 1 × working stock.

10 × MOPS electrophoresis buffer (RNAse-free)

MOPS	41.85 g
$NaOAc.3H_2O$	6.8g

Make to 800 ml with DEPC-treated water. Add 20 ml 0.5 M EDTA (pH 8.0) and adjust to pH 7.0 with 10 N NaOH. Make to 1 liter, autoclave and store in the dark at 4°C.

RNA loading buffer	**(v/v)**
Formamide	56%
12.3 M formaldehyde	19%
10 × MOPS buffer	11 %
1% Bromophenol Blue	1%
Ethidium bromide	2%

SOC mecovery media

Yeast extract	5.0 g
Tryptone	20.0 g
NaCl	0.58 g
KCl	0.186 g
$MgCl_2$	0.94 g
$MgSO_4$	1.2 g

Make up to 1 liter and autoclave before aseptically adding the following:

20% (w/v) $C_6H_{12}O_6$	18.0 ml

B) Growth media

DM25 minimal media

$K_2HPO_4.3H_2O$	7.0 g
KH_2PO_4	2.0 g
$(NH_4)_2SO_4$	1.0 g
$C_6H_5NaO_7$	0.5 g

Make up to 1 liter and autoclave before aseptically adding the following:

10% (w/v) $C_6H_{12}O_6$	0.50 ml
10% (w/v) $MgSO_4$	2.0 ml
0.2% (w/v) Thiamine (vitamin B1)	2.0 ml

Luria-Bertani medium

Tryptone	10.0 $g.l^{-1}$
Yeast extract	5.0 g. l^{-1}

NaCl	5.0 $g.l^{-1}$

Solid media contained 1.5% (w/v) agar.

5 × M9 minimal media

Na_2HPO_4	33.9 g
KH_2PO_4	15.0 g
NaCl	2.5 g
$(NH_4)_2SO_4$	5.0 g

Make up to 1 liter and autoclave. For a 1 × stock dilute 10-fold in autoclaved dd.H_2O and aseptically add the following:

20% (w/v) $C_6H_{12}O_6$	20.0 ml
1.0 M $MgSO_4$	2.0 ml
1.0 M CaCl	0.1 ml (optional)

C) Media additives

Medium additive stocks were made as follows:

Ampicillin (Ap)	100 $mg.ml^{-1}$
Chloramphenicol (Cm)	20 $mg.ml^{-1}$
Naladixic acid (Nal)	30 $mg.ml^{-1}$
Kanamycin (Km)	30 $mg.ml^{-1}$
Streptomycin (Sm)	25 $mg.ml^{-1}$
Tetracycline (Tc)	20 $mg.ml^{-1}$

These stocks were stored at -20°C

X-gal (5-bromo-4-chlro-3-indolyl-β-galactoside)

X-gal (2% w/v)	0.2 g

Dimethylformamide 10 ml

The solution was stored in aliquots at -70°C.

D) Mathematical calculations

Calculation of the number of DNA molecules in a sample of known concentration.

$$\text{Number of plasmid molecules} = \frac{6.02 \times 10^{23}(\text{molecules/mol}) \times \text{DNA amount (g)}}{\text{DNA length (bp)} \times 660(\text{g/mol/bp})}$$

Calculation for the percentage per generation fitness impact.

$$\text{Per generation fitness impact} = 1 - \left(\frac{\text{Selection rate constant}}{\text{Average number of generations per day}}\right) \times 100$$

The selection rate constant is obtained from a linear regression of the log ratio of P+ to P- cells over time, and the number of generation per day is calculated from the following formula:

$$\text{Number of generations per day} = \left(\frac{\text{Log10(CFUf)} - \text{Log10(CFUi)}}{0.301}\right),$$

where CFU_i and CFU_f are the number of CFU/ml at the beginning and end of each 24 h growth cycle, respectively.

APPENDIX B: Bacterial strains and plasmids used in this study

Strains and plasmids	Description	Reference or source
Strains		
DH5α	ϕ80*dlacZΔM15 endA1 recA1 gyrA96 thi-1 hsdR17* (r_k^-, m_k^+) *relA1 supE44 deoR* Δ(*lacZYA-argF*) *U196*	Promega Corp., Madison, Wis.
CSH56	F^- *ara* Δ(*lac pro*) *supD nalA thi*	Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
EC100D *pir*$^+$	*F*$^-$ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ*80dlacZΔM15* Δ*lacX74 recA1 endA1 araD139* Δ(*ara, leu*)*7697 galU galK* λ- *rpsL nupG pir*$^+$(*DHFR*)	Epicentre® Biotechnologies
GW125a	*recA, polA* mutant of AB1157	Dorrington and Rawlings (1989)
JM109	*endA1 gyrA96 hsdR17*(r_k^-, m_k^+) *mcrB*$^+$ *recAl relA1 supE44 thi-1* Δ*(lac-proAB) F'*[*traD36 proAB lacI*^{q}*ZΔM15*]	Yanish-Perron *et al.* (1985)
S17-1	*recA pro hsdR* (RP4-2 Tc::Mu Km::Tn*7*)	Simon *et al,* (1983)
Plasmid vectors		
EZ-Tn5™	Km^R, R6K γ-*ori*	Epicentre® Biotechnologies
pACYC177	Ap^R, Km^R, p15A replicon, cloning vector	Chang and Cohen, 1978
pBAD28	Ap^R, Cm^R, arabinose-inducible expression vector, pACYC184 replicon	Guzman *et al*. (1995)
pBR322	Ap^R, Tc^R, ColE1 replicon, cloning vector	Bolivar *et al.* (1977)
pGEM-T®	Ap^R, T-Tailed PCR product cloning vector	Promega Corp., Madison, Wis.

pGEM-T®easy	Ap^R, T-Tailed PCR product cloning vector	Promega Corp., Madison, Wis.
pOU82	Ap^R, *LacZYA*, R1 replicon	Gerdes *et al.* (1985)
pUC19	Ap^R, *LacZ'*, ColE1 replicon, cloning vector	Yanish-Perron *et al.* (1985)
Plasmid constructs		
pACYC177ΔAp^R	Km^R,Amp^s, 683-bp deletion of a BamHI-ScaI fragment from pACYC177	This study
pACYC177-TA(Ap^R)	Ap^R, 731-bp PCR fragment containing pRAS3.1 *pemIK*-like genes (nt position 8 819 to 8 088)* cloned into the XhoI-BamHI sites of pACYC177	This study
pACYC177-TA(Km^R)	Km^R, 731-bp PCR fragment containing pRAS3.1 *pemIK*-like genes (nt position 8 819 to 8 088)* cloned into the BamHI-ScaI sites of pACYC177	This study
pBAD28-mobCDEorf3	Ap^R Cm^R, 2.7-kb ApaLI-ScaI fragment containing pRAS3.1 *mobCDE* and *orf3* cloned behind P_{BAD} promoter	This study
pBAD28-mobDEorf3	Ap^R Cm^R, 2.7-kb HindIII-ScaI fragment from pRAS3.1::mobC containing *mobDE* and *orf3* cloned behind P_{BAD} promoter	This study
pBAD28-orf3	Ap^R Cm^R, 1.25-kb PstI-ScaI fragment from pRAS3.1::mobE containing *orf3* cloned behind P_{BAD} promoter	This study
pBAD28-repAC	Ap^R Cm^R, 2.7-kb SalI-StuI fragment containing pRAS3.1 *repAC* cloned behind P_{BAD} promoter	This study
pBAD28-repB	Ap^R Cm^R, 1212-bp PCR fragment containing pRAS3.1 *repB* (nt position 9 891 to 8 688)* cloned behind P_{BAD} promoter	This study
pBAD28-repBAC	Cm^R, 3.5-kb PvuI-SphI fragment containing pRAS3.1 *pemIK*-like and *repAC* genes cloned into pBAD28-repB after inactivation of the pBAD28 PvuI site	This study
pBAD28-repC	Ap^R Cm^R, 1015-bp PCR fragment containing the pRAS3.1 *repC* (nt position 6 903 to 5 889)* cloned behind P_{BAD} promoter	This study
pGEM-OriV3.1	Ap^R, 742-bp PCR fragment containing pRAS3.1 *oriV* (nt position 3 123 to 2 387)* cloned into pGEM-T®	This study
pGEM-GAPA	Ap^R, 294-bp PCR fragment containing part of the *E. coli gapA* gene cloned into pGEM-T®easy	This study
pIE1108Cm	Cm^R, pIE1107 replicon with non-essential *oriVa* deleted and St^R and Km^R genes replaced by Cm^R gene	Gardner *et al*. (2001)

pIE1130	Cm^R Km^R Sm^R Su^R, natural 10 687-bp IncQ-like plasmid isolated from piggery manure	Smalla *et al.* (2000)
pOriTF14	Ap^R, 203-bp HindIII-NcoI fragment containing pTC-F14 *oriT* cloned into pUC19	Van Zyl *et al.* (2003)
pOriTFC2	Ap^R, 208-bp HhaI-HhaI fragment containing pTF-FC2 *oriT* cloned into pUC19	Van Zyl *et al.* (2003)
pOriT-RAS3	Ap^R, 196-bp PCR fragment containing pRAS3.1 *oriT* (nt position 11 820 through 0 to 176)* cloned into pGEM-T®	This study
pOU82-TA	Ap^R, 731-bp PCR fragment containing pRAS3.1 *pemIK*-like genes (nt position 8 819 to 8 088)* cloned into pOU82	This study
pR6K.3.1.repC$^{\Delta}$	Km^R, pRAS3.1::tet with *repC* and *tetR* truncated by NheI-NheI deletion	This study
pR6K.3.2.repC$^{\Delta}$	Km^R, pRAS3.2::tet with *repC* and *tetR* truncated by NheI-NheI deletion	This study
pRAS3.1	Tc^R, natural 11 851-bp plasmid isolated from *Aeromonas salmonicida* subsp. s*almonicida* with four iterons and five 6-bp repeats	L'Abee-Lund and Sørum (2002)
pRAS3.1.34	Tc^R, pRAS3.1 derivative with three iterons obtained by random ligation of short iteron fragments after BstEII digestion and four 6-bp repeats from pRAS3.2 by exchange of a 2.9-kb HindIII-PvuI region	This study
pRAS3.1.35	Tc^R, pRAS3.1 derivative with three iterons obtained by random ligation of short iteron fragments after BstEII digestion	This study
pRAS3.1.35Km	Km^R, pRAS3.1.35 with Tc^R replaced by Km^R from pSKm2 at the BamHI-EcoRV sites	This study
pRAS3.1.44	Tc^R, pRAS3.1 derivative with four 6-bp repeats from pRAS3.2 by exchange of 2.9-kb HindIII-PvuI region	This study
pRAS3.1.54	Tc^R, pRAS3.1.55 derivative with four 6-bp repeats from pRAS3.2 by exchange of 2.9-kb HindIII-PvuI region	This study
pRAS3.1.55	Tc^R, pRAS3.1 derivative with five iterons obtained by random ligation of short iteron fragments after BstEII digestion	This study
pRAS3.1.55Km	Km^R, pRAS3.1.55 with Tc^R replaced by Km^R from pSKm2 at the BamHI-EcoRV sites	This study
pRAS3.1.74	Tc^R, pRAS3.1.75 derivative with four 6-bp repeats from pRAS3.2 by exchange of 2.9-kb HindIII-PvuI region	This study
pRAS3.1.75	Tc^R, pRAS3.1 derivative with seven iterons obtained by random ligation of short iteron fragments	This study

	after BstEII digestion	
pRAS3.1.75Km	Km^R, pRAS3.1.75 with Tc^R replaced by Km^R from pSKm2 at the BamHI-EcoRV sites	This study
pRAS3.1Km	Km^R, pRAS3.1 with Tc^R replaced by Km^R from pSKm2 at the BamHI-EcoRV sites	This study
pRAS3.1::mobC	Km^R Tc^R, pRAS3.1 with *mobC* interrupted by EZ-Tn5 at position 296	This study
pRAS3.1::mobD	Km^R Tc^R, pRAS3.1 with *mobD* interrupted by EZ-Tn5 at position 1082	This study
pRAS3.1::mobE1	Km^R Tc^R, pRAS3.1 with *mobE* interrupted by EZ-Tn5 at position 1586	This study
pRAS3.1::mobE2	Km^R Tc^R, pRAS3.1 with *mobE* interrupted by EZ-Tn5 at position 1614	This study
pRAS3.1::orf3	Km^R Tc^R, pRAS3.1 with *orf3* interrupted by EZ-Tn5 at position 2089	This study
pRAS3.1::repB	Km^R Tc^R, pRAS3.1 with *repB* interrupted by EZ-Tn5 at the PvuI site	This study
pRAS3.1::tetAR	Km^R, pRAS3.1 with *tetAR* interrupted by EZ-Tn5 at the SphI-SphI sites	This study
pRAS3.2	Tc^R, natural 11 823-bp plasmid isolated from atypical *Aeromonas salmonicida* with three iterons and four 6-bp repeats	L'Abee-Lund and Sørum (2002)
pRAS3.2Km	Km^R, pRAS3.2 with Tc^R replaced by Km^R from pSKm2 at the BamHI-EcoRV sites	This study
pRAS3.2::tetAR	Km^R, pRAS3.2 *tetAR* interrupted by EZ-Tn5 at the SphI-SphI sites	This study
pTF-FC2Cm	Cm^R, natural pTF-FC2 plasmid with chloramphenicol resistance gene cloned into into the Tn*5467* transposon, called pDR412 in previous manuscripts.	Rawlings *et al.* (1984)
pTF-FC2Tet	Tc^R, Cm^R of pDR412 replaced by Tc^R of pACYC184 at the XbaI and EcoRV sites	G. Matcher
pTC-F14Cm	Cm^R, natural pTC-F14 plasmid with Cm^R inserted in at the BamHI site	Gardner *et al.* (2001)
pTC-F14Km	Km^R, pTC-F14Cm with Cm^R replaced by Km^R from Tn*5*	Van Zyl *et al.* (2003)
R6K-OriV3.1	Km^R, pRAS3.1 *oriV* from pGEM-OriV3.1 transferred to EZ-Tn5	This study
RSF1010K	Km^R, 1-1704-bp of RSF1010 replaced by Tn*903*	G. Ziegelin

* The nucleotide (nt) positions refer to the positions on pRAS3.1 to which the PCR fragments correspond

Ap, ampicillin; Cm, chloramphenicol; Km, kanamycin; Tc, tetracycline

APPENDIX C: Primers Used in this Study

Primers	Primer Sequence
E.coliGAPA Fwd	5'-TGTTAGACGCTGATTACATGG-3'
E.coliGAPA Rev	5'-CTTTAACGAACATCGGAGTGT-3'
pRAS3A Fwd	5'-GGAGCCACTATCGACTACG-3'
pRAS3A Rev	5'-GAAGCAGCCCAGTAGTAGG-3'
pRAS3MobC Fwd	5'-ACACAACAGAGCAGCTAGA-3'
pRAS3MobC Rev	5'-TCTGGTCAAGCGTGTATCC-3'
pRAS3MobE Fwd	5'-GCATCAGCGGAAGCAGCC-3'
pRAS3MobE Rev	5'-GCCTATCGCACTTCGCC-3'
pRAS3ORF3 Fwd	5'-CCGTTCGATCTGGTAGACC-3'
pRAS3ORF3 Rev	5'-GTTCTTCCATGTCTCGACG-3'
pRAS3OriT Fwd	5'-CTTGCAGGATGAGCCAGAC-3'
pRAS3OriT Rev	5'-TGGTTGCGGAGTTGACAG-3'
pRAS3OriV Fwd	5'-GTCGAATTCGTACATTATGTTTCG-3'
pRAS3OriV Rev	5'-ATAGGTACCAGTCTTTTCCATCC-3'
pRAS3OriV2 Fwd	5'-ATATCTAATCGATAGGTGAACCA-3'
pRAS3REPB Fwd[1]	5'-ACGAATTCATGTGCGGGAAG-3'
pRAS3REPB Rev[1]	5'-TCACTGCAGTGCAACATTGTA-3'
pRAS3REPB2 Fwd	5'-GCAACTATCAGGCCATCAT-3'
pRAS3REPB2 Rev	5'-TTGGGCTTGCGGTTCTC-3'
pRAS3REPC2 Rev[1]	5'-TATCTGCAGCTTGAACAGGTG-3'
pRAS3REPC3 Fwd[1,2]	5'-TAGAATTC***AGGAGG***AGGGCT**ATG**ACTCAGCAGC-3'
pRAS3SS2 Fwd[1]	5'-GAATTCAGTGGGAGAAGCTGGAAG-3'
pRAS3SS2 Rev[1]	5'-GGATCCGGAATGGTGTAGATCGTT-3'
R6KKANR Fwd	5'-CCATTCTCACCGGATTCAG-3'
R6KKANR Rev	5'-TCACCGAGGCAGTTCCATA-3'

[1] Primer includes endonuclease restriction site (underlined)

[2] Primer overlaps the start codon (**bold**) and includes an artificial ribosomal binding site (***bold italics***)

APPENDIX D: Plasmid maps

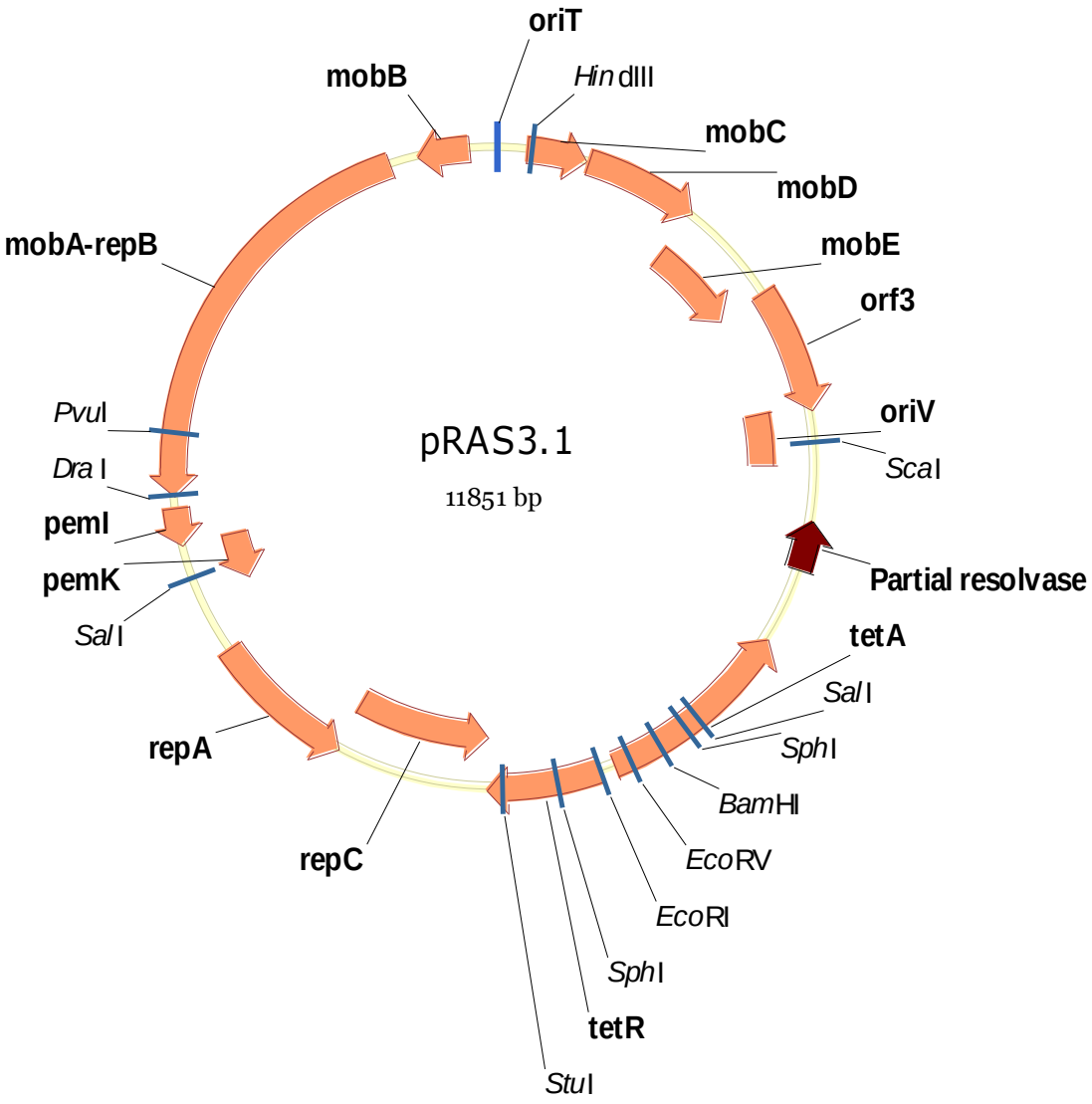


FIG. A: A circular map of pRAS3.1 (L'Abee-Lund and Sørum, 2002).

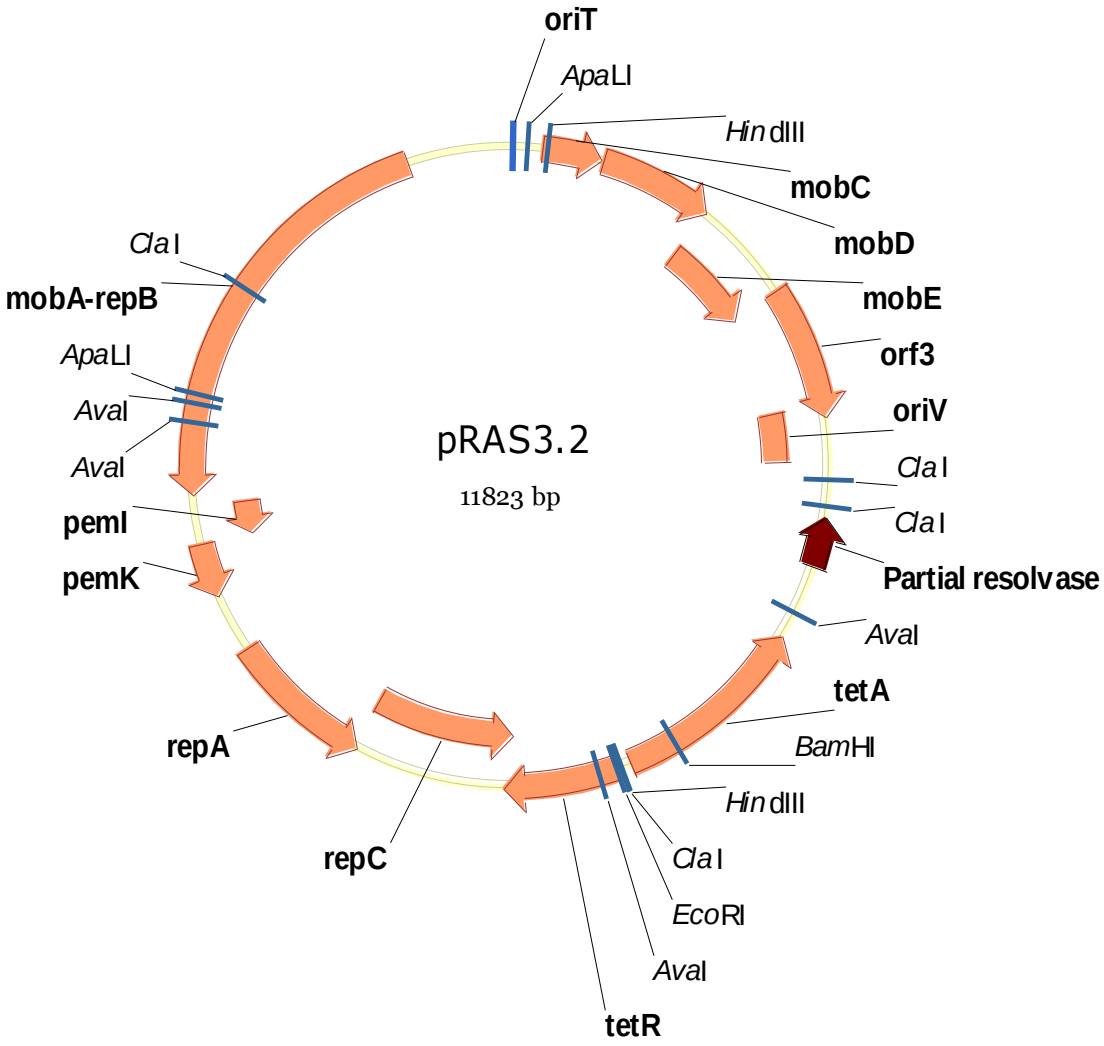


FIG. B: A circular map of pRAS3.2 (L'Abee-Lund and Sørum, 2002).

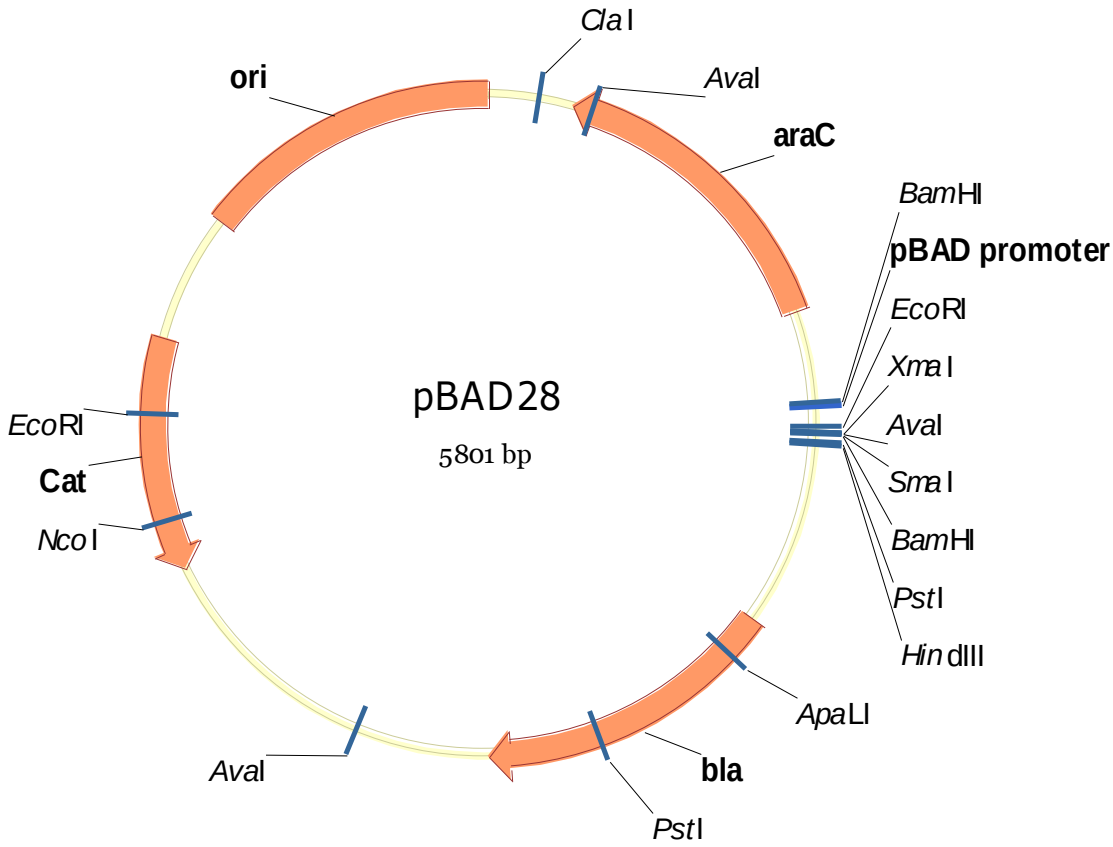


FIG. C: A circular map of pBAD28 (Guzman *et al*., 1995)

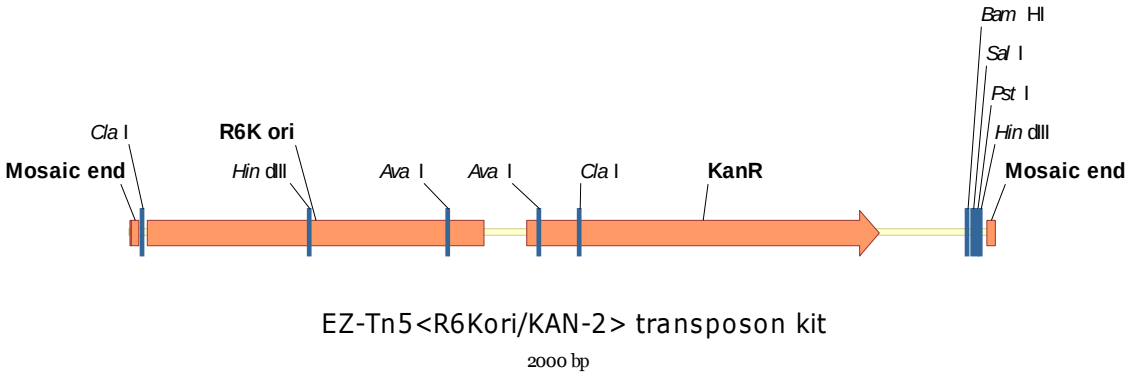


FIG. D: A linear map of EZ-Tn5 (Epicenter®).

APPENDIX E: Annotation of pRAS3.2

Universal code

```
1021      GCTGCGCACGGCGGTAGAGGAACGCGGCGGGAAATCGGTGACATTCCCGGACATGGCCGA
            L  R  T  A  V  E  E  R  G  G  K  S  V  T  F  P  D  M  A  D

1081      CCGTGTGAGCGATGACATTTACAGCAAGCGCATGAGCATTGCGGTAGCACTCAAGGAACT
            R  V  S  D  D  I  Y  S  K  R  M  S  I  A  V  A  L  K  E  L

1141      GCCCATTGGCAACCGTGCGGAGCTTACGCGGTGGCGCAACGAGGCCAAAAAGGTATTGGA
            P  I  G  N  R  A  E  L  T  R  W  R  N  E  A  K  K  V  L  E
          RBS          mobE
1201      GGCCGTAGTCCATGAGTGACCTAGACGACAGCTTTGCAAAGCTGCTCGGCAGGCAACCTA
            A  V  V  H  E  *
                        M  S  D  L  D  D  S  F  A  K  L  L  G  R  Q  P

1261      GTGACGCGGAGCGGCAAAGTCTTTACCGGGTTCGGGATGCGCTCGGCCTGAAAAACAATG
          S  D  A  E  R  Q  S  L  Y  R  V  R  D  A  L  G  L  K  N  N

1321      ATGCGCTGTGGCTGGTGCTCATGGCCTTGCAACACTACCAAGGCCAGTATGAGAAGTTCC
          D  A  L  W  L  V  L  M  A  L  Q  H  Y  Q  G  Q  Y  E  K  F

1381      CGCAAGCGATAGCACAAGCGGCCAAAGATACCTTGGTCAATTTCAAGGTAACGGCGGATG
          P  Q  A  I  A  Q  A  A  K  D  T  L  V  N  F  K  V  T  A  D

1441      CAACCGTGAAAGCATCAGCGGAAGCAGCCAAGGCGGATTTGGCGCAAGCGGTAGCAGCAG
          A  T  V  K  A  S  A  E  A  A  K  A  D  L  A  Q  A  V  A  A

1501      CAGCGCAAGAGGTCGCGCACAACACATCGGCAAAGCAGATGTGGCAGTGGGCGGCGGGCT
          A  A  Q  E  V  A  H  N  T  S  A  K  Q  M  W  Q  W  A  A  G

1561      GCATCGCGGTGGCCTTCCTGTGCGTCGGCCTGTTCGGCTGGTACATGCACAGCAGCGGCA
          C  I  A  V  A  F  L  C  V  G  L  F  G  W  Y  M  H  S  S  G

1621      AGGATTCGGGCTATCAGGCCGGATACGGCGCGGGATACACCGAAGCCAAGGACGAGAAAG
          K  D  S  G  Y  Q  A  G  Y  G  A  G  Y  T  E  A  K  D  E  K

1681      CGGCGGCGGCATGGGCCAACACGCCAGAGGGGCGCTTGGCCTACCGCTTCGCGCAAACGG
          A  A  A  A  W  A  N  T  P  E  G  R  L  A  Y  R  F  A  Q  T

1741      GAAGCCTCGCCAGCTTGGCGAAGTGCGATAGGCCGGGGTGGTACGTCGAAAAAGGCGTGT
          G  S  L  A  S  L  A  K  C  D  R  P  G  W  Y  V  E  K  G  V
                                                RBS        orf3
1801      GCTACGTCAAGCCCGCATCGGATGGCACCTACGGATGGAGGTTGCCATGAGAATCAAGAT
          C  Y  V  K  P  A  S  D  G  T  Y  G  W  R  L  P  *
                                                      M  R  I  K  I

1861      CAAAGGAGAAATCACGGCAGAACGCCTAGCGGAAGCACTGCACGCAGCAGCAGAAAAATA
            K  G  E  I  T  A  E  R  L  A  E  A  L  H  A  A  A  E  K  Y

1921      CGAAGCTGTGAGGCCGGGGCACAAGGTCTATGGGGCCAATCTCTACCTAACCGCCTTCGA
            E  A  V  R  P  G  H  K  V  Y  G  A  N  L  Y  L  T  A  F  D

1981      TGCTGATGGTCTGCCGTTCGATCTGGTAGACCATCGCGGCGAACCGCTTTCGATCACCAT
            A  D  G  L  P  F  D  L  V  D  H  R  G  E  P  L  S  I  T  I

2041      CGAGGCCAAGTCTGGCGAACTGGTAAAACCAGCACTCACAGCAGAGGGTGAAGCACACCG
            E  A  K  S  G  E  L  V  K  P  A  L  T  A  E  G  E  A  H  R

2101      ACAGAAAGCCAAAGAGGAAGCACGAAGGCAAGCTGAAGAAGCGGAAGCTGAAGCACAGCG
```

```
          Q  K  A  K  E  E  A  R  R  Q  A  E  E  A  E  A  E  A  Q  R

2161      CAGACACAGGCAAACCTTAGACGAATACGAACAGGAGCGGCAAAAACGACGGAAAAAGGA
            R  H  R  Q  T  L  D  E  Y  E  Q  E  R  Q  K  R  R  K  K  E

2221      AGCCGAAGCAAGAAAGCAATTTGAGGACGCGAACGCAATCACTGCCGAACTCTTGAAAAC
            A  E  A  R  K  Q  F  E  D  A  N  A  I  T  A  E  L  L  K  T

2281      CATGCCGGAACGCTTCATCGACGAACTGAACAAAACCGTGCAAGGCGTTTGGGACGACCT
            M  P  E  R  F  I  D  E  L  N  K  T  V  Q  G  V  W  D  D  L

2341      CAAGCCAACCGAAACGCAAGGCAAGAAGAAGGGCCAGCCGAAAGCGCTACCAGTCTTTTC
            K  P  T  E  T  Q  G  K  K  K  G  Q  P  K  A  L  P  V  F  S

2401      CATCCATGCGGACGGGCTGGTGCTTTCCGTCGAGACATGGAAGAACCCGCGCCGGGTTCT
            I  H  A  D  G  L  V  L  S  V  E  T  W  K  N  P  R  R  V  L

2461      CAATCCGCTTTGCACACTCCAGCACGGCGAAATAGCCCCATTCTGGATGCACGAAGCGTG
            N  P  L  C  T  L  Q  H  G  E  I  A  P  F  W  M  H  E  A  W

2521      GTTAGAAGCCATGAGGCGGATAGTCGATCTGCTGGATACGCTCACCGCAGCCCCGGCAGA
            L  E  A  M  R  R  I  V  D  L  L  D  T  L  T  A  A  P  A  E
                                                                5/5-bp
2581      GGCCCTAGAAAGCCAGTAGAGGCGTTTTCGCCTGCAAGGAAGGGGGTAGCGGCGGGGCGG
            A  L  E  S  Q  *
    6/6-bp 6/6-bp 6-bp tail
2641      GGGCGGGGG***TGTGTC***TGCAAGGGCTTTAGCCCGCGCAGGGCGAAGCCCGAAGGCAAGGCT

                                  15-bp Conserved region
                                       ssiA/ssiB
2701      TCAGACGACGCCTAAAGCGGCGGCTGAAAGCCGCGCATGGTAGAAGTTCTTTAGAACTTC

          A + T-rich region
2761      TTAGACCGGAATTTTGGAAAAACGCGAAAACTCAGTACTGACAAGGGTTTGCGGGCGTTT

                        BstEII   20/22-bp      BstEII    20/22-bp
2821      TTGTGGCTTGTCAGGTTACCACCTGCGCCGGGGGTGGTTACCACCTGCGCCGGGGGTGGT

          BstEII  221/22-bp[A]
2881      TACCACCTGCGCCGGGGGCTTGAATTAAATTAGATAAATCTCTTATTGATAGACAAACAC

2941      ACTACACTTAATACCCATTATGTTAAAAACGACAGGATTCATGAAAATCATCAGAATTGT

3001      GCAGGATGGTTCACCTATCGATTAGATATGGGGTCTGAAGGCCAATAGAACGAAAACGTA

3061      CGTTAGTGAAGTAACTGTCTGATATATCGAAACATAATGTACATTGGAAAACGCCATCAA

3121      AACGGTGTCTTTTTAATCGAAAATTGGCACTTAACGGACTTTCTTGTCTACTAATCGATT

3181      TTTCAGACGATAATTTTCCTCTTACCGTACGTTTTCGTACGGTTGGACCTCAAACCCCTT
                                                                    K
```

```
3241      TGAGCGTGGCAAGCTGCACAAAGGGATAGGGCGGCACCAGAGAAAAATCACTCAGGGTCA
            L  T  A  L  Q  V  F  P  Y  P  P  V  L  S  F  D  S  L  T  L

3301      ATGCCAGCGCTTCGTTAATACAGATGTAGGTGTTCCACAGGGTAGCCAGCAGCATCCTGC
            A  L  A  E  N  I  C  I  Y  T  N  W  L  T  A  L  L  M  R  R

3361      GATGCAGATCCGGAACATAATGGTGCAGGGCGCTGACTTCCGGTGCCAGCAGATATTTTG
            H  L  D  P  V  Y  H  H  L  A  S  V  E  P  A  L  L  Y  K  P

3421      GCAGTTTGCCTTGGATCAGAGCCATCTGACGCAGGGCTAGTGCAGCCGGATAGTCAATAG
            L  K  G  Q  I  L  A  M  Q  R  L  A  L  A  A  P  Y  D  I  A
                                                          Partial resolvase
3481      CTACCGGCAGCGGTGCGGACTGTTGTAACTCAGAATAAGAAATGAGGCCGCTCATGGCGT
            V  P  L  P  A  S  Q  Q  L  E  S  Y  S  I  L  G  S  M

3541      TGACTCTCAGTCATAGTATCGTGGTATCACCGGTTGGTTCCACTCTCTGTTGCGGGCAAC

3601      TTCAGCAGCACGTAGGGGACTTCCGCGTTTCCAGACTTTACGAAACACGGAAACCGAAGA

3661      CCATTCATGTTGTTGCTCAGGTCGCAGACGTTTTGCAGCAGCAGTCGCTTCACGTTCGCT

3721      CGCGTATCGGTGATTCATTCTGCTAACCAGTAAGGCAACCCCGCCAGCCTAGCCGGGTCC
                                                                AvaI
3781      TCAACGACAGGAGCACGATCATGCGCACCCGTGGCCAGGACCCAACGCTGCCCGAGATGC

3841      GCCGCGTGCGGCTGCTGGAGATGGCGGACGCGATGGATATGTTCTGCCAAGGGTTGGTTT

3901      GCGCATTCACAGTTCTCCGCAAGAATTGATTGGCTCCAATTCTTGGAGTGGTGAATCCGT

3961      TAGCGAGGTGCCGCCGGCTTCCATTCAGGTCGAGGTGGCCCGGCTCCATGCACCGCGACG
                                  *  T  S  T  A  R  S  W  A  G  R  R

4021      CAACGCGGGGAGGCAGACAAGGTATAGGGCGGCGCCTACAATCCATGCCAACCCGTTCCA
            L  A  P  L  C  V  L  Y  L  A  A  G  V  I  W  A  L  G  N  W
                                                              (C)[C]
4081      TGTGCTCGCCGAGGCGGCATAAATCGCCGTGACGATCAGCGGTCCAATGATCGAAGTTAG
            T  S  A  S  A  A  Y  I  A  T  V  I  L  P  G  I  I  S  T  L

4141      GCTGGTAAGAGCCGCGAGCGATCCTTGAAGCTGTCCCTGATGGTCGTCATCTACCTGCCT
            S  T  L  A  A  L  S  G  Q  L  Q  G  Q  H  D  D  D  V  Q  R

4201      GGACAGCATGGCCTGCAACGCGGGCATCCCGATGCCGCCGGAAGCGAGAAGAATCATAAT
            S  L  M  A  Q  L  A  P  M  G  I  G  G  S  A  L  L  I  M  I

4261      GGGGAAGGCCATCCAGCCTCGCGTCGCGAACGCCAGCAAGACGTAGCCCAGCGCGTCGGC
            P  F  A  M  W  G  R  T  A  F  A  L  L  V  Y  G  L  A  D  A
                                                          (A)
4321      CGCCATGCCGGCGATAATGGCCTGCTTCTCGCCGAAACGTTTGGTGGCGGGACCAGTGAC
            A  M  G  A  I  I  A  Q  K  E  G  F  R  K  T  A  P  G  T  V
                                       (G)
```

```
4381      GAAGGCTTGAGCGAGGGCGTGCAAGATTCCGAATACCGCAAGCGACAGGCCGATCATCGT
           F  A  Q  A  L  A  H  L  I  G  F  V  A  L  S  L  G  I  M  T
             (T)                  (A)
4441      CGCGCTCCAGCGAAAGCGGTCCTCGCCGAAAATGACCCAGAGCGCTGCCGGCACCTGTCC
           A  S  W  R  F  R  D  E  G  F  I  V  W  L  A  A  P  V  Q  G
                                                (T)               (T)
4501      TACGAGTTGCATGATAAAGAAGACAGTCATAAGTGCGGCGACGATAGTCATGCCCCGCGC
           V  L  Q  M  I  F  F  V  T  M  L  A  A  V  I  T  M  G  R  A
                                       (T)          SalI
4561      CCACCGGAAGGAGCTGACTGGGTTGAAGGCTCTCAAGGGCATCGGTCGACGCTCTCCCTT
           W  R  F  S  S  V  P  N  F  A  R  L  P  M  P  R  R  E  G  K
                      (A)                                   (T)
4621      ATGCGACTCCTGCATTAGGAAGCAGCCCAGTAGTAGGTTGAGGCCGTTGAGCACCGCCGC
           H  S  E  Q  M  L  F  C  G  L  L  L  N  L  G  N  L  V  A  A
                                                                  (A)
4681      CGCAAGGAATGGTGCATGCAAGGAGATGGCGCCCAACAGTCCCCCGGCCACGGGGCCTGC
           A  L  F  P  A  H  L  S  I  A  G  L  L  G  G  A  V  P  G  A

4741      CACCATACCCACGCCGAAACAAGCGCTCATGAGCCCGAAGTGGCGAGCCCGATCTTCCCC
           V  M  G  V  G  F  C  A  S  M  L  G  F  H  R  A  R  D  E  G
                            (A)
4801      ATCGGTGATGTCGGCGATATAGGCGCCAGCAACCGCACCTGTGGCGCCGGTGATGCCGGC
           D  T  I  D  A  I  Y  A  G  A  V  A  G  T  A  G  T  I  G  A
                            BamHI (A)                     (C)
4861      CACGATGCGTCCGGCGTAGAGGATCCACAGGACGGGTGTGGTCGCCATGATCGCGTAGTC
           V  I  R  G  A  Y  L  I  W  L  V  P  T  T  A  M  I  A  Y  D

4921      GATAGTGGCTCCAAGTAGCGAAGCGAGCAGGACTGGGCGGCGGCCAAAGCGGTCGGACAG
           I  T  A  G  L  L  S  A  L  L  V  P  R  R  G  F  R  D  S  L
                         (A)  (T)                (A)
4981      TGCTCCGAGAACGGGTGCGCATAGAAATTGCATCAACGCATATAGCGCTAGCAGCACGCC
           A  G  L  V  P  A  C  L  F  Q  M  L  A  Y  L  A  L  L  V  G
         (G)                                              (C)     (T)
5041      ATAGTGACTGGCGATGCTGTCGGAATGGACGATATCCCGCAAGAGGCCCGGCAGTACCGG
           Y  H  S  A  I  S  D  S  H  V  I  D  R  L  L  G  P  L  V  P
                                                                  (A)
5101      CATAACCAAGCCTATGCCTACAGCATCCAGGGTGACGGTGCCGAGGATGACGATGAGCGC
           M  V  L  G  I  G  V  A  D  L  T  V  T  G  L  I  V  I  L  A
                    tetA(G)                          (T)(G)     (T)
5161      ATTGTTAGATTTCATACACGGTGCCTGACTGCGTTAGCAATTTAACTGTGATAAACTACC
           N  N  S  K  M
               HindIII                          EcoRI    tetR
         (CG)  (G)(A)   (AG)    (A)     (A)        (A)(AC)   (A)
5221      GCATTAAAGCTTATCGATGATAAGCTGTCAAGCATGAGAATTCGCGAATGAACAAGCTCC
                                                        M  N  K  L
                                    (A)
5281      AACGCGAGGCCGTGATCCGAACCGCGCTCGAACTGCTTAACGACGTGGGCATGGAAGGTC
          Q  R  E  A  V  I  R  T  A  L  E  L  L  N  D  V  G  M  E  G
                                    AvaI(T)
5341      TAACGACGCGCCGACTGGCTGAGCGCCTCGGGGGTGCAACAGCCAGCGCTCTACTGGCATT
          L  T  T  R  R  L  A  E  R  L  G  V  Q  Q  P  A  L  Y  W  H
            (G)          (A)
5401      TCAAGAACAAGCGTGCGTTGCTCGACGCACTTGCCGAAGCCATGCTGACGATAAATCACA
          F  K  N  K  R  A  L  L  D  A  L  A  E  A  M  L  T  I  N  H
                     (G)      (A)
5461      CGCATTCGACGCCAAGGGATGACGACGACTGGCGTTCGTTCCTGAAGGGCAATGCATGCA
          T  H  S  T  P  R  D  D  D  D  W  R  S  F  L  K  G  N  A  C
```

```
                                        (T)
5521    GTTTTCGACGGGCGTTGCTCGCTTATCGCGATGGCGCGCGTATTCATGCCGGGACGCGGT
        S  F  R  R  A  L  L  A  Y  R  D  G  A  R  I  H  A  G  T  R
                                     (T)
5581    CAGCCGCGCCGCAGATGGAAAAAGCCGACGCGCAGCTTCGCTTCCTTTGCGATGCTGGCT
        S  A  A  P  Q  M  E  K  A  D  A  Q  L  R  F  L  C  D  A  G
          (T)                                                   (T)
5641    TTTCGGCAGGTGACGCGACCTATGCGTTGATGGCAATCAGCTACTTCACCGTCGGCGCTG
        F  S  A  G  D  A  T  Y  A  L  M  A  I  S  Y  F  T  V  G  A

5701    TTCTTGAGCAGCAAGCTAGCGAGGCAGACGCCGAGGAGCGGGGCGAAGATCAGTTGACCA
        V  L  E  Q  Q  A  S  E  A  D  A  E  E  R  G  E  D  Q  L  T

5761    CCTCAGCGTCTACGATGCCGGCGCGCCTACAGAGCGCGATGAAAATCGTCTACGAAGGCG
        T  S  A  S  T  M  P  A  R  L  Q  S  A  M  K  I  V  Y  E  G

5821    GTCCGGACGCGGCATTCGAGCGAGGCCTGGCTCTCATCATCGGCGGTCTTGAACAGGTGC
        G  P  D  A  A  F  E  R  G  L  A  L  I  I  G  G  L  E  Q  V

5881    GGCTCAGTCCGGCGTCCAGTCCTGCGGGTCGAACAAATCTGGTTCTTGCATTGGCGGCTG
        R  L  S  P  A  S  S  P  A  G  R  T  N  L  V  L  A  L  A  A
           *  D  P  T  W  D  Q  P  D  F  L  D  P  E  Q  M  P  P  Q

5941    GCTCATGAGGTCGTGGTATTCGCTGGGGCGCATGTCGTGCTCACCCTCAAGGAACGGCAG
        G  S  *
         S  M  L  D  H  Y  E  S  P  R  M  D  H  E  G  E  L  F  P  L

6001    TTCCTTGGGACGACTTACCAGCACAAGCCCGCTGCTGGTCTGGTCGATGGCCCATCCCAT
         E  K  P  R  S  V  L  V  L  G  S  S  T  Q  D  I  A  W  G  M

6061    ACCGCCTTGCAGTCCGCGTGTGCGGTCTGAGATTTCGTCTAAAGCATCGCGTAGTTGACC
         G  G  Q  L  G  R  T  R  D  S  I  E  D  L  A  D  R  L  Q  G

6121    CCGCCTGACTGATCGAACCTTCGTTGATGCTTCTTCCCACGCCGCCTTGCCCCATACGTG
         R  R  V  S  R  V  K  T  S  A  E  E  W  A  A  K  G  W  V  H

6181    GGGCGCTAGCGTGTCGATGTGCGCCCCGTTGCCATTGCCAAGCGAACGGCCAAGCCGGAT
         P  A  L  T  D  I  H  A  G  N  G  N  G  L  S  R  G  L  R  I

6241    GTTTGAGCACAGCCAACAGTGCAGCAGCTTGGCGGCTTCACTCTCCAGCTCGTTGCGCTC
         N  S  C  L  W  C  H  L  L  K  A  A  E  S  E  L  E  N  R  E

6301    GAAAAGCGAAACCTTGACGTGTTGCTCTCGAAAAACCGCCCCGGTCAGGCGCGGATTGAC
         F  L  S  V  K  V  H  Q  E  R  F  V  A  G  T  L  R  P  N  V

6361    GGCAACAAAGATTTCCCCAGTCTTGTGATTGAACCGACTCGAAAGGAGATTGCAGCGGTA
         A  V  F  I  E  G  T  K  H  N  F  R  S  S  L  L  N  C  R  Y

6421    ATCCCAGCCGGTAGCGTGATTGATTTCCCGTATCTGCGCGTTACCCATTCGGTTCAGGCT
         D  W  G  T  A  H  N  I  E  R  I  Q  A  N  G  M  R  N  L  S

6481    TTCCTTGATTCTGGCAAGTGCCGCCCCGCTTTCTGGGTCGTCATATCCAGCATCAATCAG
         E  K  I  R  A  L  A  A  G  S  E  P  D  D  Y  G  A  D  I  L

6541    CAGGCTGCGCAAACTGGTTTTCTTGGTTTTCAGCGGCCTTCCATCATCGTTGTCAAAGTC
         L  S  R  L  S  T  K  K  T  K  L  P  R  G  D  D  N  D  F  D
```

```
6601      CAAAGCAAGGCGCAATTGCTGGCCTATTTCGCCGGGTGGTTTTGCCTCGATCATCAAGCC
           L  A  L  R  L  Q  Q  G  I  E  G  P  P  K  A  E  I  M  L  G

6661      GTCAATGCCAAGCTGGGCGCTTACTGCAAGCAAGATTGATTGGTCATCAGCTCCAAGCTG
           D  I  G  L  Q  A  S  V  A  L  L  I  S  Q  D  D  A  G  L  Q

6721      CTCGAACCCTTGAAAGCTGATTTCACTATCGCCGAAGCGGCGCGGCGGCACATACAGTCG
           E  F  G  Q  F  S  I  E  S  D  G  F  R  R  P  P  V  Y  L  R

6781      CCCGCGCTTCTTCCCCTTCGTGGGGACGAAAAGCCCGTCAAACACATGAGCGGGATCGAA
           G  R  K  K  G  K  T  P  V  F  L  G  D  F  V  H  A  P  D  F
                                                repC
6841      TTTGGCATAGGCCGGTTTAATACGCTGCTGAGTCATAGCCCTTGCCCCCCTTGTTGCCGC
           K  A  Y  A  P  K  I  R  Q  Q  T  M
                        *  Y  A  A  S  D  Y  G  K  G  G  K  N  G  S

6901      TCTCACGCTTGATCTCCTGTAGCTCTACCGGCAGCAACACCCCGCCGGTGTGTCGCTGAT
            E  R  K  I  E  Q  L  E  V  P  L  L  V  G  G  T  H  R  Q  Y

6961      ACCAGCGGTCTAGCGGGGTCGCGTCGTAGTTCTGCTTACTCACGCCAAAGCGGACAAAAA
            W  R  D  L  P  T  A  D  Y  N  Q  K  S  V  G  F  R  V  F  L

7021      GGCCGCGCCGCTCGTTGCCGATGGGCGCACGATCAAAGCCCCGGTCGCTCAGGCGCTCGG
            G  R  R  E  N  G  I  P  A  R  D  F  G  R  D  S  L  R  E  A

7081      CTTCCTGCTCCGTCATCTTGGCGACATAGCCGCACCATCTGGCGTTGTCGATCAGGGCAG
            E  Q  E  T  M  K  A  V  Y  G  C  W  R  A  N  D  I  L  A  S

7141      ACGCGCCCCGCGCTGCCTGCTGCTGGTCGGTCTGGCCGTCGCGGGCGCTGCCCTTGTTGA
            A  G  R  A  A  Q  Q  Q  D  T  Q  G  D  R  A  S  G  K  N  V

7201      CGTGGTGCAGGTACAGGACAGCCGCCCCGGTGGTTGCCGCGATGTGTTCCAACACGGAAA
            H  H  L  Y  L  V  A  A  G  T  T  A  A  I  H  E  L  V  S  V

7261      CAAGGTGGGCCATGTCGCCGTTGCTGTTCTCGTCGAGGCTGTGAATGCGGCTCAATGTGT
            L  H  A  M  D  G  N  S  N  E  D  L  S  H  I  R  S  L  T  D

7321      CCAGCACAATCAGCCGCGCCCCTGCGCTGTAGTCGATGACGCGGCGCAAGTGCGCCTCGT
            L  V  I  L  R  A  G  A  S  Y  D  I  V  R  R  L  H  A  E  D

7381      CCATCACGTTCAACCGCTTGCCCATGATCGGCTCAAGCATCAGGTTCTCAGCGATGGCTT
            M  V  N  L  R  K  G  M  I  P  E  L  M  L  N  E  A  I  A  E

7441      CGCGGGCCGACTGGTTGAGGTGCTGGCCGATGGCGTGGACACGCCGCACAAGGGCGGGCT
            R  A  S  Q  N  L  H  Q  G  I  A  H  V  R  R  V  L  A  P  Q

7501      GTGGATCTTCGCCAGCGAGATAGACCACGCGCCCGGTGTGCGCCGGGGCCAGTCCTACAA
            P  D  E  G  A  L  Y  V  V  R  G  T  H  A  P  A  L  G  V  L

7561      GGTCGCCGCCTGCCACACTGCAAGCGATTGACATGGCCGCTTCAAGAGCAAAAAAGCTCT
            D  G  G  A  V  S  C  A  I  S  M  A  A  E  L  A  F  F  S  K

7621      TGCCAGTTGCGCCAGGTGCAACAAGTGCGCCCACGGTTCCGGCCAAGAATCCGGGCCAGA
            G  T  A  G  P  A  V  L  A  G  V  T  G  A  L  F  G  P  W  I
                                                                repA
7681      TGAAATCAAGTTCTGGCGGCTCATTGGTGAAGGCCGCCATAATGTCTAAAGCCATGCTGT
            F  D  L  E  P  P  E  N  T  F  A  A  M  I  D  L  A  M
```

RBS

7741 TCTCCTGTGTAGAACGGTTATGGTTAGGCGCTGCCGGGTGTTCCCGCACCCTGCCGCGCC

7801 GCTTCTGCCTTCGCCTTGGCGGCATAGCCCGCTTCCAGACTGGCGGCATCGGTCGGCGTG

7861 TCGCCGTCGATGGGCAGGCCAGAAGCGAACATGAGCGCATCACGTAGGGCAATCTCTGCT

7921 TGCCCGTACTGGCGTCCCTGAAAGGCGCTGTGGCCTTCCGTGCCGGTGTCCCGCACGGTG

7981 CGCCCGCCGTCGATGATGCGCCACTCGTACCAAGCGTTTTCCGGCTCGCCGTAGACCTCG

8041 ATCATCGCGCCCAAGGAATCCGGGGCGGGAATGGTGTAGATCGTTTTCATTTCGCAATCC
* K A I G

8101 CCAAGAACACGGCCAAGCAACGCTCAACCTCAAGCAGTGCGTCCGCGTCGATGCGCCCGA
L F V A L C R E V E L L A D A D I R G F

SalI

8161 AAGCTGGCCCTACCTTGCTGCGCAGCACCGTCATCGCCTT*GTCGAC*CATCACCTGCGATG
A P G V K S R L V T M A K D V M V Q S P

8221 GCTTCTGCAAACCGTTCTCTGCGCTCGGCTGAACCGTGACGCGCAGCAACGGCGCGGCAT
K Q L G N E A S P Q V T V R L L P A A N

8281 TAAGTGTGCTGGTGACAGGCAGCACGGTAACGCTGCCGTGCTCGCTGAACTGGTTGGCCT
L T S T V P L V T V S G H E S F Q N A Q

8341 GAACCACAAGCGCAGGCCGTGGCTTGCCAAAGTCGCCCTGCATGGCGATGGTCACAAGGT
V V L A P R P K G F D G Q M A I T V L D

***pemK*-like RBS**

8401 CGCCGCGCATCATTCCGTCCAGCCGTCCACGTCTGCCAAGGCTTCATCCATGAACTGCTG
G R **M M**

* E T W G D V D A L A E D M F Q Q

8461 CATGGCGGTGTCGGCGCTGTCGGCCTGAGCCACAAGCAGGCATTGGCGGCGGCACTCCTC
M A T D A S D A Q A V L L C Q R R C E E

8521 CGCGAAGTCGGGGCGGCGCGTATCCGGCACCCAGATTTGAACCGGACGAAGCCCCGCCAT
A F D P R R T D P V W I Q V P R L G A M

***pemI*-like**

8581 GCGTAATGCATCGCGATGCTTCTGAACCCGCTGATTAACGTGTGCTGTTGCCATTTCCAA
R L A D R H K Q V R Q N V H A T A **M**

RBS **-10** **-35**

8641 GCCTCCAATTAAGTTACATGCAAC**ATTGTA**CGCACCTCGCAAGTTAC**ATGCAA**CAGGTAT

8701 TTAAAGACTCATCCCCGGCGAGTTATCGCGCTCGATCTTCTGCGCCTTGGCCTGCTCCTG
* L S M G P S N D R E I K Q A K A Q E Q

8761 CTGGCGTACAGGCTCGCGCCCTTCCAGCTTCTCCCACTGCTGGCGGTACTTCCCAAGCTC
Q R V P E R G E L K E W Q Q R Y K G L E

8821 GGCGGCTTGGCGGTCGCCTGCGGCGCTGTAGGCATAGCGGGCGGTGCGCTGCGCGTAGTC

```
          A  A  Q  R  D  G  A  A  S  Y  A  Y  R  A  T  R  Q  A  Y  D

8881      GTTCCAGTCGCGGCCCTCGTCCTTCTGCCGGGTGGCCGGGGCGCACTGGCGGATAGCGCC
           N  W  D  R  G  E  D  K  Q  R  T  A  P  A  C  Q  R  I  A  G

8941      CTCGATGGCGGCTTGATCGTGGCCGGTGACGCGCATACGCACGGCAATCATGGAATCCAA
           E  I  A  A  Q  D  H  G  T  V  R  M  R  V  A  I  M  S  D  L

9001      GCGGGACAAGTCCACCTCGCCGCCACGCTGCCGCTTGAGCACGTCGCGATAATGCCGCTG
           R  S  L  D  V  E  G  G  R  Q  R  K  L  V  D  R  Y  H  R  Q

9061      GTAGGCGTCGATCGCGCTGCCGCTGGCCGCTGCAAGCTCCAAGGCGGGCTTGGCTTTGCT
           Y  A  D  I  A  S  G  S  A  A  A  L  E  L  A  P  K  A  K  S
              AvaI
9121      GCGCTCGGGCTGCTGCGCCTTCAAGGCCGCTTGCCGCTGATACTCGGCGTCGATCTGGCT
           R  E  P  Q  Q  A  K  L  A  A  Q  R  Q  Y  E  A  D  I  Q  S
                                                                AvaI
9181      GGACAAGGCCAGCGCCTTGATGCACTCCCGGCGCTCGGCCTTGAGCAAGCGCACCTCGGG
           S  L  A  L  A  K  I  C  E  R  R  E  A  K  L  L  R  V  E  P

9241      ATAGCTGCCGTCCTCCCGCTGGTGCTTGGGCTTGCGGTTCTCGTAGCCGGGTGCACGGTG
           Y  S  G  D  E  R  Q  H  K  P  K  R  N  E  Y  G  P  A  R  H

9301      CGGGTGGATGGCTCCCGATAGCTTCGGGTCGCCGTATTCACGGTTCAGGGCATCGCTCAG
           P  H  I  A  G  S  L  K  P  D  G  Y  E  R  N  L  A  D  S  L

9361      CCGGTTGCCCACGTCCTTATCGTGGGCCGTCCCCAGCTTCGGCACGGTGATGATGGCCTG
           R  N  G  V  D  K  D  H  A  T  G  L  K  P  V  T  I  I  A  Q

9421      ATAGTTGCCGGGGCTGGATTCAAGCACGGCGGCGGGCTGGTAGCCGTCGCGGATAAGCCG
           Y  N  G  P  S  S  E  L  V  A  A  P  Q  Y  G  D  R  I  L  R

9481      CTCCAGCTTCTCCCGGTTCATGTCGTCGATGAGGATGTGATGCTTCTTGTCCGATAGCGG
           E  L  K  E  R  N  M  D  D  I  L  I  H  H  K  K  D  S  L  P

9541      CGTGTAGTAGAGGTTTTCGCCCCGGCGCTGTAGGCGCTGCATCTCCGGCGTGCGCTGCTC
           T  Y  Y  L  N  E  G  R  R  Q  L  R  Q  M  E  P  T  R  Q  E

9601      GATCTCCTGCGGTGTGAAGCCCCGCGTGATGCCGTCCTTCTTGTCGAGGATGAAGGTTTG
           I  E  Q  P  T  F  G  R  T  I  G  D  K  K  D  L  I  F  T  Q

9661      CTTGCTGCCGTCCTCGCGCATCTTGATAGACGTTACCCGGTAACGCTCTGCGCCTACCGC
           K  S  G  D  E  R  M  K  I  S  T  V  R  Y  R  E  A  G  V  A
                                                 repB              RBS
9721      TTCCGCGTACCGCTCGAACTGCTTTAGCTGCTCCGATTTCATCGCCTGCACCCTTTCCTG
           E  A  Y  R  E  F  Q  K  L  Q  E  S  K  M  A  Q  V  R  E  Q

9781      CTGTATCCGCTGCCGCTCCTGCTGGATGCTCTCCTGAAGCTCTGGATTTGTGATCTTGAA
           Q  I  R  Q  R  E  Q  Q  I  S  E  Q  L  E  P  N  T  I  K  F

9841      GCCGTGCTCTGCTGCTAACTTCCCGCACATGGCCTTGTATTCGTCGTTCCCCATCACGGT
           G  H  E  A  A  L  K  G  C  M  A  K  Y  E  D  N  G  M  V  T

9901      GAAGCTGCCCCACTTCTGCGCCGATAGCTGTAACGCGGCAAGCGTGCTGTCCCGGTTGCG
           F  S  G  W  K  Q  A  S  L  Q  L  A  A  L  T  S  D  R  N  R

9961      CCAGTCGTGAATATCGATGCTTTTGCCCTTGTCCACGAAAGACACACCCCCGCCCCGCCC
```

```
          W  D  H  I  D  I  S  K  G  K  D  V  F  S  V  G  G  G  R  G

10021     CGCTTCTTCCTTCCGGCTGTAATGCACCTGCTGGCCCACGATTTCAGGCTGATAGGCGCG
           A  E  E  K  R  S  Y  H  V  Q  Q  G  V  I  E  P  Q  Y  A  R

10081     AATGTCTCGCGGCGTCGGCGGCTCGCTGCGGTCGCCCTCGATGCGCTGCGGCTCGCTGGC
           I  D  R  P  T  P  P  E  S  R  D  G  E  I  R  Q  P  E  S  A

10141     CCGGTGTCGCCACTGCTCGGCAAGCTCCGGGCTTTTCTGCATCCGCTGCCACTGCTCCAG
           R  H  R  W  Q  E  A  L  E  P  S  K  Q  M  R  Q  W  Q  E  L

10201     GTCGGGATAGGGTCGGAACCGCTGCCGGTGCTGCTCGCGCTCCTTCTGGTGCTTCTCCTT
           D  P  Y  P  R  F  R  Q  R  H  Q  E  R  E  K  Q  H  K  E  K

10261     CAAGGCCGCTTTCTCTGCGGCCTGCTCGGCGGCAATCACGCTGCGCATGGCGTTGAGCAC
           L  A  A  K  E  A  A  Q  E  A  A  I  V  S  R  M  A  N  L  V

10321     TTCGCCCTTGCCTTTCCAGTTTCCGCGCATGAGTTCATCGCGGCGGGCCTTCTGCTGCTC
           E  G  K  G  K  W  N  G  R  M  L  E  D  R  R  A  K  Q  Q  E

10381     GGCAAGTTGCTTGCGCTCCTGCTCCTGCCGCTTGTCCTGTGCCAGCTTGTCCGCGTTCTT
           A  L  Q  K  R  E  Q  E  Q  R  K  D  Q  A  L  K  D  A  N  K

10441     CTCGGCGTAGTGTGCCTTGCGCCCGGTGATGTAGTCCTTCCATCCCGGCACGTCTGGCTT
           E  A  Y  H  A  K  R  G  T  I  Y  D  K  W  G  P  V  D  P  K

10501     GATCGGCTCCGGCTCGCGCGGCGCTACCTGCTGACGCTGCGGCGCTGGCTGGTATTCACC
           I  P  E  P  E  R  P  A  V  Q  Q  R  Q  P  A  P  Q  Y  E  G

10561     TAGCCGCTTCTGCATCTTGGCAAGGCTGGCGTTGCGGTCTACATCGCTGGCCTTCACGCC
           L  R  K  Q  M  K  A  L  S  A  N  R  D  V  D  S  A  K  V  G

10621     AACATCCCCGACAAACACGGTCGCCCCGCTGCCGGTCTTTTCGTAGCGCATCCCCTTGGC
           V  D  G  V  F  V  T  A  G  S  G  T  K  E  Y  R  M  G  K  A

10681     TGCTAACTCGCGGTGCAACTGCTCCCACGTCTGCGCCTGCTTGATGATGGCTGCGCCGTC
           A  L  E  R  H  L  Q  E  W  T  Q  A  Q  K  I  I  A  A  G  D

10741     CTCGATGGCGATGCGGTGGGCTGACTTCTCCCCGGTGCGGTTCTCCATGTCCCGCTTCTT
           E  I  A  I  R  H  A  S  K  E  G  T  R  N  E  M  D  R  K  K

10801     CTGGTCGGGCTGGCGTGGCTTTTCCGGGTCGTATGGCGCTCGGCCTAGCTCCCCGTTCTC
           Q  D  P  Q  R  P  K  E  P  D  Y  P  A  R  G  L  E  G  N  E

10861     CAACACCTGATAGCGCCCGTTTTGCTCACGCTGCCAGCCTTGGGCGTGCTCAATCCGTGC
           L  V  Q  Y  R  G  N  Q  E  R  Q  W  G  Q  A  H  E  I  R  A

10921     AATGGCTTTGTGTGCCAGTTCAATATCAAATCCACGGTTCTTCTCAACAATCTTGAGCGT
           I  A  K  H  A  L  E  I  D  F  G  R  N  K  E  V  I  K  L  T

10981     TTCGGGATGCACCCGATTGATGACGATGTGCAGGTGGATATTGTCCGTGTCGGAATGAAG
           E  P  H  V  R  N  I  V  I  H  L  H  I  N  D  T  D  S  H  L

11041     GCCATAGATGGCTTGGTGGTCTTTCCACCCAAGCTCATCCATGAAGATGCTAACCGCTTC
           G  Y  I  A  Q  H  D  K  W  G  L  E  D  M  F  I  S  V  A  E

11101     CTCTACTTGCTCCGGGCTTGGCTGTTCGCCTTCGCGCCAACTCAAGACGTAGTGATTGAT
```

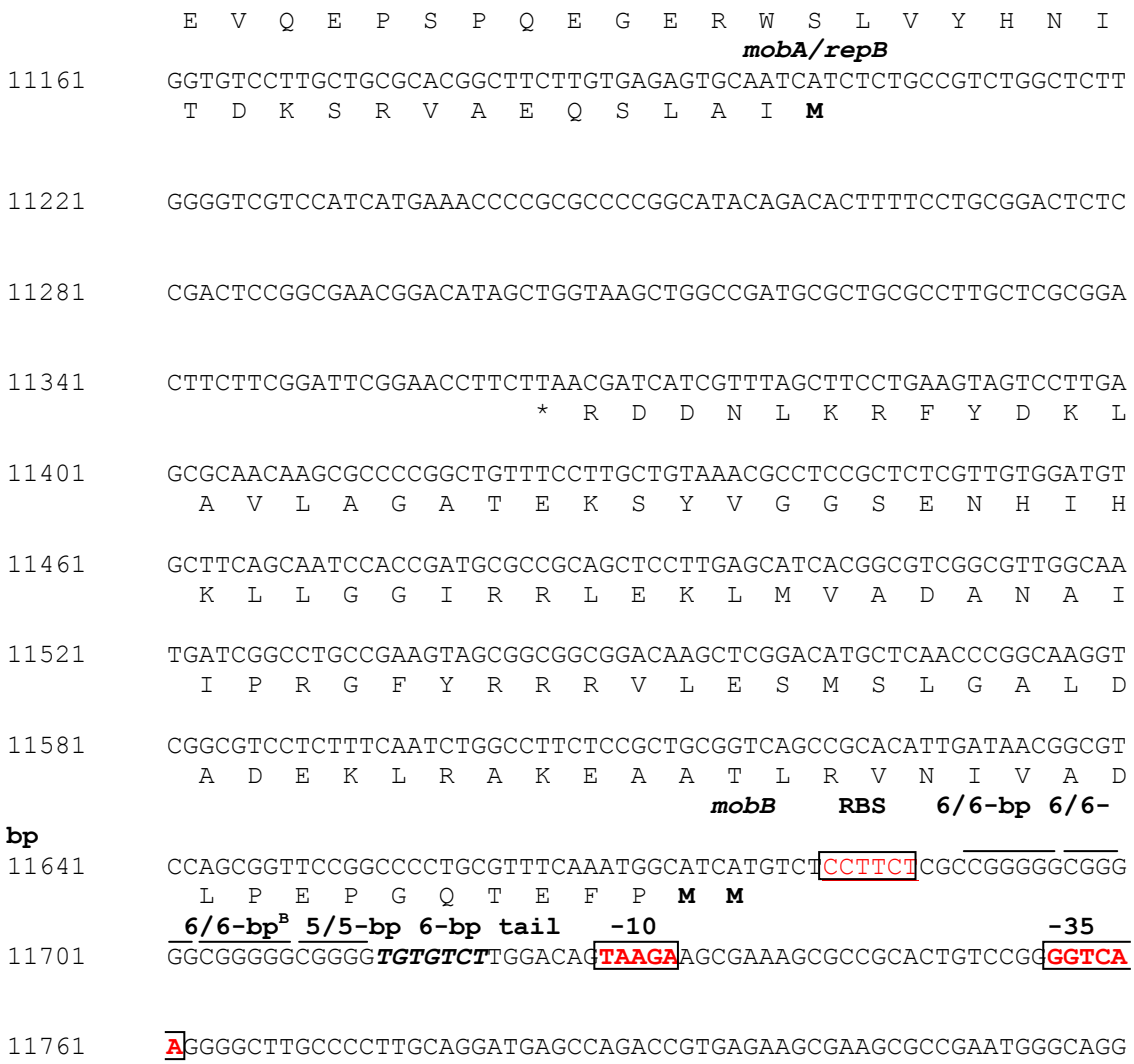


FIG. E: Annotated sequence of pRAS3.2 and includes the mutational differences that has given rise to pRAS3.1. The nucleotides at which pRAS3.1 differs from pRAS3.2 are indicated above the pRAS3.2 nucleotide sequence in brackets. All the promoter regions and transcriptional start sites, annotated sequences within the *oriV* other than the iterons, and *nic* site within the *oriT* are putative. The position of the additional iteron and 6-bp repeat of pRAS3.1 are indicated by [A] and [B] respectively. The nucleotide sequences of pRAS3.1 (accession number AY043298.1) and pRAS3.2 (accession number NC_003124.1) were obtained from the GenBank database.

REFERENCES

Allen, S. P. and H. P. Blaschek. 1990. Plasmid copy number and stability determination in *Clostridium perfringens* transformants. FEMS Microbiol. Lett. **72**:323-327.

Aoki, T. and A. Takahashi. 1986. Tetracycline-resistant gene of a non-transferable R plasmid from fish-pathogenic bacteria *Aeromonas salmonicida*. Bull. Jpn. Soc. Scient. Fish. **52**:1913-1917.

Albritton, W. L., J. L. Brunton, L. Slaney, and I. MacLean. 1982. Plasmid-mediated sulfonamide resistance in *Haemophilus ducreyi*. Antimicrob. Agents Chemother. **21**:159-165.

Amitai, S., Y. Yassin, and H. Engelberg-Kulka. 2004. MazF-mediated cell death in *Escherichia coli*: a point of no return. J. Bacteriol. **186**:8295-8300.

Andersson, D. I. and B. R. Levin. 1999. The biological cost of antibiotic resistance. Curr. Opin. Microbiol. **2**:489-493.

Applebee, M. K., M. J. Herrgard, and B. O. Palsson. 2008. Impact of Individual Mutations on Increased Fitness in Adaptively-evolved strains of *Escherichia coli*. J. Bacteriol. **190**:5087-5094.

Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl. 1993. Current protocols in molecular biology. Wiley Interscience, New York.

0Bahassi, E. M, M.H. O'Dea, N. Allali, J. Messens, M. Gellert, M. Couturier. 1999. Interactions of CcdB with DNA gyrase. Inactivation of Gyra, poisoning of the gyrase-DNA complex, and the antidote action of CcdA. J. Biol. Chem. **274**:10936-10944.

Banerjee, S. K., B. T. Luck, H. Y. Kim, and V. N. Iyer. 1992. Three clustered origins of replication in a promiscuous-plasmid replicon and their differential use in a PolA+ strain and a delta PolA strain of Escherichia coli K-12. J. Bacteriol. **174**:8139-8143.

Barth, P. T. and N. J. Grinter. 1974. Comparison of the deoxyribonucleic acid molecular weights and homologies of plasmids conferring linked resistance to streptomycin and sulfonamides. J. Bacteriol. **120**:618-630.

Becker, E. C. and R. J. Meyer. 1997. Acquisition of resistance genes by the IncQ plasmid R1162 is limited by its high copy number and lack of a partitioning mechanism. J. Bacteriol. **179**:5947-5950.

Becker, E. C. and R. J. Meyer. 2002. MobA, the DNA strand transferase of plasmid R1162. J. Biol. Chem. **277**:14575-14580.

Becker, E. C. and R. J. Meyer. 2003. Relaxed specificity of the R1162 nickase: a model for evolution of a system for conjugative mobilization of plasmids. J. Bacteriol. **185**:3538-3546.

Bergstrom, C. T., M. Lipsitch, and B. R. Levin. 2000. Natural selection, infectious transfer and the existence conditions for bacterial plasmids. Genetics **155**:1505-1519.

Bhattacharjee, M. K. and R. J. Meyer. 1991. A segment of a plasmid gene required for conjugal transfer encodes a site-specific, single-strand DNA endonuclease and ligase. Nucl. Acids Res. **19**:1129-1137.

Bhattacharjee, M., X. M. Rao, and R. J. Meyer. 1992. Role of the origin of transfer in termination of strand transfer during bacterial conjugation. J. Bacteriol. **174**:6659-6665.

Bignell, C. and C. M. Thomas. 2001. The bacterial ParA-ParB partitioning proteins. J. Biotechnol. **91**:1-34.

Bingle, L. E. H. and C. M. Thomas. 2001. Regulatory circuits for plasmid survival. Curr. Opin. Microbiol. **4**:194-200.

Bjellqvist, B., G. J. Hughes, Ch. Pasquali, N. Paquet, F. Ravier, J.-C. Sanchez, S. Frutiger, and D. F. Hochstrasser. 1993. The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences. Electrophoresis **14**:1023-1031.

Bjellqvist, B., B. Basse, E. Olsen, and J. E. Celis. 1994. Reference points for comparisons of two-dimensional maps of proteins from different human cell types defined in a pH scale where isoelectric points correlate with polypeptide compositions. Electrophoresis **15**:529-539.

Boe, L., K. Gerdes, and S. Molin. 1987. Effects of genes exerting growth inhibition and plasmid stability on plasmid maintenance. J. Bacteriol. **169**:4646-4650.

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow. 1977. Construction and characterization of new cloning vehicle. II. A multipurpose cloning system. Gene **2**:95-113.

Bönemann, G., M. Stiens, A. Pühler, and A. Schlüter. 2006. Mobilizable IncQ-related plasmid carrying a new quinolone resistance gene, *qnrS2*, isolated from the bacterial community of a wastewater treatment plant. Antimicrob. Agents Chemother. **50**:3075-3080.

Bouma, J. and R. Lenski. 1988. Evolution of a bacteria/plasmid association. Nature **335**:351-352.

Bowers, L. M., R. Krüger, and M. Filutowicz. 2007. Mechanism of origin activation by monomers of R6K-encoded pi protein. J. Mol. Biol. **368**:928-938.

Bowers, L. M. and M. Filutowicz. 2008. Cooperative binding mode of the inhibitors of R6K Rreplication, pi dimers. J. Mol. Biol. **377**:609-615.

Brasch, M. A. and R. J. Meyer. 1987. A 38 base-pair segment of DNA is required in cis for conjugative mobilization of broad host-range plasmid R1162. J. Mol. Biol. **198**:361-369.

Bravo, A., G. de Torrontegui, and R. Diaz. 1987. Identification of components of a new stability system of plasmid R1, ParD, that is close to the origin of replication of this plasmid. Mol. Gen. Gen. **210**:101-110.

Brenowitz, M., D. F. Senear, and R. E. Kingston. 2001. DNase I footprint analysis of protein-DNA binding, Unit 12.4. *In* F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, and J. A. Smith (ed.), Current Protocols in Molecular Biology. John Wiley & Sons, Inc. New Jersey.

Brown, A. M. C. and N. S. Willetts. 1981. A physical and genetic map of the IncN plasmid R46. Plasmid **5**:188-201.

Bryan, L. E., H. M. V. Elzen, and J. T. Tseng. 1972. Transferable drug resistance in *Pseudomonas aeruginosa*. Antimicrob. Agents Chemother. **1**:22-29.

Bustin, S. A., V. Benes, J. A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M. W. Pfaffl, G. L. Shipley, J. Vandesompele, and C. T. Wittwer. 2009. The MIQE guidelines:

Minimum Information for Publication of Quantitative Real-Time PCR Experiments. Clin Chem **55**:611-622.

Cabezón, E., E. Lanka, and F. de la Cruz. 1994. Requirements for mobilization of plasmids RSF1010 and ColE1 by the IncW plasmid R388: *trwB* and RP4 *traG* are interchangeable. J. Bacteriol. **176**:4455-4458.

Carlton, C. B. and B. J. Brown. 1981. Gene mutation, p. 222-242. *In* P. Gerhardt (ed.), Maunual of methods for general bacteriology. American Society for Microbiology, Washington, D.C.

Cascales, E., K. Atmakuri, Z. Liu, A. N. Binns, and P. J. Christie. 2005. *Agrobacterium tumefaciens* oncogenic suppressors inhibited T-DNA and VirE2 protein substrate binding to the VirD4 coupling protein. Mol. Microbiol. **58**:565-579.

Cesareni, G., M. Helmer-Citterich, and L. Castagnoli. 1991. Control of ColE1 plasmid replication by antisense RNA. Trends Genet. **7**:230-235.

Chang, A. C. and S. N. Cohen . 1978. Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. J. Bacteriol. **134**:1141-1156.

Chattoraj, D., K. Cordes, and A. Abeles. 1984. Plasmid P1 replication: Negative control by repeated DNA sequences. Proc. Natl Acad. Sci. U.S.A. **81**:6456-6460.

Chattoraj, D. K., K. M. Snyder, and A. L. Abeles. 1985. P1 plasmid replication: Multiple functions of RepA protein at the origin. Proc Proc. Natl Acad. Sci. U.S.A. **82**:2588-2592.

Chattoraj, D. K. 2000. Control of plasmid DNA replication by iterons: No longer paradoxical. Mol. Microbiol. **37**:467-476.

Chen, L., Y. Chen, D. W. Wood, and E. W. Nester. 2002. A new type IV secretion system promotes conjugal transfer in *Agrobacterium tumefaciens*. J. Bacteriol. **184**:4838-4845.

Clennel, A. M., B. Johnston, and D. E. Rawlings. 1995. Structure and function of Tn5467, a Tn21-like transposon located on the *Thiobacillus ferrooxidans* broad-host-range plasmid pTF-FC2. Appl. Environ. Microbiol. **61**:4223-4229.

Cooper, T. F. and J. A. Heinemann. 2000. Postsegregational killing does not increase plasmid stability but acts to mediate the exclusion of competing plasmids. Proc. Natl Acad. Sci. U.S.A. **97**:12643-12648.

Crosa, J. H. 1980. Three origins of replication are active in vivo in the R plasmid RSF1040. J. Biol. Chem. **255**:11075-11077.

Dahlberg, C. and L. Chao. 2003. Amelioration of the cost of conjugative plasmid carriage in *Eschericha coli* K12. Genetics **165**:1641-1649.

Daly, M., L. Villa, C. Pezzella, S. Fanning, and A. Carattoli. 2005. Comparison of multidrug resistance gene regions between two geographically unrelated *Salmonella* serotypes. J. Antimicrob. Chemother. **55**:558-561.

Das, N., M. Valjavec-Gratian, A. N. Basuray, R. A. Fekete, P. P. Papp, J. Paulsson, and D. K. Chattoraj. 2005. Multiple homeostatic mechanisms in the control of P1 plasmid replication. Proc. Natl Acad. Sci. U.S.A. **102**:2856-2861.

Davis, B. D. and E. S. Mingioli. 1950. Mutants of *Escherichia coli* requiring methionine or vitamin B12. J. Bacteriol. **60**:17-28.

Dawkins, R. 1976. The selfish gene. Oxford University Press, Oxford.

De Castro, E., C. J. A. Sigrist, A. Gattiker, V. Bulliard, S. Petra, P. S. Langendijk-Genevaux, E. Gasteiger, A. Bairoch, and N. Hulo. 2006. ScanProsite: detection of PROSITE signature matches and ProRule-associated functional and structural residues in proteins. Nucl. Acids Res. **34**:W362-W365.

De Gelder, L., J. M. Ponciano, P. Joyce, and E. M. Top. 2007. Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. Microbiology **153**:452-463.

De Gelder, L., J. J. Williams, J. M. Ponciano, M. Sota, and E. M. Top. 2008. Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. Genetics **178**:2179-2190.

De Graaff, J., J. H. Crosa, F. Heffron, and S. Falkow. 1978. Replication of the nonconjugative plasmid RSF1010 *in Escherichia coli* K-12. J. Bacteriol. **134**:1117-1122.

Deane, S. M. and D. E. Rawlings. 2004. Plasmid evolution and interaction between the plasmid addiction stability systems of two related broad-host-range IncQ-like plasmids. J. Bacteriol. **186**:2123-2133.

Del Solar, G. and M. Espinosa. 2000. Plasmid copy number control: an ever-growing story. Mol. Microbiol. **37**:492-500.

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**:434-464.

Derbyshire, K. M. and N. S. Willetts. 1987. Mobilization of the non-conjugative plasmid RSF1010: A genetic analysis of its origin of transfer. Mol. Gen. Gen. **206**:154-160.

Dionisio, F., I. C. Conceicão, A. C. R. Marques, L. Fernandes, and I. Gordo. 2005. The evolution of a conjugative plasmid and its ability to increase bacterial fitness. Biol. Lett. **1**:250-252.

Dorrington, R. A. and D. E. Rawlings. 1989. Identification and sequence of the basic replication region of a broad-host-range plasmid isolated from *Thiobacillus ferrooxidans*. J. Bacteriol. **171**:2735-2739.

Dorrington, R. A. and D. E. Rawlings. 1990. Characterization of the minimum replicon of the broad-host-range plasmid pTF-FC2 and similarity between pTF-FC2 and the IncQ plasmids. J. Bacteriol. **172**:5697-5705.

Dugan, J., D. D. Rockey, L. Jones, and A. A. Andersen. 2004. Tetracycline resistance in *Chlamydia suis* mediated by genomic islands inserted into the Chlamydial inv-like gene. Antimicrob. Agents Chemother. **48**:3989-3995.

Engberg, B. and K. Nordström. 1975. Replication of R-factor R1 in Escherichia coli K-12 at different growth rates. J. Bacteriol. **123**:179-186.

Enne, V. I., A. A. Delsol, J. M. Roe, and P. M. Bennett. 2004. Rifampicin resistance and its fitness cost in *Enterococcus faecium*. J. Antimicrob. Chemother. **53** :203-207.

Fernández-López, R., M. P. Garcillán-Barcia, C. Revilla, M. Lázaro, L. Vielva, and F. de la Cruz. 2006. Dynamics of the IncW genetic backbone imply general trends in conjugative plasmid evolution. FEMS Microbiol. Rev. **30**:942-966.

Fling, M. E., J. Kope, and C. Richards. 1988. Characterization of plasmid pAZ1 and the type III dihydrofolate reductase gene. Plasmid **19**:30-38.

Fong, S. S., A. R. Joyce, and B. Ø. Palsson. 2005. Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states. Genome Res. **15**:1365-1372.

Fozo, E. M., M. R. Hemm, and G. Storz. 2008. Small toxic proteins and the antisense RNAs that repress them. Microbiol. Mol. Biol. Rev. **72**:579-589.

Frame, R. and J. O. Boshop . 1971. The number of sex-factors per chromosome in *Escherichia coli*. Biochem. J. **121**:93-103.

Francia, M. V., A. Varsaki, M. P. Garcillán-Barcia, A. Latorre, C. Drainas, and F. de la Cruz. 2004. A classification scheme for mobilization regions of bacterial plasmids. FEMS Microbiol. Rev. **28**:79-100.

Frey, J. and M. Bagdasarian . 1989. The molecular biology of IncQ plasmids, p. 79-93. *In* C. M. Thomas (ed.), Promiscuous Plasmids of Gram Negative Bacteria. Academic Press, London.

Frey, J., M. M. Bagdasarian, and M. Bagdasarian. 1992. Replication and copy number control of the broad-host-range plasmid RSF1010. Gene **113**:101-106.

Gardner, M. N., S. M. Deane, and D. E. Rawlings. 2001. Isolation of a new broad-host-range IncQ-like plasmid, pTC-F14, from the acidophilic bacterium *Acidithiobacillus caldus* and analysis of the plasmid replicon. J. Bacteriol. **183**:3303-3309.

Gardner, M. N. 2003. An investigation into the replicon of a broad host range mobilizable plasmid from the moderately thermophilic bacterium *Acidithiobacillus caldus*. Thesis (PhD). University of Stellenbosch.

Gardner, M. N. and D. E. Rawlings. 2004. Evolution of compatible replicons of the related IncQ-like plasmids, pTC-F14 and pTF-FC2. Microbiol. **150**:1797-1808.

Gasset-Rosa, F., T. Diaz-López, R. Lurz, A. Prieto, M. E. Fernández-Tresguerres, and R. Giraldo. 2008. Negative regulation of pPS10 plasmid replication: origin pairing by zipping-up DNA-bound RepA monomers. Mol. Microbiol. **68**:560-572.

Gasteiger, E., C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel, and A. Bairoch. 2005. Protein identification and analysis tools on the ExPASy server , p. 571-608. *In* J. M. Walker (ed.), The Proteomics Protocols Handbook. Humana Press, New Jersey.

Geibel, S., S. Banchenko, M. Engel, E. Lanka, and W. Saenger. 2009. Structure and function of primase RepB' encoded by broad-host-range plasmid RSF1010 that replicates exclusively in leading-strand mode. Proc. Natl. Acad. Sci. U.S.A. **106**:7810-7815.

Gerdes, K., J. E. Larsen, and S. Molin. 1985. Stable inheritance of plasmid R1 requires two different loci. J. Bacteriol. **161**:292-298.

Gerdes, K., P. B. Rasmussen, and S. Molin. 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proc. Natl Acad. Sci. U.S.A. **83**:3116-3120.

Gerdes, K., S. K. Christensen, and A. Løbner-Olesen. 2005. Prokaryotic toxin-antitoxin stress response loci. Nat. Rev. Microbiol. **3**:371-382.

Germino, J. and D. Bastia. 1983. Interaction of the plasmid R6K-encoded replication initiator protein with its binding sites on DNA. Cell **34**:125-134.

Glick, B. R. 1995. Metabolic load and heterologous gene expression. Biotechnol. Adv. **13**:247-261.

Götz, A., R. Pukall, E. Smit, E. Tietze, R. Prager, H. Tschape, J. D. van Elsas, and K. Smalla. 1996. Detection and characterization of broad-host-range plasmids in environmental bacteria by PCR. Appl. Environ. Microbiol. **62**:2621-2628.

Gormley, E. P. and J. Davies. 1991. Transfer of plasmid RSF1010 by conjugation from *Escherichia coli* to *Streptomyces lividans* and *Mycobacterium smegmatis*. J. Bacteriol. **173**:6705-6708.

Grinter, N. J. and P. T. Barth. 1976. Characterization of SmSu plasmids by restriction endonuclease cleavage and compatibility testing. J. Bacteriol. **128**:394-400.

Guerry, P., J. Van Embden, and S. Falkow. 1974. Molecular nature of two nonconjugative plasmids carrying drug resistance genes. J. Bacteriol. **117**:619-630.

Guzman, L. M., D. Belin, M. J. Carson, and J. Beckwith. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J. Bacteriol. **177**:4121-4130.

Haines, A. S., K. Jones, M. Cheung, and C. M. Thomas. 2005. The IncP-6 Plasmid Rms149 consists of a small mobilizable backbone with multiple large insertions. J. Bacteriol. **187**:4728-4738.

Hampshire, A. J., D. A. Rusling, V. J. Broughton-Head, and K. R. Fox. 2007. Footprinting: A method for determining the sequence selectivity, affinity and kinetics of DNA-binding ligands. Methods **42**:128-140.

Haring, V., P. Scholz, E. Scherzinger, J. Frey, K. Derbyshire, G. Hatfull, N. S. Willetts, and M. Bagdasarian. 1985. Protein RepC is involved in copy number control of the broad host range plasmid RSF1010. Proc. Natl Acad. Sci. U.S.A. **82**:6090-6094.

Hazan, R., B. Sat, M. Reches, and H. Engelberg-Kulka. 2001. Postsegregational killing mediated by the P1 phage "addiction module" *phd-doc* requires the *Escherichia coli* programmed cell death system *mazEF*. J. Bacteriol. **183**:2046-2050.

Henderson, D. and R. J. Meyer. 1996. The primase of broad-host-range plasmid R1162 is active in conjugal transfer. J. Bacteriol. **178**:6888-6894.

Henderson, D. and R. Meyer . 1999. The MobA-linked primase is the only replication protein of R1162 required for conjugal mobilization. J. Bacteriol. **181**:2973-2978.

Honda, Y., H. Sakai, H. Hiasa, K. Tanaka, T. Komano, and M. Bagdasarian. 1991. Functional division and reconstruction of a plasmid replication origin: Molecular dissection of the oriV of the broad-host-range plasmid RSF1010. P. Natl Acad. Sci. U. S. A. **88**:179-183.

Honda, Y., T. Nakamura, K. Tanaka, A. Higashi, H. Sakai, T. Komano, and M. Bagdasarian. 1992. DnaG-dependent priming signals can substitute for the two essential DNA initiation signals in *oriV* of the broad host-range plasmid RSF1010. Nucl. Acids Res. **20**:1733-1737.

Honda, Y., T. Akioka, S. Takebe, K. Tanaka, D. Miao, A. Higashi, T. Nakamura, Y. Taguchi, H. Sakai, T. Komano, and M. Bagdasarian. 1993. Mutational analysis of the specific priming signal essential for DNA replication of the broad host-range plasmid RSF1010. FEBS Lett. **324**:67-70.

Huggett, J., K. Dheda, S. Bustin, and A. Zumla. 2005. Real-time RT-PCR normalisation; strategies and considerations. Genes Immun .**6**:279-284.

Hulo, N., A. Bairoch, V. Bulliard, L. Cerutte, E. De Castro, P. S. Langendijk-Genevaux, M. Pagni, and S. J. A. Sigrist. 2006. The PROSITE database. Nucl. Acids Res. **34**:D227-D230.

Ito, H., H. Ishii, and M. Akiba. 2004. Analysis of the complete nucleotide sequence of an *Actinobacillus pleuropneumoniae* streptomycin-sulfonamide resistance plasmid, pMS260. Plasmid **51**:41-47.

Jiang, Y., J. Pogliano, D. R. Helinski, and I. Konieczny. 2002. ParE toxin encoded by the broad-host-range plasmid RK2 is an inhibitor of *Escherichia coli* gyrase. Mol. Microbiol. **44**:971-979.

Jonassen, I., J. F. Collins, and D. Higgins. 1995. Finding flexible patterns in unaligned protein sequences. Protein Sci. **4**:1587-1595.

Jonassen, I. 1997. Efficient discovery of conserved patterns using a pattern graph. Comput. Appl. Biosci. **13**:509-522.

Katashkina, J., T. Kuvaeva, I. Andreeva, A. Skorokhodova, I. Biryukova, I. Tokmakova, L. Golubeva, and S. Mashko. 2007. Construction of stably maintained non-mobilizable derivatives of RSF1010 lacking all known elements essential for mobilization. BMC Biotechnol. **7**:80.

Kim, K. and R. J. Meyer. 1985. Copy number of the broad host-range plasmid R1162 is determined by the amounts of essential plasmid-encoded proteins. J. Mol. Biol. **185**:755-767.

Kim, K. and R. J. Meyer. 1986. Copy-number of broad host-range plasmid R1162 is regulated by a small RNA. Nucl. Acids Res. **14**:8027-8046.

Kim, Y. J. and R. J. Meyer . 1991. An essential iteron-binding protein required for plasmid R1162 replication induces localized melting within the origin at a specific site in AT-rich DNA. J. Bacteriol. **173**:5539-5545.

Kittell, B. L. and D. R. Helinski. 1991. Iteron inhibition of plasmid RK2 replication in vitro: Evidence for intermolecular coupling of replication origins as a mechanism for RK2 replication control. Proc. Natl. Acad. Sci. U.S.A. **88**:1389-1393.

Kok, M., A. C. Arnberg, and B. Witholt. 1989. Single-stranded circular DNA generated from broad host range plasmid R1162 and its derivatives. Plasmid **21**:238-241.

Krakauer, D. C. and J. B. Plotkin. 2002. Redundancy, antiredundancy, and the robustness of genomes. Proc. Natl. Acad. Sci. U.S.A. **99**:1405-1409.

Krone, S. M., R. Lu, R. Fox, H. Suzuki, and E. M. Top. 2007. Modelling the spatial dynamics of plasmid transfer and persistence. Microbiology **153**:2803-2816.

Krüger, R., S. A. Rakowski, and M. Filutowicz. 2004. Participating elements in the replication of iteron-containing plasmids, p. 25-46. *In* B. E. Funnell and G. J. Phillips (ed.), Plasmid Biology. ASM Press, Washington, DC.

L'Abee-Lund, T. M. and H. Sørum. 2002. A global non-conjugative Tet C plasmid, pRAS3, from *Aeromonas salmonicida*. Plasmid **47**:172-181.

Lawley, T., B. M. Wilkins, and L. S. Frost. 2004. Bacterial conjugation in Gram-negative bacteria, p. 203-226. *In* G. Phillips and B. E. Funnell (ed.), Plasmid Biology. ASM Press, Washington, DC.

Lessl, M., D. Balzer, W. Pansegrau, and E. Lanka. 1992. Sequence similarities between the RP4 Tra2 and the Ti VirB region strongly support the conjugation model for T-DNA transfer. J. Biol. Chem. **267**:20471-20480.

Lessl, M., D. Balzer, K. Weyrauch, and E. Lanka. 1993. The mating pair formation system of plasmid RP4 defined by RSF1010 mobilization and donor-specific phage propagation. J. Bacteriol. **175**:6415-6425.

Lin-Chao, S., W. Chen, and T. Wong. 1992. High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. Mol. Microb. **6**:3385-3393.

Lin, L. S. and R. J. Meyer. 1986. Directly repeated, 20-bp sequence of plasmid R1162 DNA is required for replication, expression of incompatibility, and copy-number control. Plasmid **15**:35-47.

Lin, L. S., Y. J. Kim, and R. J. Meyer. 1987. The 20 bp, directly repeated DNA sequence of broad host range plasmid R1162 exerts incompatibility *in vivo* and inhibits R1162 DNA replication *in vitro*. Mol. Gen. Gen. **208**:390-397.

Lisser, S. and H. Margalit . 1993. Compilation of *E. coli* mRNA promoter sequences. Nucl. Acids Res. **21**:1507-1516.

Liu, M., Y. Zhang, M. Inouye, and N. Woychik. 2008. Bacterial addiction module toxin Doc inhibits translation elengation through its association with the 30S ribosomal subunit. Proc. Natl. Acad. Sci. U.S.A. **105**:5885-5890.

Lee, C., J. Kim, S. G. Shin, and S. Hwang. 2006a. Absolute and relative qPCR quantification of plasmid copy number in *Escherichia coli*. J. Biotechnol. **123**:273-280.

Lee, C. L., D. S. Ow, and S. K. Oh. 2006b. Quantitative real-time polymerase chain reaction for determination of plasmid copy number in bacteria. J. Microbiol. Meth. **65**:258-267.

Lenski, R. E. and J. E. Bouma. 1987. Effects of segregation and selection on instability of plasmid pACYC184 in *Escherichia coli* B. J. Bacteriol. **169**:5314-5316.

Lenski, R. E. 1992. Relative fitness: Its estimation and its significance for environmental applications of microorganisms, p. 183-198. *In* M. A. Levin, R. J. Seidler, and M. Rogul (ed.), Microbial ecology: Principles, Methods, and Applications. McGraw-Hill, New York.

Lenski, R. E., S. C. Simpson, and T. T. Nguyen. 1994. Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. J. Bacteriol. **176**:3140-3147.

Lessl, M., D. Balzer, R. Lurz, V. L. Waters, D. G. Guiney, and E. Lanka. 1992. Dissection of IncP conjugative plasmid transfer: definition of the transfer region Tra2 by mobilization of the Tra1 region *in trans*. J. Bacteriol. **174**:2493-2500.

Lewington, J. and M. J. Day. 1986. A rapid electrophoretic method for the measurement of plasmid copy number. Lett. Appl. Microbiol. **3**:109-112.

Logan, J., K. Edwards, and N. Saunders. 2009. *Real-Time PCR : Current technology and applications*. Caister Academic Press. Norfolk.

Masuda, Y., K. Miyakawa, Y. Nishimura, and E. Ohtsubo. 1993. *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. J. Bacteriol. **175**:6850-6856.

Maeser, S., P. Scholz, S. Otto, and E. Scherzinger. 1990. Gene F of plasmid RSF1010 codes for a low-molecular-weight repressor protein that autoregulates expression of the *repAC* operon. Nucl. Acids Res. **18**:6215-6222.

Matcher, G. F. and D. E. Rawlings. 2009. The effect of the location of the proteic post-segregational stability system within the replicon of plasmid pTF-FC2 on the fine regulation of plasmid replication. Plasmid **62**:98-107.

Masai, H. and K. Arai. 1989. Leading strand synthesis of R1 plasmid replication *in vitro* is primed by primase alone at a specific site downstream of *oriR*. J. Biol. Chem. **264**:8082-8090.

McEachern, M. J., M. Filutowicz, and D. R. Helinski. 1985. Mutations in direct repeat sequences and in a conserved sequence adjacent to the repeats result in a defective replication origin in plasmid R6K. Proc. Natl Acad. Sci. U.S.A. **82**:1480-1484.

Mendoza-Vargas, A., L. Olvera, M. Olvera, R. Grande, L. Vega-Alvarado, B. Taboada, V. Jimenez-Jacinto, H. Salgado, K. Juáírez, B. Contreras-Moreira, A. M. Huerta, J. Collado-Vides, and E. Morett. 2009. Genome-wide identification of transcription start sites, promoters and transcription factor binding sites in *E. coli*. PLoS ONE **4**:e7526.

Meyer, R., M. Hinds, and M. Brasch. 1982. Properties of R1162, a broad-host-range, high-copy-number plasmid. J. Bacteriol. **150**:552-562.

Meyer, R. 2009. Replication and conjugative mobilization of broad host-range IncQ plasmids. Plasmid **62**:57-70.

Miao, D. M., Y. Honda, K. Tanaka, A. Higashi, T. Nakamura, Y. Taguchi, H. Sakai, T. Komano, and M. Bagdasarian. 1993. A base-paired hairpin structure essential for the functional priming signal for DNA replication of the broad host range plasmid RSF1010. Nucl. Acids Res. **21**:4900-4903.

Miao, D. M., H. Sakai, S. Okamoto, K. Tanaka, M. Okuda, Y. Honda, T. Komano, and M. Bagdasarian. 1995. The interaction of RepC initiator with iterons in the replication of the broad host-range plasmid RSF1010. Nucl. Acids Res. **23**:3295-3300.

Modi, R. I. and J. Adams. 1991. Coevolution in bacterial-plasmid populations. Evolution **45**:656-667.

Needleman, D. J. 2008. Plasmid segregation: Is a total understanding within reach? Curr. Biol. **18**:R212-R214.

Niedenzu, T., D. Röleke, G. Bains, E. Scherzinger, and W. Saenger. 2001. Crystal structure of the hexameric replicative helicase RepA of plasmid RSF1010. Journal of Molecular Biology **306**:479-487.

Nordström, K. and S. J. Austin. 1989. Mechanisms that contribute to the stable segregation of plasmids. Annu. Rev. Genet. **23**:37-69.

Norman, A., H. H. Lars, and S. J. Sørensen. 2009. Conjugative plasmids: Vessels of the communal gene pool. Phil. Trans. R. Soc. B. **364**:2275-2289.

Novick, R. P. 1987. Plasmid incompatibility. Microbiol. Mol. Biol. Rev. **51**:381-395.

Pal, S. K., R. J. Mason, and D. K. Chattoraj. 1986. P1 plasmid replication: Role of initiator titration in copy number control. J. Mol. Biol. **192**:275-285.

Pal, S. K. and D. K. Chattoraj. 1988. P1 plasmid replication: Initiator sequestration is inadequate to explain control by initiator-binding sites. J. Bacteriol. **170**:3554-3560.

Pandey, D. P. and K. Gerdes. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucl. Acids Res. **33**:966-976.

Pansegrau, W., G. Ziegelin, and E. Lanka. 1987. The origin of conjugative IncP plasmid transfer: Interaction with plasmid-encoded products and the nucleotide sequence at the relaxation site. Biochim. Biophys. Acta **951**:365-374.

Park, K., E. Han, J. Paulsson, and D. K. Chattoraj. 2001. Origin pairing (‘handcuffing') as a mode of negative control of P1 plasmid copy number. EMBO J. **20**:7323-7332.

Parker, C. and J. R. Meyer . 2002. Selection of plasmid molecules for conjugative transfer and replacement strand synthesis in the donor. Mol. Microbiol. **46**:761-768.

Parker, C., X. l. Zhang, D. Henderson, E. Becker, and R. Meyer. 2002. Conjugative DNA synthesis: R1162 and the question of rolling-circle replication. Plasmid **48**:186-192.

Paulsson, J. 2002. Multileveled selection on plasmid replication. Genetics **161** :1373-1384.

Paulsson, J. and D. K. Chattoraj. 2006. Origin inactivation in bacterial DNA replication control. Mol. Microbiol. **61**:9-15.

Perwez, T. and R. J. Meyer . 1999. Stabilization of the relaxosome and stimulation of conjugal transfer are genetically distinct functions of the R1162 protein MobB. J. Bacteriol. **181**:2124-2131.

Pfaffl, M. W., G. W. Horgan, and L. Dempfle. 2002. Relative expression software tool (REST(C)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. Nucl. Acids Res. **30**:e36.

Pfaffl, M. W. 2004. Quantification strategies in real-time PCR, p. 87-112. *In* S. A. Bustin (ed.), A-Z of quantitative PCR. International University Line (IUL), La Jolla, CA, USA.

Pickett, M. A., J. S. Everson, P. J. Pead, and I. N. Clarke. 2005. The plasmids of Chlamydia trachomatis and Chlamydophila pneumoniae (N16): Accurate determination of copy number and the paradoxical effect of plasmid-curing agents. Microbiol. **151**:893-903.

Poirel, L., A. Carrër, J. D. Pitout, and P. Nordmann. 2009. Integron mobilization unit as a source of mobility of antibiotic resistance genes. Antimicrob. Agents Chemother. **53**:2492-2498.

Prentki, P., M. Chandler, and L. Carp. 1977. Replication of the prophage P1 during the cell cycle of *Escherichia coli*. Mol. Gen. Gen. **152**:71-76.

Projan, S. J., S. Carleton, and R. P. Novick. 1983. Determination of plasmid copy number by fluorescence densitometry. Plasmid **9**:182-190.

Providenti, M. A., J. M. O'Brien, R. J. Ewing, E. S. Paterson, and M. L. Smith. 2006. The copy-number of plasmids and other genetic elements can be determined by SYBR-Green-based quantitative real-time PCR. J. Microbiol. Methods **65**:476-487.

Pushnova, E. A., M. Geier, and Y. S. Zhu. 2000. An easy and accurate agarose gel assay for quantitation of bacterial plasmid copy numbers. Anal. Biochem. **284**:70-76.

Rawlings, D. E., I. Pretorius, and D. R. Woods. 1984. Expression of a *Thiobacillus ferrooxidans* origin of replication in *Escherichia coli*. J. Bacteriol. **158**:737-738.

Rawlings, D. E., I. M. Pretorius, and D. R. Woods. 1986. Expression of *Thiobacillus ferrooxidans* plasmid functions and the development of genetic systems for the Thiobacilli. Biotechnol. Bioeng. Symp.**16**:281-287.

Rawlings, D. E. 1999. Proteic toxin-antitoxin, bacterial plasmid addiction systems and their evolution with special reference to the pas system of pTF-FC2. FEMS Microbiol. Lett. **176**:269-277.

Rawlings, D. E. 2001. The molecular genetics of *Thiobacillus ferrooxidans* and other mesophilic, acidophilic, chemolithotrophic, iron- or sulfur-oxidizing bacteria. Hydrometallurgy **59**:187-201.

Rawlings, D. E. and E. Tietze. 2001. Comparative biology of IncQ and IncQ-Like plasmids. Microbiol. Mol. Biol. Rev. **65**:481-496.

Rawlings, D. E. 2005. The evolution of pTF-FC2 and pTC-F14, two related plasmids of the IncQ-family. Plasmid **53**:137-147.

Reith, M., R. Singh, B. Curtis, J. Boyd, A. Bouevitch, J. Kimball, J. Munholland, C. Murphy, D. Sarty, J. Williams, J. Nash, S. Johnson, and L. Brown. 2008. The genome of Aeromonas salmonicida subsp. salmonicida A449: insights into the evolution of a fish pathogen. BMC Genomics **9**:427.

Reynolds, M. G. 2000. Compensatory evolution in rifampin-resistant *Escherichia coli*. Genetics **156**:1471-1481.

Roberts, R. C. and D. R. Helinski. 1992. Definition of a minimal plasmid stabilization system from the broad-host-range plasmid RK2. J. Bacteriol. **174**:8119-8132.

Rohrer, J. and D. E. Rawlings. 1992. Sequence analysis and characterization of the mobilization region of a broad-host-range plasmid, pTF-FC2, isolated from *Thiobacillus ferrooxidans*. J Bacteriol. **174**:6230-6237.

Rotger, R. and C. Nombela. 1983. Characterization of penicillin-resistant □-lactamase-producing strains of *Neisseria gonorrhaoeae* isolated in Spain (1978-1981). Microbiol. Esp. **36**:115-122.

Rotger, R., F. Rubio, and C. Nombela. 1986. A Multi-resistance plasmid isolated from commensal *Neisseria* species is closely related to the Enterobacterial plasmid RSF1010. J Gen Microbiol **132**:2491-2496.

Saano, A. K. and V. V. Zinchenko. 1987. A new IncQ plasmid R89S: Properties and genetic organization. Plasmid **17**:191-201.

Sakai, H. and T. Komano. 1996. DNA replication of IncQ broad-host-range plasmids in gram-negative bacteria. Biosc. Biotech. Biochem. **60**:377-382.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sarmah, A. K., M. T. Meyer, and A. B. A. Boxall. 2006. A global perspective on the use, sales, exposure pathways, occurrence, fate and effects of veterinary antibiotics (VAs) in the environment. Chemosphere **65**:725-759.

Scherzinger, E., M. M. Bagdasarian, P. Scholz, R. Lurz, B. Rückert, and M. Bagdasarian. 1984. Replication of the broad host range plasmid RSF1010: Requirement for three plasmid-encoded proteins. Proc. Natl Acad. Sci. U.S.A. **81**:654-658.

Scherzinger, E., V. Haring, R. Lurz, and S. Otto. 1991. Plasmid RSF1010 DNA replication *in vitro* promoted by purified RSF1010 RepA, RepB and RepC proteins. Nucl. Acids Res. **19**:1203-1211.

Scherzinger, E., V. Kruft, and S. Otto. 1993. Purification of the large mobilization protein of plasmid RSF1010 and characterization of its site-specific DNA-cleavage/DNA-joining activity. Eur. J. Biochem. **217**:929-938.

Scherzinger, E., G. Ziegelin, M. Bárcena, J. M. Carazo, R. Lurz, and E. Lanka. 1997. The RepA protein of plasmid RSF1010 is a replicative DNA helicase. J. Biol. Chem. **272**:30228-30236.

Scholz, P., V. Haring, B. Wittmann-Liebold, K. Ashman, M. Bagdasarian, and E. Scherzinger. 1989. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. Gene **75**:271-288.

Schmidt, T., K. Friehs, and E. Flaschel. 1996. Rapid determination of plasmid copy number. J. Biotechnol. **49**:219-229.

Schrag, S. J., V. Perrot, and B. R. Levin. 1997. Adaptation to the fitness costs of antibiotic resistance in *Escherichia coli*. Proc. Biol. Sci. **264**:1287-1291.

Schröder, 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.

Shultzaberger, R. K., Z. Chen, K. A. Lewis, and T. D. Schneider. 2007. Anatomy of *Escherichia coli* σ70 promoters. Nucl. Acids Res. **35**:771-788.

Simon, R., U. Priefer, and A. Puhler. 1983. A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in Gram negative bacteria. Nat. Biotech. **1**:784-791.

Škulj, M., V. Okršlar, Š. Jalen, S. Jevševar, P. Slanc, B. Štrukelj, and V. Menart. 2008. Improved determination of plasmid copy number using quantitative real-time PCR for monitoring fermentation processes. Microb. Cell Fact. **7**:6.

Smalla, K., H. Heuer, A. Götz, D. Niemeyer, E. Krögerrecklenfort, and E. Tietze. 2000. Exogenous isolation of antibiotic resistance plasmids from piggery manure slurries reveals a high prevalence and diversity of IncQ-like plasmids. Appl. Environ. Microbiol. **66**:4854-4862.

Smith, A. S. G. and D. E. Rawlings. 1997. The posion-antidote stability system of the broad-host-range *Thiobacillus ferrooxidans* plasmid pTF-FC2. Mol. Microbiol. **26**:961-970.

Smith, A. S. G. and D. E. Rawlings. 1998a. Efficiency of the pTF-FC2 pas poison-antidote stability system in *Escherichia coli* is affected by the host strain, and antidote degradation requires the Lon protease. J. Bacteriol. **180**:5458-5462.

Smith, A. S. G. and D. E. Rawlings. 1998b. Autoregulation of the pTF-FC2 proteic poison-antidote plasmid addiction system (pas) is essential for plasmid stabilization. J. Bacteriol. **180**:5463-5465.

Sozhamannan, S. and D. K. Chattoraj. 1993. Heat shock proteins DnaJ, DnaK, and GrpE stimulate P1 plasmid replication by promoting initiator binding to the origin. J. Bacteriol. **175**:3546-3555.

Stalker, D. M., R. Kolter, and D. R. Helinski. 1979. Nucleotide sequence of the region of an origin of replication of the antibiotic resistance plasmid R6K. Proc. Natl. Acad. Sci. U.S.A. **76**:1150-1154.

Stewart, F. M. and B. R. Levin. 1977. The population biology of bacterial plasmids: A priori conditions for the existence of conjugationally transmitted factors. Genetics **87**:209-228.

Sýkora, P. 1992. Macroevolution of plasmids: A model for plasmid speciation. J. Theor. Biol. **159**:53-65.

Taguchi, Y., K. Tanaka, Y. Honda, D. M. Miao, H. Sakai, T. Komano, and M. Bagdasarian. 1996. A *dnaA* box can functionally substitute for the priming signals in the *oriV* of the broad host-range plasmid RSF1010. FEBS Lett. **388**:169-172.

Tao, L., R. E. Jackson, and Q. Cheng. 2005. Directed evolution of copy number of a broad host range plasmid for metabolic engineering. Metab. Eng. **7** :10-17.

Tanaka, K., K. Kino, Y. Taguchi, D. M. Miao, Y. Honda, H. Sakai, T. Komano, and M. Bagdasarian. 1994. Functional difference between the two oppositely oriented priming signals essential for the initiation of the broad host-range plasmid RSF1010 DNA replication. Nucl. Acids Res. **22**:767-772.

Tettelin, H. and T. Feldblyum. 2009. Bacterial genome sequencing. Methods Mol. Biol. **551**:231-247.

Thomas, C. M. 2000. Paradigms of plasmid organization. Mol. Microbiol. **37**:485-491.

Thomas, C. M. 2004. Evolution and population genetics of bacterial plasmids, p. 509-528. *In* B. E. Funnell and G. J. Phillips (ed.), Plasmid Biology. ASM Press, Washington, D.C.

Tietze, E. 1998. Nucleotide sequence and genetic characterization of the novel IncQ-like plasmid pIE1107. Plasmid **39**:165-181.

Tietze, E., H. Tschäpe, and W. Voigt. 1989. Characterization of new resistance plasmids belonging to incompatibility group IncQ. J. Basic Microbiol. **29**:695-706.

Tsuchimoto, S., H. Ohtsubo, and E. Ohtsubo. 1988. Two genes*, pemK* and *pemI*, responsible for stable maintenance of resistance plasmid R100. J. Bacteriol. **170**:1461-1466.

Tsuchimoto, S., Y. Nishimura, and E. Ohtsubo. 1992. The stable maintenance system pem of plasmid R100: degradation of PemI protein may allow PemK protein to inhibit cell growth. J. Bacteriol. **174**:4205-4211.

Tsutsui, H., A. Fujiyama, T. Murotsu, and K. Matsubara. 1983. Role of nine repeating sequences of the mini-F genome for expression of F-specific incompatibility phenotype and copy number control. J. Bacteriol. **155**:337-344.

Turgeon, N., C. Laflamme, J. Ho, and C. Duchaine. 2008. Evaluation of the plasmid copy number in *B. cereus* spores, during germination, bacterial growth and sporulation using real-time PCR. Plasmid **60**:118-124.

Van Melderen, L. and M. Saavedra De Bast. 2009. Bacterial toxin-antitoxin systems: More than selfish entities? PLoS Genet. **5**:e1000437.

Van Zyl, L. J., S. M. Deane, and D. E. Rawlings. 2003. Analysis of the mobilization region of the broad-host-range IncQ-like plasmid pTC-F14 and its ability to interact with a related plasmid, pTF-FC2. J. Bacteriol. **185**:6104-6111.

Van Zyl, L. J., S. M. Deane, L. A. Louw, and D. E. Rawlings. 2008. Presence of a family of plasmids (29 to 65 Kilobases) with a 26-kilobase common region in different strains of the sulfur-oxidizing bacterium *Acidithiobacillus caldus*. Appl. Environ. Microbiol. **74**:4300-4308.

Vogel, J., H. L. Andrews, S. K. Wong, and R. R. Isberg. 1998. Conjugative transfer by the virulence system *of Legionella pneumophila*. Science **279**:873-876.

Watve, M. M., N. Dahanukar, and M. G. Watve. 2010. Sociobiological control of plasmid copy number in bacteria. PLoS ONE **5**:e9328.

Whittle, G., M. E. Katz, E. H. Clayton, and B. F. Cheetham. 2000. Identification and characterization of a native *Dichelobacter nodosus* plasmid, pDN1. Plasmid **43**:230-234.

Williams, D. R. and C. M. Thomas. 1992. Active partitioning of bacterial plasmids. J Gen Microbiol **138**:1-16.

Willitts, N. and C. Crowther. 1981. Mobilization of the non-conjugative IncQ plasmid RSF1010. Genet. Res. **37**:311-316.

Willson, P. J., H. G. Deneer, A. Potter, and W. Albritton. 1989. Characterization of a streptomycin-sulfonamide resistance plasmid from *Actinobacillus pleuropneumoniae*. Antimicrob. Agents Chemother. **33**:235-238.

Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mpl8 and pUC19 vectors. Gene **33**:103-119.

Yoichi, H., S. Hiroshi, and K. Tohru. 1988. Two single-strand DNA initiation signals located in the oriV region of plasmid RSF1010. Gene **68**:221-228.

Zhang, S. and R. Meyer. 1997. The relaxosome protein MobC promotes conjugal plasmid mobilization by extending DNA strand separation to the nick site at the origin of transfer. Mol. Microbiol. **25**:509-516.

Zhang, Y., J. Zhang, K. P. Hoeflich, M. Ikura, G. Qing, and M. Inouye. 2003. MazF cleaves cellular mRNAs specifically at ACA to block protein synthesis in *Escherichia coli*. Mol. Cell **12**:913-923.

Zhang, X., S. Zhang, and R. J. Meyer. 2003. Molecular handcuffing of the relaxosome at the origin of conjugative transfer of the plasmid R1162. Nucl. Acids Res. **31**:4762-4768.

Zhang, J., Y. Zhang, L. Zhu, M. Suzuki, and M. Inouye. 2004. Interference of mRNA function by sequence-specific endoribonuclease PemK. J. Biol. Chem. **279**:20678-20684.

Zhou, H. and R. J. Meyer. 1990. Deletion of sites for initiation of DNA synthesis in the origin of broad host-range plasmid R1162. J. Mol. Biol. **214**:685-697.