Chapter 3

Mosaicism of the Large Natural *Escherichia coli* Plasmid pRK100

Marjanca Starčič Erjavec[1,2], Wim Gaastra[2], Jos P. M. van Putten[2], and Darja Žgur-Bertok[1]

[1]Department of Biology, Biotechnical Faculty, University of Ljubljana, Večna pot 111, 1000 Ljubljana, Slovenia

[2]Department of Infectious Diseases and Immunology, Utrecht University, P.O. Box 80.165, 3508 TD Utrecht, The Netherlands

Submitted for publication

ABSTRACT

In search for the evolutionary origin of the conjugative F-like plasmid pRK100, we determined the plasmid's functional replication region(s) and performed targeted genetic analysis of the plasmid. Construction of minireplicons via ligation of Tn*1725* with plasmid fragments and targeted cloning of putative replication regions, followed by sequence analysis indicated two functional replication regions, a plasmid F-like RepFIB and a plasmid R1-like RepFIIA replication region. Partial nucleotide sequencing of regions of the plasmid revealed genes that encode a putative enterochelin iron uptake system previously associated with an *Escherichia coli* pathogenicity island, PAI III_{536}, and the pColV-like aerobactin genes. In addition, a homologue of the plasmid R100-like *rmoA* gene was found that exhibits strong similarity to the Hha/YmoA class of modulators of gene expression. PCR and hybridisation experiments further demonstrated that pRK100 harbours IS*2* and IS*3* insertion sequences that may have facilitated in the acquisition of elements from other DNA molecules. These data together with the previous identification of a F-like *tra* region and a pColIa-like colicin Ia, indicate that pRK100 has a highly mosaic structure with elements derived from many different known large natural plasmids, and also from chromosome.

Key words: replication region, RepFIB, RepFIIA, F-like plasmid, ColV-like plasmid, construction of minireplicons, transposable element, iron uptake system

1. INTRODUCTION

Plasmids are self-replicating extrachromosomal DNA elements that may embody 1% to more than 10% of the genome of a given bacterial strain. To propagate and maintain themselves, plasmids have developed sophisticated replication functions that are encoded by distinct gene replication regions. These regions typically contain the origin of replication (generically termed *ori*), a gene encoding the protein involved in the initiation of replication, and genes whose products are required for the control of replication (18). To further ensure their stability in a bacterial population plasmids also encode systems responsible for resolution of plasmid multimers, equipartitioning at cell division, and post-segregational killing of plasmid free cells that arise in a population.

Conjugative plasmids can transfer themselves to other bacteria via the plasmid-encoded conjugation machinery. In addition they can facilitate the transfer of non-mobilisable genetic elements. Conjugative plasmids have been demonstrated to be important gene transfer vehicles in the spread of resistances to antibiotics and chemicals as well as of metabolic properties and virulence genes. The growing evidence from genome sequence analysis that microbial evolution has to an important degree been shaped by horizontal gene transfer has further renewed the interest in plasmid conjugation and the properties of large natural conjugative plasmids.

Genetic and functional analyses of the various kinds of large natural conjugative plasmids that have been isolated suggest that they are generally larger than 40 kb in size and may have evolved from the integration into small plasmids of distinct chromosomal or plasmid elements and/or fusion of plasmids. Major classes of large natural plasmids are the IncFI and IncFII plasmids. This strict classification however, may turn out to be rather artificial as members of the plasmid families appear to have exchanged genetic information.

In order to learn more about the dynamic nature and the evolution of large natural conjugative plasmids, we investigated the plasmid pRK100. pRK100 is a 145 kb F-like conjugative plasmid that was discovered in a uropathogenic *Escherichia coli* strain. The plasmid harbours the transposon Tn*5431* that carries genes that confer resistance to ampicillin and tetracycline, and genes that encode the colicins V and Ia and an aerobactin iron uptake system (1), suggesting that it has acquired genetic elements from different origins. In the present study, we further investigated the mosaicism of pRK100 by determining the active origin(s) of replication, and targeted hybridisation and sequencing based on typical characteristics of other large natural plasmids.

2. MATERIALS AND METHODS

Bacterial strains, plasmids and growth conditions

The bacterial strains and plasmids used in this study and their characteristics are listed in Table 1. Bacteria were grown in Luria-Bertani (LB) medium with aeration at 37 °C. Ampicillin (Ap, 100 μg/ml), tetracycline (Tc, 10 μg/ml), kanamycin (Km, 30 μg/ml), chloramphenicol (Cm, 50 μg/ml), and nalidixic acid (Nal, 25 μg/ml) were added to the growth media, when appropriate.

Table 1. **Bacterial strains and plasmids**

Strain or DNA	Relevant features	Reference or source*
Strains		
KS533	Rough:K1:H7 harbouring pRK100	23
HB101	*hsdR hsdM recA13 supE44 leuB6 lacZ proA2*	D. Ehrlich[1]
CL225	HB101 harbouring pRK100	D. Žgur-Bertok
DH5α	Φ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *endA1 recA1 hsdR17 deoR*	BRL Life
	thi-1 supE44 gyrA96 relA1	Technologies
RU4404	MM294::Tn*1725* Cm^r *thi endA hsdR*	21
RU4405	MM294::Tn*1731* Tc^r *thi endA hsdR*	21
PC1994	*thyA argA bio pheA endA polA*	NCCB
Plasmids		
pMW2	plasmid with kanamycin cassette; Ap^r, Kn^r	M. Wösten[2]
pBluescript SK–	*E. coli* vector plasmid; Ap^r	Stratagene
pGEM-T Easy	T-vector for cloning of PCR products; Ap^r	Promega
pRK100	colV colIa natural plasmid; Ap^r, Tc^r	1
pS2	pBluescript SK- with the 1731 bp *Sal*I fragment of pRK100	This study
pS3	pBluescript SK- with the 2156 bp *Sal*I fragment of pRK100	This study
pS10	pBluescript SK- with the 2062 bp *Sal*I fragment of pRK100	This study
pSA11	pGEM-T Easy with the 1204 bp PCR product of pS3 (Sal11-A/Sal11-B)	This study
pSA14	pGEM-T Easy with the 839 bp PCR product of pS10 (Sal14-A/Sal14-B)	This study

*[1] Laboratoire Genetique Microbienne, INRA, Domain de Vilvert, Jouy en Josas, France
[2] *Department of Infectious Diseases and Immunology, Utrecht University, The Netherlands*

General DNA manipulation techniques

Plasmid DNA isolation, ligation, transformation and stringent hybridisation experiments were performed using standard methods (17). Restriction endonuclease digestions were carried out as specified by the manufacturer (Promega, Madison, WI). DNA fragments were isolated from agarose gels using the QIAquick Gel Extraction kit (Qiagen, Valencia, CA). DNA sequencing was performed using the ABI dye terminator cycle sequencing kit in an ABI PRISM™ 310 Genetic Analyzer automated sequencer. Sequences were analysed using Lasergene software (DNAstar, Inc.). The primers used for sequencing are listed in Table 2.

PCR amplification

The polymerase chain reactions (PCR) were performed in a 50 µl PCR reaction mixture with 20 pmol of forward and reverse primer (Table 2), 5 µl of purified template DNA, 0.2 mM of dNTP mixture (Pharmacia), 0.625 U Taq DNA polymerase (Promega) and 1× PCR buffer (Promega). The PCR conditions are given in Table 2. Each PCR started with a prolonged denaturation step (4 min, 94 °C), and ended with a prolonged extension step (10 min, 72 °C) after the last cycle. The pGEM-T Easy system (Promega) was used for cloning of the PCR products.

Table 2. **Oligonucleotide primers and PCR conditions**

Oligonucleotide primer*	Nucleotide sequence	PCR conditions
for RepFIIA PCR:		
FinO-1	5'-CGGATAAAGGCAGAACTTCAGGC-3'	(94 °C-1:00, 58 °C-0:30,
Sal14-B	5'-GAAAGGCGGCACTCTGTTGT-3'	72 °C-4:00) 30×
Sal11-A	5'-CCGGAAAGTTATATGACAGT-3'	(94 °C-1:00, 55 °C-0:30,
Sal14-B	5'-GAAAGGCGGCACTCTGTTGT-3	72 °C-3:00) 35×
for RepFIIA sequencing:		
Sal11-B	5'-GCTACGCCACAAAGTAAAGT-3'	
Sal14-A	5'-CCCTGAAGTGACCTCCTCTG-3'	
565-PS3	5'-CTCAATGAGGAAGTCACTGC-3'	
568-PS10	5'-AAACCTATTCACTGCCTGTC-3'	
for RepFIB PCR:		
RepA-f[1]	5'-GGAATTCTCGCTGCAAACCTTGTCACT-3'	(94 °C-0:30, 63 °C-0:30,
RepA-r[1]	5'-GGAATTCGGAGATCCTGCGTACACTGCCT-3'	72 °C-1:30) 30×
RepFIB-5	5'-GCGGACAATCCAAATGGTGA-3'	(94 °C-1:00, 55 °C-0:30,
RepFIB-6	5'-ATATGCTGTTCGCCACCCTC-3'	72 °C-3:00) 30×
for RepFIIA sequencing:		
FIB2-1	5'-CAATAACAACACCGTACAACC-3'	
FIB2-2	5'-ACAGTTTATGTTCAGCGGGAT-3'	
RepFIB-3	5'-TATCAGTAACATGCCACAGC-3'	
RepFIB-4	5'-TCCCGTAACCTGATGCTGAG-3'	
for RepFIC PCR:		
RepC-1	5'-TTCATCTAGTTTGGCGACGAGG-3'	(94 °C-1:00, 63 °C-1:00,
RepC-2	5'-GCTGTCTTTCGGGCTGATTTCT-3'	72 °C-2:00) 30×
for RepZ PCR:		
RepZ-1	5'-CCATATAACGCAGTACACTGGA-3'	(94 °C-1:00, 58 °C-0:30,
RepC-2	5'-GCTGTCTTTCGGGCTGATTTCG-3'	72 °C-1:30) 30×
for RepFIA PCR:		
RepFIA-f[2]	5'-CTCACTGAGGCGGCATATAGTC-3'	(94 °C-0:30, 63 °C-0:30,
RepFIA-r[2]	5'-ATGGAAGTGATATCGCGGAAGG-3'	72 °C-1:00) 30×
for aerobactin PCR:		
aer1	5'-TACCGGATTGTCATATGCAGACCGT-3'	(94 °C-0:30, 62 °C-0:30,
aer2	5'-AATATCTTCCTCCAGTCCGGAGAAG-3'	72 °C-0:50) 30×
for *traM* and *finP* PCR:		
FinP-1	5'-TATTGAGAAGCGTCGACAGG-3'	(94 °C-1:00, 55 °C-1:00,
FinP-2	5'-TGACGAACATGAGCAGCATC-3'	72 °C-1:00) 30×
for *traJ* PCR:		
PtraJ-1	5'-CGGGATCC-TCCAAAAAATGATGATGAAT-3'	(94 °C-1:00, 55 °C-0:30,
TraYN-2	5'-GCAGAACGTGTACCAAATCTT-3'	72 °C-1:00) 30×
for *traY* PCR:		
TraY-f[1]	5'-GGAATTCAAGATTTGGTACACGTTCTGC-3'	(94 °C-1:00, 63 °C-1:00,
TraY-r[1]	5'-GGAATTCCTTCCTCTTTATCTGCCTCCC-3'	72 °C-2:00) 30×
for *traD* PCR:		
TraD-f[1]	5'-GGAATTCCAGATTGCGTCCATGCGTATCC-3'	(94 °C-1:00, 63 °C-1:00,
TraD-r[1]	5'-GGAATTCATCACCACACATATCACCGCGC-3'	72 °C-1:00) 30×

Oligonucleotide primer*	Nucleotide sequence	PCR conditions
for *traI* PCR:		
TraI-1	5'-ACAGCGAATATACGTGACGG-3'	(94 °C-0:30, 57 °C-0:30,
FinO-4	5'-CGTGGTGACATTGATGATGG-3'	72 °C-3:00) 30×
for *finO* PCR:		
FinO-f[1]	5'-GGAATTCGAAGCGACCGGTACTGACACTG-3'	(94 °C-1:00, 63 °C-1:00,
FinO-r[1]	5'-GGAATTCGCCTGAAGTTCTGCCTTTATCCG-3	72 °C-2:00) 30×
for *ccdB* PCR:		
CcdB-1	5'-GAGAGCCGTTATCGTCTGTT-3'	(94 °C-1:00, 55 °C-1:00,
CcdB-2	5'-CTGAGATCAGCCACTTCTTC-3'	72 °C-1:00) 30×
for IS*2* PCR:		
IS2-1	5'-ACTTAACCCATTACAAGCCCGC-3'	(94 °C-0:30, 61 °C-0:30,
IS2-2	5'-AACCTGCTGTACCGCCATCGAA-3'	72 °C-1:00) 30×
for IS*3* PCR:		
IS3-1	5'-AGTTCAGCATCAAAGCAATGTG-3'	(94 °C-0:30, 59 °C-0:30,
IS3-2	5'-GCAGGCATTATCGTAGCAGCAA-3'	72 °C-1:00) 30×

*[1] Nucleotide sequence is based on Boyd, *et al.* (4)

[2] Nucleotide sequence is based on Mulec, *et al.* (12)

Construction of minireplicons

To construct minireplicons using transposon Tn*1725* (21), pRK100 (Ap^r, Tc^r) was first conjugally transferred to strain RU4404 that carries Tn*1725* (Cm^r) on the chromosome (21). Transconjugants were selected on LB plates with Cm, Ap, and Tc and incubated at 30 °C for several days to allow transposition of Tn*1725* into pRK100. pRK100 plasmids with inserted transposon (pRK100::Tn*1725*) were selected by conjugal transfer to DH5α (Nal^r) on LB plates containing Cm, Ap, Tc, and Nal. Several pRK100::Tn*1725* plasmids from independent experiments were isolated, partially digested with *Sal*I, self-ligated with T4-ligase (Gibco), and introduced into DH5α by electrotransformation. Transformants carrying minireplicons were selected on LB plates containing Cm.

Construction of minireplicons using PCR products of the putative replication regions was carried out by either cloning the PCR product into pGEM-T Easy, or by joining the DNA polymerase I polished PCR product with the *Sma*I fragment of pMW2 that harbours the kanamycin resistance gene. Transformants carrying minireplicons were selected on LB plates containing Km.

The obtained minireplicons were analysed by restriction enzyme digestion analysis, PCR and sequencing, as described above.

3. RESULTS

Identification of a functional RepFIB region

To determine the functional replication region(s) of pRK100, we initially utilised the randomly inserting transposon Tn*1725* (Cm^r) (21). For this purpose, pRK100 was transferred into the Tn*1725* carrying *E. coli* strain RU4404. After induction of the transposition event by growth at 30°C, pRK100 was re-isolated from strain RU4404, and partially digested. Self-ligation of the obtained plasmid fragments and re-introduction of the ligated plasmid fragments into DH5α, yielded several Cm resistant transformants. Plasmid analysis indicated that these transformants carried Tn*1725* (Cm^r) containing minireplicons with sizes between 19.3 and 24 kb.

Assuming that pRK100 contains a replication region that resembles one of the identified *rep* regions in other large natural plasmids, the pRK100 derived functional replication region of the minireplicons was investigated by PCR using RepFIIA, RepFIA, RepFIB, RepFIC, and RepZ specific primers (Table 2). Partial nucleotide sequencing of the obtained PCR products indicated that all of the minireplicons, which were obtained in independent experiments, harboured a RepFIB-like replication region.

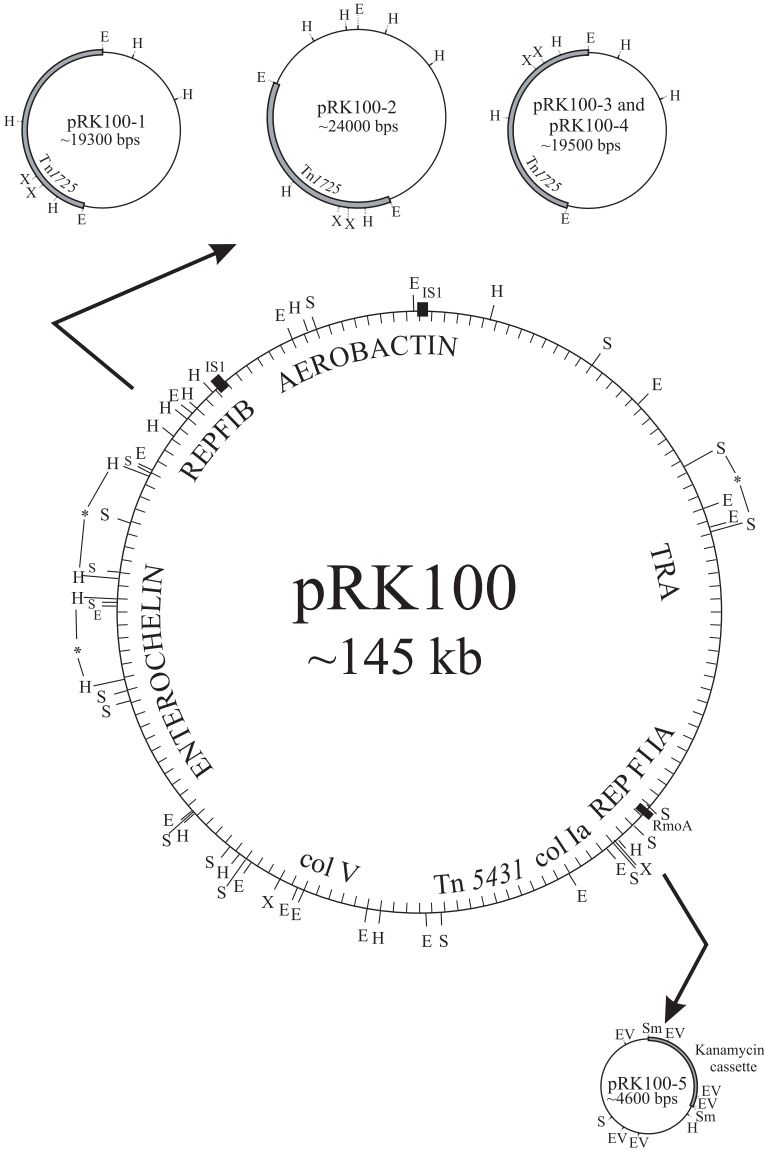

Fig. 1. **Map of pRK100 and the constructed minireplicons**
The large map depicts the known coding regions of pRK100 and the position of the newly identified replication regions. The smaller maps represent the RepFIIA (pRK100-5) and RepFIB (pRK100-1, pRK100-2, pRK100-3, pRK100-4) minireplicons. The restriction sites for *Eco*RI (E), *Eco*RV (EV), *Sal*I (S), *Hin*dIII (H), *Sma*I (Sm) and *Xho*I (X) are indicated.

Further nucleotide sequencing of the pRK100-1 RepFIB replicon using primer walking showed that RepFIB of pRK100 was highly similar (96%) to RepFIB replication regions of the plasmids F, pColV-K30, pO157, R124 and other related plasmids (data not shown). Restriction analysis of the obtained RepFIB minireplicons located the RepFIB region in proximity of the aerobactin encoding region on the pRK100 plasmid map (Fig. 1).

Identification of a functional RepFIIA region

Previous hybridisation experiments suggested that pRK100 may carry a RepFIC replication region downstream of the transfer *tra* region (1). In order to further investigate the nature and functionality of this region, which was not obtained as a minireplicon with Tn*1725* approach, we PCR amplified the region with the primers FinO-1 and Sal14-B that encompass the downstream *finO* gene of the *tra* operon and the downstream ColIa region of plasmid pRK100, respectively (Fig. 1). This yielded a 4.5 kb PCR product, which was sequenced using primer walking. Sequence analysis using the BLAST algorithm (www.ncbi.nlm.nih.gov.) revealed that the obtained sequence was 95% similar to the RepFIIA replication region of plasmid R1 of *Escherichia coli* and 94% similar to that of the virulence plasmid pWR100 of *Shigella flexneri*.

To assess the possible functionality of the newly identified RepFIIA region of the pRK100, we amplified the region with the primers Sal11-A and Sal14-B (Table 2) and attempted to clone the amplified fragment into pGEM-T Easy. Introduction of these constructs into either a *polA*+ or *polA*– *E. coli* background yielded no transformants. As an alternative approach, the RepFIIA PCR product was ligated to the kanamycin resistance cassette derived from the plasmid pMW2. Electrotransformation of these constructs into *E. coli* DH5α yielded several Km resistant transformants. Restriction enzyme digestion analysis, PCR and nucleotide sequencing confirmed the successful construction of the RepFIIA minireplicon (Fig. 1). Together, these results strongly suggest that pRK100 has two intact replicons, RepFIB and RepFIIA.

Heterogeneous origin of pRK100

The presence of two intact replication regions, RepFIB and RepFIIA, with high similarity to those present on the large natural plasmids F, pColV-K30, pO157 and R124, and pWR100 and R1, respectively, and the mapping of RepFIB and RepFIIA to different parts of pRK100 (Fig. 1), suggested that pRK100 may constitute a plasmid chimera composed of elements from different plasmids. To obtain further evidence for the mosaic nature and origin of pRK100, we tested the plasmid for the presence of several other large natural plasmid specific regions, and sequenced several thus far unidentified DNA regions.

Since large F-like plasmids frequently harbour multiple insertion sequences that may facilitate in the acquisition of elements from other plasmids, PCR reactions with primers specific for either IS*2* or IS*3* were performed. PCR products were obtained in both reactions indicating that at least one copy of each insertion sequence is carried by pRK100. PCR amplification and hybridisation experiments aimed at the detection of *ccdB*, which encodes the CcdB protein of the post-segregational CcdAB killing system in F plasmid (9), yielded no positive results (data not shown), indicating that pRK100 has not acquired this killing system.

To gain more information about the nature and origin of the region upstream of RepFIB, we cloned and partially sequenced a 1731 bp *Sal*I restriction fragment (pS2). Hybridisation experiments mapped this fragment to the correct region (data not shown). Sequence analysis indicated that the fragment carried two adjacent open reading frames that were 95% and 98% similar at the amino acid level to the *iroD* and *iroC* genes, respectively, of the uropathogenic *E. coli* strain 536 pathogenicity island PAI III_{536} (Genbank accession number X16664). The *iroD* gene encodes a putative ferric ente-

rochelin esterase, while the *iroC* gene product is an ABC transport protein. These data suggest that this part of pRK100 may be of chromosomal origin.

Nucleotide sequence analysis of a 2156 bp *Sal*I restriction fragment (pS3) that mapped by hybridisation to a region downstream of the *tra* region (Fig. 1) revealed an open reading frame that was 95% identical to the RmoA protein encoded on plasmid R100. In R100, this protein is a modulator of R100 conjugal transfer (15). The protein sequence of R100 RmoA is highly homologous to chromosomally encoded protein Hha (14) that downregulates the expression of the virulence factor haemolysin in *E. coli* (13), and to YmoA that modulates the expression of Yop proteins and the YadA adhesin in *Yersinia enterocolitica* (5). Both Hha and YmoA belong to a protein family that downregulates gene expression in enterobacteria (11).

Sequence analysis of the PCR product obtained with primers aer1 and aer2 demonstrated the presence on pRK100 of genes encoding an aerobactin iron uptake system consistent with earlier reported DNA hybridisation data (1). BLAST analysis of the obtained aerobactin showed that the region was 100% identical to the sequence found on *E. coli* pColV-K30, and 97% identical to sequences found on the *Shigella flexneri* SHI-2 pathogenicity island and the *Shigella boydii* 0-1392 aerobactin island SHI-3.

Mosaic nature of pRK100

The newly identified traits of pRK100 described above, together with the previous identification of a F-like *tra* region and a pColIa-like colicin Ia (1), indicate that pRK100 has a highly mosaic structure with elements derived from the chromosome and other known large natural plasmids. In Figure 2, the mosaic nature of pRK100 as well as that of other large natural plasmids is illustrated by a systematic comparison of the identified pRK100 elements with sequences present in other large natural plasmids.

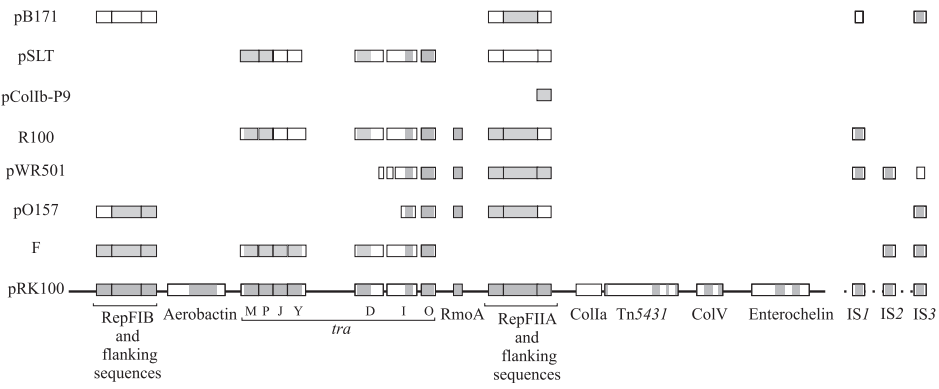

Fig. 2. **Mosaicism of pRK100**

The obtained pRK100 sequences with the assigned functions were analysed using the BLAST algorithm for their similarity with other plasmids. In the Figure, the regions of pRK100 that are present in one or more of the other plasmids are marked as boxes (note that the differences in box sizes are only approximate). The shaded parts of the boxes are very similar (>95%) to the obtained pRK100 nucleotide sequence. The nonshaded parts represent sequences that were not compared due to the fact that the corresponding pRK100 regions were not sequenced. The smaller size of distinct boxes in some of the other plasmids denotes that part of the region is missing. Only the presence of IS sequences is denoted. Note that only large natural plasmids whose complete nucleotide sequences are available in Genbank are represented.

4. DISCUSSION

Conjugative plasmids are increasingly recognised as important vehicles of horizontal gene transfer and knowledge of their evolution and mechanism of action is important in understanding the spread of traits of medical and ecological significance. The data obtained in the present study indicate the large natural conjugative plasmid pRK100 as a plasmid mosaic that is composed of a diverse set of plasmid and chromosomal genetic elements. The most prominent evidence of the highly mosaic plasmid structure of pRK100 was the demonstration of the presence of two intact replication regions, RepFIB and RepFIIA. The RepFIB replication region was found to be most similar (96%) to the RepFIB replication region of plasmid F and other related plasmids. In these plasmids, this region typically contains a single initiator gene *repA* that is flanked by a series of repeat elements. The *repA* promoter is located within the upstream repeat elements and the expression of the gene is autoregulated by the binding of RepA to the repeats, which presumably prevents RNA polymerase access to the promoter (19).

The sequence of the RepFIIA replication region of pRK100 was most similar (95%) to the RepFIIA of plasmid R1. IncFII replicons (originally designated RepFIIA) typically consist of *repA2* encoding a repressor, the *copA* gene that encodes an antisense RNA molecule, a *repA1* gene whose protein initiates plasmid replication by binding to the downstream *ori*, the *tapA* gene encoding a short leader peptide, and a *repA4* region. In this system, the RepA2 repressor is assumed to regulate transcription of *repA1* mRNA, while the antisense RNA CopA, which is complementary to the leader region of *repA1* mRNA (CopT), regulates translation. When CopA binds to CopT, *tapA,* which is necessary for RepA1 synthesis, is not expressed (3). The *repA4* appears to be important for the stability of plasmid maintenance (10).

The presence of two replication regions with distinctive mechanisms of action within one large natural plasmid is not unique and has previously been reported for several plasmids, such as pO157 and pB171 (2). The successful construction of minireplicons carrying either RepFIB or RepFIIA indicates that both replication regions of pRK100 are functional in *E. coli*. The identification of the replication regions required different experimental strategies. The RepFIB minireplicon was isolated by religation of plasmid pRK100 digestion fragments carrying Tn*1725*. The fact that we were not able to isolate RepFIIA minireplicons using this approach may be due to the apparent preference of Tn*3*-like transposons (such as Tn*1725*) (8) for the RepFIIA replicon (16). The alternative approach of ligating the PCR generated RepFIIA fragment with DNA carrying a kanamycin resistance gene circumvented this problem and resulted in the formation of functional RepFIIA minireplicons. The functionality of the identified replication regions in the setting of pRK100 remains to be defined.

Previous DNA hybridisation experiments suggested that pRK100 may harbour a RepFIC replication region belonging to the IncFII extended family of replicons (1). The inability to obtain a PCR product with primers specific for RepFIC indicated that the RepFIC probe may have cross-hybridised with another replication region. Indeed it is known that the used pULB2440 derived RepFIC probe is not highly specific and thus may hybridise with different replication regions including Rep9, RepFIIA, RepI1, RepB/O and RepK (6). The mapping of RepFIIA to the same position in pRK100, as previously suggested for RepFIC, further supports the idea that pRK100 carries RepFIIA and not RepFIC.

The presence of genes that exhibit a strong similarity with sequences present on of a wide variety of other plasmids further underlined the mosaic nature of pRK100 and suggests that during its evolution pRK100 apparently sequestered genetic information from many different sources. Typical examples of the mosaicism of pRK100 are the *tra* region that enables conjugal transfer and the RepFIB replication region that appear to be acquired from a F-like plasmid, the replication region RepFIIA that is probably derived from a R1-like plasmid, the aerobactin uptake system and the colicin V deter-

minants likely originating from a pColV-like plasmid, the *Shigella flexneri* SHI-2 pathogenicity island and/or the *Shigella boydii* 0-1392 aerobactin island SHI-3, and the colicin Ia encoding region derived from a ColIa harbouring plasmid into which transposon Tn*5431* had inserted (24). Particularly noteworthy is the newly identified putative enterochelin-based iron uptake system on pRK100 that is most similar (98%) to sequences on the pathogenicity island PAI III_{536} of the uropathogenic *E. coli* strain 536. This PAI also harbours the *cvaB* gene required for colicin V export (7). The two convergent operons required for colicin V synthesis, export and immunity are encoded by ColV plasmids (22) as well as by pRK100.

An additional noteworthy finding was that partial nucleotide sequencing revealed that pRK100 carries a gene that is very similar to the *rmoA* gene on plasmid R100 (Fig. 2). This in conjunction with the finding that *rmoA* is situated between the promoter distal region of the *tra* operon and the RepFIIA replication region in both pRK100 and R100 suggests that these plasmid regions are related. RmoA seems to be involved in modulation of plasmid transfer in response to some environmental factors (15) and comparative sequence analysis has shown the same gene organisation in a number of other characterised F-like plasmids. The DNA sequences of the corresponding region in plasmid R1 remains to be defined.

The backbone of pRK100 has previously been suggested to be similar to that of plasmid F (1). Our data suggest that this may be only partially true as we found no evidence for the presence on pRK100 of the CcdAB post-segregational killing system of plasmid F. It can be imagined that in pRK100 the colicinogenicity and conjugal transfer are the important factors in the maintaining plasmid stability among the population. Under this scenario, bacteria that have lost the plasmid are no longer immune and can be killed by plasmid-harbouring cells producing the colicin, while bacteria that have lost the plasmid can serve as recipients in conjugal transfer. An additional factor that may contribute to persistence of the plasmid is the presence of Tn*5431*. The resolvase of the transposon may function as a multimer resolution system and thus prevent dimerisation of plasmids and unequal plasmid distribution to daughter cells (20).

The mechanism(s) that contributed to the mosaicism of pRK100, are unknown. However, hybridisation and PCR experiments showed that pRK100 harbours several insertion sequences including IS*1* (1) and, as demonstrated in this study, IS*2* and IS*3* elements that can mediate recombination and rearrangement events. Some of the genes, such as the enterochelin uptake system, appear to originate from a chromosomal pathogenicity island. Pathogenicity islands (PAI) range in size from approximately 20-120 kb and have been proposed to be horizontally transferred via temperate bacteriophages. Extensive analysis of the PAIs of uropathogenic *E. coli* strains indicate that they contain mosaic sequences (7). The discovery of plasmid encoded enterochelin and aerobactin may imply that conjugative plasmids could be involved in the flux of chromosomal PAI sequences and, possibly, their horizontal transfer. To this end, the highly mosaic plasmid structure of pRK100 may be considered as an important player in the development of genome plasticity by acting as an intermediate in the transfer of a variety of virulence, antibiotic resistance, or metabolic traits between populations.

5. REFERENCES

1. **Ambrožič, J., A. Ostroveršnik, M. Starčič, I. Kuhar, M. Grabnar, and D. Žgur-Bertok.** 1998. *Escherichia coli* ColV plasmid pRK100: genetic organization, stability and conjugal transfer. *Microbiology* **144**: 343–352.
2. **Bergquist, P. L., S. Saadi, and W. K. Maas.** 1986. Distribution of basic replicons having homology with RepFIA, RepFIB and RepFIC among IncF group plasmids. *Plasmid* **15**:19–34.

3. **Blomberg, P., K. Nordström, and E. G. H. Wagner.** 1992. Replication control of plasmid R1: RepA synthesis is regulated by CopA RNA through inhibition of leader peptide translation. *EMBO J* **11**:2675–2683.
4. **Boyd, E. F., C. W. Hill, S. M. Rich, and D. L. Hartl.** 1996. Mosaic structure of plasmids from natural populations of *Escherichia coli*. *Genetics* **143**:1091–1100.
5. **Cornelis, G. R., C. Sluiters, I. Delor, D. Geib, K. Kaniga, C. Lambert de Rouvroit, M. P. Sory, J. C. Vanooteghem, and T. Michiels.** 1991. *ymoA*, a *Yersinia enterocolitica* chromosomal gene modulating the expression of virulence functions. *Mol Microbiol* **5**:1023–1034.
6. **Couturier, M., F. Bex, P. L. Bergquist, and W. K. Maas.** 1988. Identification and classification of bacterial plasmids. *Microbiol Rev* **52**:375–395.
7. **Dobrindt, U., G. Blum-Öhler, T. Hartsch, G. Gottschalk, E. Z. Ron, R. Fünfstück, and J. Hacker.** 2001. S-Fimbria-encoding determinant sfa_1 is located on pathogenicity island III_{536} of uropathogenic *Escherichia coli* strain 536. *Infect. Immun.* **69**:4248–4256.
8. **Heffron, F.** 1983. Tn*3* and its relatives, p. 223–260. *In* J. A. Shapiro (editor), Mobile genetic elements. Academic Press, New York.
9. **Helinski, D. R., A. E. Toukdarian, and R. P. Novick.** 1996. Replication control and other stable maintenance mechanisms of plasmids, p. 2295–2324. *In* F. C. Neidhardt, *et al.* (editors), *Escherichia coli* and *Salmonella*, Cellular and molecular biology, 2nd edition. ASM Press, Washington, D. C.
10. **Jiang, T., Y.-N. Min, W. Liu, D. D. Womble, and R. H. Rownd.** 1993. Insertion and deletion mutations in the *repA4* region of the IncFII plasmid NR1 cause unstable inheritance. *J Bacteriol* **175**:5350–5358.
11. **Mikulskis, A. V., and G. R. Cornelis.** 1994. A new class of proteins regulating gene expression in enterobacteria. *Mol Microbiol* **11**:77–86.
12. **Mulec, J., M. Starčič, and D. Žgur-Bertok.** 2002. F-like plasmid sequences in enteric bacteria of diverse origin, with implication of horizontal transfer and plasmid host range. *Curr Microbiol* **44**:231–235.
13. **Nieto, J. M., M. Carmona, S. Bolland, Y. Jubete, F. de la Cruz, and A. Juárez.** 1991. The *hha* gene modulates haemolysin expression in *Escherichia coli*. *Mol Microbiol* **5**:1285–1293.
14. **Nieto, J. M., and A. Juárez.** 1996. A new member of Hha/YmoA class of bacterial regulators in plasmid R100 of *Escherichia coli*. *Mol Microbiol* **19**:407.
15. **Nieto, J. M., A. Prenafeta, E. Miquelay, S. Torrades, and A. Juárez.** 1998. Sequence, identification and effect on conjugation of the *rmoA* gene of plasmid R100-1. *FEMS Microbiol Lett* **169**:59–66.
16. **Picken, R. N., A. J. Mazaitis, and W. K. Maas.** 1984. High incidence of transposon Tn*3* insertions into a replication control gene of the chimeric R/Ent plasmid pCG86 of *Escherichia coli*. *J Bacteriol* **160:** 430–433.
17. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
18. **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.
19. **Spiers, A. J., N. Bhana, and P. L. Bergquist.** 1993. Regulatory interactions between RepA, an essential replication protein, and the DNA repeats of RepFIB from plasmid P307. *J Bacteriol* **175**:4016–4024.
20. **Thomas, C. M.** 2000. Paradigms of plasmid organization. *Mol Microbiol* **37**:485–491.
21. **Ubben, D., and R. Schmitt.** 1986. Tn*1721* derivatives for transposon mutagenesis, restriction mapping and nucleotide sequence analysis. *Gene* **41**:145–152.
22. **Waters, V. L., and J. H. Crosa.** 1991. Colicin V virulence plasmids. *Microbiol Rev* **55**:437–450.
23. **Žgur-Bertok, D., E. Modrič, and M. Grabnar.** 1990. Aerobactin uptake system, ColV production, and drug resistance encoded by a plasmid from a urinary tract infection *Escherichia coli* strain of human origin. *Can J Microbiol* **36**:297–299.
24. **Žgur-Bertok, D., J. Ambrožič, and M. Grabnar.** 1996. Tn*5431* arose by transposition of Tn*3* into Tn*1721*. *Can J Microbiol* **42**:1274–1276.