

## ORIGINAL ARTICLE

# Shifts in the host range of a promiscuous plasmid through parallel evolution of its replication initiation protein

Masahiro Sota<sup>1,3,4</sup>, Hirokazu Yano<sup>1,4</sup>, Julie M Hughes<sup>1</sup>, Gary W Daughdrill<sup>1,5</sup>, Zaid Abdo<sup>2</sup>, Larry J Forney<sup>1</sup>, and Eva M Top<sup>1</sup>

<sup>1</sup>Department of Biological Sciences, Initiative for Bioinformatics and Evolutionary Studies, University of Idaho, 252 Life Sciences South, Moscow, ID, USA and <sup>2</sup>Departments of Mathematics and Statistics, Initiative for Bioinformatics and Evolutionary Studies, University of Idaho, Brink Hall, Moscow, ID, USA

**The ability of bacterial plasmids to adapt to novel hosts and thereby shift their host range is key to their long-term persistence in bacterial communities. Promiscuous plasmids of the incompatibility group P (IncP)-1 can colonize a wide range of hosts, but it is not known if and how they can contract, shift or further expand their host range. To understand the evolutionary mechanisms of host range shifts of IncP-1 plasmids, an IncP-1 $\beta$  mini-replicon was experimentally evolved in four hosts in which it was initially unstable. After 1000 generations in serial batch cultures under antibiotic selection for plasmid maintenance (kanamycin resistance), the stability of the mini-plasmid dramatically improved in all coevolved hosts. However, only plasmids evolved in *Shewanella oneidensis* showed improved stability in the ancestor, indicating that adaptive mutations had occurred in the plasmid itself. Complete genome sequence analysis of nine independently evolved plasmids showed seven unique plasmid genotypes that had various kinds of single mutations at one locus, namely, the N-terminal region of the replication initiation protein TrfA. Such parallel evolution indicates that this region was under strong selection. In five of the seven evolved plasmids, these *trfA* mutations resulted in a significantly higher plasmid copy number. Evolved plasmids were found to be stable in four other naive hosts, but could no longer replicate in *Pseudomonas aeruginosa*. This study shows that plasmids can specialize to a novel host through trade-offs between improved stability in the new host and the ability to replicate in a previously permissive host.**

*The ISME Journal* (2010) 4, 1568–1580; doi:10.1038/ismej.2010.72; published online 3 June 2010

**Subject Category:** Evolutionary genetics

**Keywords:** plasmid; host range; evolution; replication; stability; copy number

## Introduction

Plasmids are important agents of bacterial adaptation, as they are often transferable between phylogenetically distinct hosts and code for a variety of ecologically important phenotypic traits (Frost *et al.*, 2005; Thomas and Nielsen, 2005; Schlüter *et al.*, 2007). As a result, they are responsible for the rapid spread of antibiotic resistance, a severe problem in

human health care that increasingly restricts our ability to effectively treat serious, sometimes life-threatening infections (Levy and Marshall, 2004; Strahilevitz *et al.*, 2009). In spite of the critical role of plasmids in public health care, there is a dearth of information with regard to the range of hosts in which these plasmids can replicate and be stably maintained, and how that host range might evolve over time.

Plasmids with a broad host range (BHR) are important mediators of gene transfer between distantly related bacteria (Guiney, 1984; Thomas and Helinski, 1989). In contrast to the so-called narrow-host-range ‘specialist’ plasmids (Guiney, 1982), BHR plasmids have evolved various strategies that enable them to replicate autonomously in a much wider range of phylogenetically distinct hosts (Thomas and Helinski, 1989; Toukdarian, 2004). Well-known examples of such ‘generalist’ plasmids are those of the incompatibility group P-1 (IncP-1), which can transfer to and replicate in strains of  $\alpha$ -,  $\beta$ - and  $\gamma$ -proteobacteria (Guiney, 1984; Thomas and Helinski, 1989; Schlüter *et al.*, 2007).

Correspondence: Eva M Top, Department of Biological Sciences, Initiative for Bioinformatics and Evolutionary Studies, University of Idaho, 252 Life Science South, PO Box 443051, Moscow, Idaho 83844-3051, USA.

E-mail: evatop@uidaho.edu

<sup>3</sup>Current address: Research & Development Center, NAGASE Co. & LTD., 2-2-3 Murotani, Nishi-ku, Kobe 651-2241, Japan.

<sup>4</sup>These authors contributed equally to this work.

<sup>5</sup>Current address: Department of Cell Biology, Microbiology and Molecular Biology, Center for Biomolecular Identification and Targeted Therapeutics, University of South Florida, 3720 Spectrum Boulevard, Suite 321, Tampa, Florida 33612.

Received 14 December 2009; revised 19 April 2010; accepted 23 April 2010; published online 3 June 2010

Several studies have analyzed the effects of point mutations, deletions or transposon insertions generated *in vitro* on the host range of plasmids. For example, single amino-acid changes in either the replication initiation protein RepA of the narrow-host-range *Pseudomonas* plasmid pPS10 or in *Escherichia coli* DnaA were shown to expand the plasmid's replication range (Maestro *et al.*, 2002, 2003). A key factor was the ability of the plasmid Rep protein to productively interact with proteins of the host. The prototype BHR IncP-1 $\alpha$  plasmid RK2 (= RP4) encodes both a long and short version of the Rep protein TrfA, which differ by an additional 98 amino acids at the N-terminus of the longer form (Shingler and Thomas, 1984). The short form (TrfA-33, also referred to as TrfA2) has been shown to be responsible for the core function of replication initiation and plasmid copy number control, and is sufficient for plasmid replication in strains of *E. coli* and *Pseudomonas putida* (Durland *et al.*, 1990; Lin and Helinski, 1992; Blasina *et al.*, 1996; Konieczny and Helinski, 1997). In contrast, plasmid replication in *Pseudomonas aeruginosa* requires the long form of the protein (TrfA-44, hereafter called TrfA1; Durland and Helinski, 1987; Fang and Helinski, 1991). Collectively, these findings suggest that small changes in key genes allow plasmids to shift or expand their host range. However, it is not known whether these kinds of *in vitro*-generated genetic variants would naturally arise and successfully compete in bacterial populations.

The range of hosts in which a BHR plasmid can be stably maintained in the absence of positive selection for the plasmid was found to be narrower than its replication range. We have coined the term 'long-term host range' to reflect the range of hosts in which plasmids can be maintained for at least 100 generations in the absence of selection (De Gelder *et al.*, 2007). As persistence of plasmids during periods without selection likely determines their fate in bacterial communities, we will hereafter refer to this long-term host range simply as the plasmid host range. There are at least five factors that can affect the stable maintenance of a plasmid: (i) efficient plasmid replication, (ii) stable plasmid inheritance during cell division through effective partitioning, (iii) postsegregational killing efficiency, (iv) plasmid 'cost' to host fitness and (v) reinfection of plasmid-free cells by transmissible plasmids (Bergstrom *et al.*, 2000; Slater *et al.*, 2008). Expansion, contraction or shift in plasmid host range may result in new avenues for plasmids and the traits they encode. In a worst-case scenario, the range of pathogens in which current multidrug resistance or virulence plasmids are stably maintained could thus expand in the future. Therefore, we need a better understanding of the evolutionary mechanisms of plasmid host range shifts.

Previous studies have shown that the fitness cost of plasmids to their host can be attenuated over evolutionary time by genetic change(s) in the host

chromosome, the plasmid or both (Bouma and Lenski, 1988; Modi and Adams, 1991; Modi *et al.*, 1991; Lenski *et al.*, 1994; Turner *et al.*, 1998; Dahlberg and Chao, 2003; Turner, 2004; Dionisio *et al.*, 2005; Heuer *et al.*, 2007; De Gelder *et al.*, 2008). However, in most cases, the molecular basis of plasmid–host adaptation remains unexplained, and it is unknown whether the host range of the evolved plasmids shifted simultaneously. In this study, we examined whether and by which evolutionary mechanisms an IncP-1 plasmid can expand, contract or shift its host range. To accomplish this, we experimentally evolved four hosts that carried the same model IncP-1 mini-replicon in the presence of selection for the plasmid, and compared the stability of evolved and ancestral plasmids in evolved, ancestral and other naive hosts. The plasmid became more stable in all coevolved hosts after 1000 generations. In one host, *Shewanella oneidensis* MR-1, the improved stability was always due to various kinds of single genetic changes in the N-terminal region of the plasmid-encoded TrfA1 protein. These mutations were under strong selection during plasmid–host specialization, and resulted in a shift in host range.

## Materials and methods

### *Bacterial strains, plasmids and media*

Bacterial strains and plasmids used in this study are listed in Table 1. All strains were cultured at 30 °C in Luria–Bertani broth (LB) or LB agar (Ausubel *et al.*, 1988). Kanamycin (Km) was added to the media at a final concentration of 100 mg l<sup>-1</sup>. Plasmid pMS0506 (Figure 1) was constructed as follows. The cryptic IncP-1 $\beta$  plasmid, pBP136 (Kamachi *et al.*, 2006), was digested with endonucleases *Sph*I and *Xba*I. The resulting 11.6-kb fragment carrying the genes for plasmid control and maintenance, but not those for plasmid transfer, was ligated with a Km resistance gene from the cloning vector pUC4K (Taylor and Rose, 1988) and with the *oriT* sequence of the IncP-1 $\alpha$  plasmid RP4 (Pansegrau *et al.*, 1994), both of which were amplified by PCR with appropriate primers. This plasmid, which has a total size of 13 117 bp (see Supplementary material for DNA sequence), was then introduced into several strains by transformation.

### *Evolution experiments*

Two experimental evolution protocols were used. The 'non-host-switching protocol' was carried out as follows. The four host strains, *Acinetobacter baumannii* ATCC 19606, *Pseudomonas koreensis* R28, *P. putida* H2 and *S. oneidensis* MR-1, were first grown in LB medium for 100 generations as described previously (De Gelder *et al.*, 2008). Hereafter, the strains are referred to as 19606, MR-1, H2 and R28, respectively. A single colony was

**Table 1** Bacterial strains and plasmids used in this study

Name	Relevant characteristics	References
<b>Strains</b>		
<i>Agrobacterium tumefaciens</i> C58	$\alpha$ -proteobacterium	Wood <i>et al.</i> (2001)
<i>Acinetobacter baumannii</i> ATCC 19606	$\gamma$ -proteobacterium, type strain, low G+C content (ca. 40%)	ATCC <sup>a</sup>
<i>P. aeruginosa</i> KG2512	PAO1 (ATCC 15692) <i>aph</i> , deletion in the aminoglycoside 3'-phosphotransferase type IIb ( <i>aph</i> ) gene, high sensitivity to kanamycin or neomycin	N. Goto, Kyoto Pharm. University, Japan
<i>P. koreensis</i> R28	$\gamma$ -proteobacterium, unstable maintenance of IncP-1 plasmids	De Gelder <i>et al.</i> (2007)
<i>P. putida</i> H2	$\gamma$ -proteobacterium, unstable maintenance of IncP-1 plasmids	Heuer <i>et al.</i> (2007)
<i>P. alcaligenes</i> JCM 5967	$\gamma$ -proteobacterium, type strain	JCM <sup>b</sup>
<i>Cupriavidus necator</i> JMP228	$\beta$ -proteobacterium, cured derivative of JMP134	Top <i>et al.</i> (1995)
<i>Shewanella oneidensis</i> MR-1	$\gamma$ -proteobacterium, type strain, low G+C content (ca. 47%)	ATCC
<i>Sinorhizobium meliloti</i> RM1021	$\alpha$ -proteobacterium	Diaz <i>et al.</i> (1994)
Evo-Sh4	<i>S. oneidensis</i> MR-1 derivative isolated from lineage 4 of the non-host-switching protocol after 1000 generations.	This study
Evo-Sh14	<i>S. oneidensis</i> MR-1 derivative isolated from lineage 4 of the host-switching protocol after 1000 generations	This study
<b>Plasmids</b>		
pBP136	IncP-1 $\beta$ , no accessory genes with known phenotypes	Kamachi <i>et al.</i> (2006)
pMS0506	Km <sup>r</sup> , <i>oriT</i> , pBP136 derivative lacking transfer genes	This study
pEvo-Pk1 <sup>c</sup>	pMS0506 derivative isolated from a <i>P. koreensis</i> R28 clone from lineage 1.	This study
pEvo-Pp4 <sup>c</sup>	pMS0506 derivative isolated from a <i>P. putida</i> H2 clone from lineage 4.	This study
pEvo-Sh1 <sup>c</sup>	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 1, evolved under the non-host-switching protocol; 129-bp deletion in <i>trfA</i> (position 97-225) <sup>d</sup> .	This study
pEvo-Sh2	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 2, evolved under the non-host-switching protocol; identical to pEvo-Sh1.	This study
pEvo-Sh3	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 3, evolved under the non-host-switching protocol; 231-bp deletion in <i>trfA</i> (position 24-254).	This study
pEvo-Sh4	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 4, evolved under the non-host-switching protocol; identical to pEvo-Sh1.	This study
pEvo-Sh5	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 5, evolved under the non-host-switching protocol; G-to-C mutation at position 92 in <i>trfA</i> , resulting in R31P.	This study
pEvo-Sh11	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 1, evolved under the host-switching protocol; G-to-A mutation at position 73 in <i>trfA</i> , resulting in A25 T.	This study
pEvo-Sh13	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 3, evolved under the host-switching protocol; 15-bp deletion in <i>trfA</i> (position 93-107).	This study
pEvo-Sh14	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 4, evolved under the host-switching protocol; 688-bp duplication in <i>trfA</i> (position 228-915).	This study
pEvo-Sh15	pMS0506 derivative isolated from a <i>S. oneidensis</i> MR-1 clone from lineage 5, evolved under the host-switching protocol; 13-bp duplication in <i>trfA</i> (position 210-232).	This study
pUC4K	Ap <sup>r</sup> , Km <sup>r</sup> , cloning vector	Taylor and Rose (1988)
pSTV28	Cm <sup>r</sup> , cloning vector	Takara Bio Inc., Shiga, Japan
pHY923	pSTV28 carrying a 1.6 kb <i>kleE</i> gene region of pBP136, and a 2.4 kb <i>atpB</i> gene region of the <i>S. oneidensis</i> MR-1 chromosome.	This study

<sup>a</sup>American Type Culture Collection.

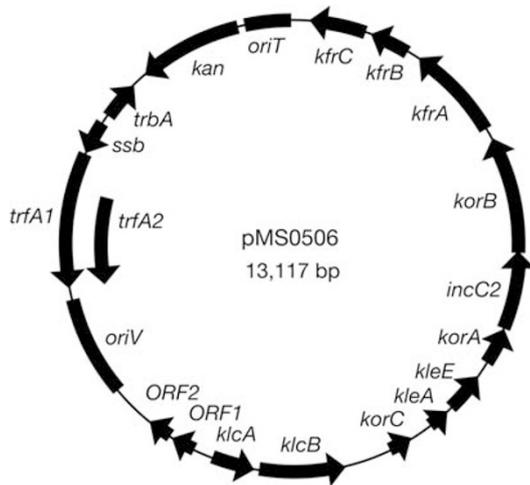
<sup>b</sup>Japan Collection of Microorganisms.

<sup>c</sup>All evolved plasmids were obtained from clones isolated after 1,000 generations (100 days).

<sup>d</sup>In this study nucleotide position 1 was defined as the first nucleotide of the *trfA1* start codon. The *trfA* gene has two overlapping in-frame open reading frames, *trfA1* (position 1-1224) and *trfA2* (position 370-1224).

isolated from each strain, and plasmid pMS0506 was introduced into the isolates by electroporation. The presence of plasmid in the transformants was confirmed by gel electrophoresis of the cell lysates,

except for strain H2, in which it was confirmed by PCR targeting the *trfA* gene because of high plasmid instability in this strain. One resulting transformant from each strain was used to inoculate five tubes



**Figure 1** Map of plasmid pMS0506 (13 117 bp). This plasmid was derived from the IncP-1 $\beta$  plasmid pBP136 (Kamachi *et al.*, 2006) by removing the transfer and multimer-resolution genes and inserting an *oriT* site and a Km resistance gene. See Materials and methods for further details. The *trfA1*, *trfA2* and *ssb* genes are involved in replication; *incC2* and *korB* are required for partitioning; other maintenance functions are encoded by *kleA*, *kleC*, *klcA*, *klcB*, *korC*, *kfrB* and *kfrC*, whereas *korA*, *kfrA* and *trbA* are responsible for transcriptional regulation. ORF1 and ORF2 code for unknown functions.

containing 5 ml LB with Km to select for the plasmid (LB (Km)). The cells were cultured for 24 h (defined as generation 0), and after every subsequent period of 24 h, approximately 4.88  $\mu$ l of each culture was transferred into 5 ml fresh LB (Km) for a total of 100 days (1000 generations). This ratio of volumes was chosen because it results in a 1026-fold dilution, allowing 10 cell doublings in a 24 h time period ( $2^{10} = 1024$ ).

To promote plasmid evolution, rather than host evolution or plasmid–host coevolution, strain MR-1 (pMS0506) was also cultured under a ‘host-switching protocol’, wherein the plasmids were transformed back into the ancestral strain every 100 generations. Each time, plasmid DNA was extracted from 1.8 ml of the replicate-evolving populations and dissolved in 10  $\mu$ l sterile water. All plasmid DNA was used to transform ancestral MR-1 cells by electroporation, and all transformants (typically 450–1800) were cultured in LB (Km). The MR-1 host-switching protocol was carried out with four independent replicate lineages (Sh11, Sh13, Sh14 and Sh15, corresponding to isolated plasmids pEvo-Sh11, pEvo-Sh13, pEvo-Sh14 and pEvo-Sh15); all other protocols had five replicate lineages.

#### Stability experiments

Plasmid stability experiments were carried out as described previously (De Gelder *et al.*, 2008), but using the same serial dilution regime as in the evolution experiment. The 1000-generation-old populations of the four strains were used to test the stability of evolved plasmids in their coevolved

hosts. To investigate the stability of evolved plasmids in ancestral hosts, a single colony of a coevolved host was selected randomly to extract the residing plasmid; the plasmid was then introduced into the ancestral host by electroporation, and five transformants were tested. The same approach was used to test the stability of pMS0506 in two evolved plasmid-free clones of *S. oneidensis*, Evo-Sh4 and Evo-Sh14, and of pMS0506 and two evolved plasmids in four naive hosts. Evolved plasmid-free clones were obtained by screening the corresponding 1000-generation-old cultures for Km-sensitive clones. Logistic regression with a model selection approach (Agresti, 2002) was used to determine significant differences between stability patterns.

#### Growth rate analysis

The maximum growth rate of ancestral host MR-1 containing an evolved plasmid, and of MR-1 with ancestral plasmid pMS0506, was compared as follows. The cultures, grown overnight in LB (Km) to stationary phase, were diluted 100-fold in 100  $\mu$ l of fresh medium in a 96-well microtiter plate, which was incubated at 30  $^{\circ}$ C with agitation for 24 h in a BIO-TEK Power Wave HT shaker incubator (BioTek instruments Inc., Winooski, VT, USA) (eight replicates per clone). The optical density at 600 nm ( $OD_{600}$ ) in each well was recorded every 10 min, and the maximum growth rate ( $r_{MAX}$ ) was determined using local polynomial regression to the growth curves in R (referred to as ‘loess’), and then calculating the maximum slope of the fitted smooth line (Cleveland *et al.*, 1992; R-Development-Core-Team, 2009). In this study, we defined the relative fitness simply as the ratio of maximum growth rates of host MR-1 with evolved versus ancestral plasmid:  $r_{MAX} [MR-1(pEvo)]/r_{MAX} [MR-1(pMS0506)]$ , with pEvo being one of the seven unique evolved plasmid genotypes.

#### DNA methodology

Standard methods were used for extraction of plasmid DNA, DNA digestion with restriction endonucleases, ligation, gel electrophoresis and transformation of *E. coli* cells (Ausubel *et al.*, 1988). To amplify partial *trfA* sequences, colony PCR was performed on 10–15 colonies that were randomly selected from each of the nine lineages at generation 1000. Each colony was suspended in 100  $\mu$ l sterile distilled and deionized water and 1  $\mu$ l of this suspension was used in PCR reactions (25  $\mu$ l total volume), which were carried out with 2X PCR Master Mix (Fermentas, International Inc., Burlington, Ontario, Canada). The primers used were *trfA1f*, 5'-ATGACGAACAACGAGTTCAACGAGC-3'; *trfA1r*, 5'-CAGCACCTGGGCGAATACCA-3'; *trfAsw4f*, 5'-GTGTCGCTGATCCGGCGCTT-3'; *trfAsw4r*, 5'-AGAGCTGCATGTCCTTGATCT-3'. The latter two were used to confirm the large duplication found in pEvo-Sh14

by amplifying a 308-bp region that spanned the junction between the duplicated sequences. The PCR-amplified DNA fragments were purified using a GeneJET PCR Purification Kit (Fermentas) according to the manufacturers' protocols.

The nucleotide sequence of the ancestral and nine evolved plasmids (by primer walking), as well as of *trfA* amplicons, was determined with an ABI PRISM model 3730 DNA analyzer (Applied Biosystems, Foster City, CA, USA). Sequence analysis was performed with software programs GENETYX (SDC Inc., Tokyo, Japan), CLC Sequence Viewer (CLC bio A/S, Cambridge, MA, USA), FinchTV (Geospiza Inc., Seattle, WA, USA) and BLAST (National Center for Biotechnology Information, Bethesda, MD, USA).

#### Plasmid copy number analysis

SYBR-green-based quantitative PCR (qPCR) was used to determine the copy numbers of pMS0506 and its evolved variants in ancestral host MR-1. The copy number of a 158-bp sequence within the *kleE* gene of pMS0506 was defined as the plasmid copy number (*kleE* fragment), and that of a 167-bp sequence within the chromosomal gene *atpB* was used as an internal control (*atpB* fragment, located 13-kb upstream from *dnaA* in the MR-1 chromosome (Genbank accession no. NC\_004347)). Plasmid pHY923, which carried the *kleE* and *atpB* genes at a 1:1 ratio, was used as a copy number reference (Table 1).

The qPCR reactions were performed at a 20  $\mu$ l reaction scale using Fast SYBR Green Master Mix and the 7900HT Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). The primers used were *kleEF*, 5'-CGTCTCGATATAC AAGCCCA-3'; *kleER*, 5'-TTGCACCCATAACACTC CA-3'; *atpBF*, 5'-GTTGGTTCCATACATCAGCG-3'; *atpBR*, 5'-AACTGACGCTACAGCCCTT-3'. Data were analyzed using the SDS 2.3 software (Applied Biosystems) provided with the instrument.

Total DNA of MR-1 clones digested by restriction enzymes was used as qPCR template to eliminate PCR bias caused by the topology of DNA (Providenti *et al.*, 2006; Chen *et al.*, 2007). Total DNA was isolated from cultures in the exponential growth phase ( $OD_{600} = 0.6$ ) using the GenElute Bacterial genomic DNA Kit (Sigma-Aldrich, St Louis, MO, USA). Purified DNA was digested by *EcoRI* and *PstI*, and precipitated with ethanol before PCR. Serial dilutions of digested pHY923 were also subjected to qPCR to generate a standard curve for both the *kleE* and *atpB* fragments. The copy number of the two gene fragments in total DNA was then calculated on the basis of the Ct value and the standard curve ( $R^2 > 0.997$ ); for each qPCR plate, a standard curve was generated for each gene fragment. Absence of nonspecific amplification was confirmed by analyzing the dissociation curves. The plasmid copy number was calculated as the ratio of the copy numbers of *kleE* and *atpB*.

#### Prediction of protein structure

The secondary structure of TrfA1 was predicted using NNPREPDICTION (<http://www.cmpharm.ucsf.edu/~nomi/nnpredict.html>) (Kneller *et al.*, 1990) and PredictProtein (<http://www.predictprotein.org/>) (Rost *et al.*, 2004). A tertiary structure class of 'none' was used for the prediction. The presence of intrinsically unstructured regions of full-length TrfA1 was predicted using IUPred (<http://iupred.enzim.hu/>) (Dosztanyi *et al.*, 2005).

## Results

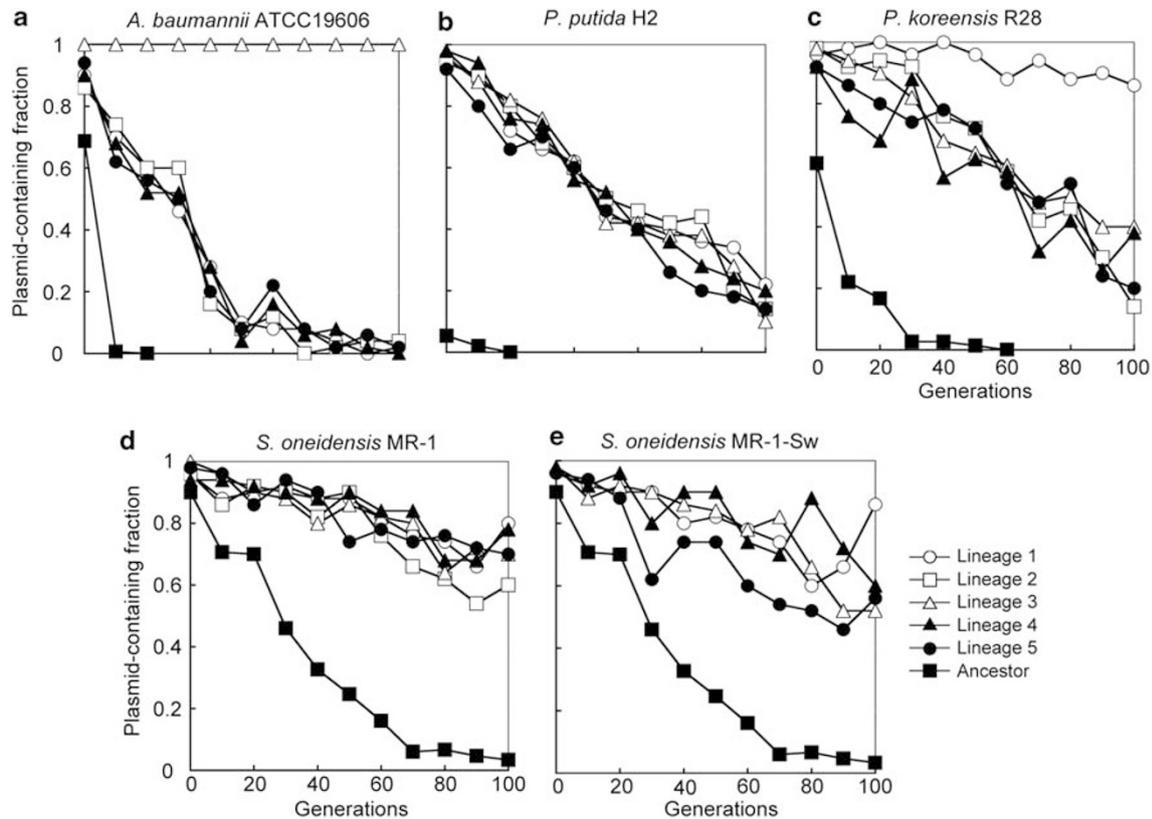
#### Improved stability of evolved plasmids in their coevolved hosts

One of the ways by which a plasmid can adapt to a novel host and thereby shift, contract or expand its host range is by increasing its stability. To investigate whether and how the stability of an IncP-1 plasmid can improve after long-term maintenance in different hosts under antibiotic selection, we experimentally evolved four strains (*A. baumannii* ATCC 19606, *P. koreensis* R28, *P. putida* H2 and *S. oneidensis* MR-1) after introducing the IncP-1 $\beta$  mini-replicon pMS0506 (Figure 1). Hosts were chosen because (i) plasmid pMS0506 was unstable in all these strains (Figure 1, filled squares), in contrast to its high stability in *E. coli* EC100 (data not shown); (ii) strains 19606 and MR-1 have a much lower GC content than the plasmid and are important species from a clinical and bioremediation perspective; and (iii) plasmid adaptation to strains H2 and R28 has been studied by us previously (Heuer *et al.*, 2007; De Gelder *et al.*, 2008). The use of a mini-replicon instead of a wild-type plasmid allowed cost-efficient resequencing of multiple plasmids from independently evolved lineages and allowed us to focus on plasmid replication and stable inheritance without the confounding effects of plasmid transfer. After 1000 generations, the stability of plasmids in their coevolved hosts was determined.

For all strains and protocols, the evolved plasmids were much more stable in their coevolved hosts than the ancestral plasmid in the respective ancestral hosts, and the plasmid stability patterns were quite similar between lineages of a given strain, even under the two different protocols (Figure 2). Plasmid evolution in strain 19606 was not further investigated because the plasmid was no longer present in the evolved lineage that showed stable maintenance of Km resistance (Figure 2a), and because the other four lineages contained only small plasmids (5–8 kb), suggesting large deletions. Overall, the improved plasmid stability in all four evolved hosts suggests that the plasmid, the host or both had undergone evolutionary changes.

#### Plasmid or host adaptation caused improved plasmid stability

To determine whether the improved stability of the evolved plasmids in their coevolved hosts was, at

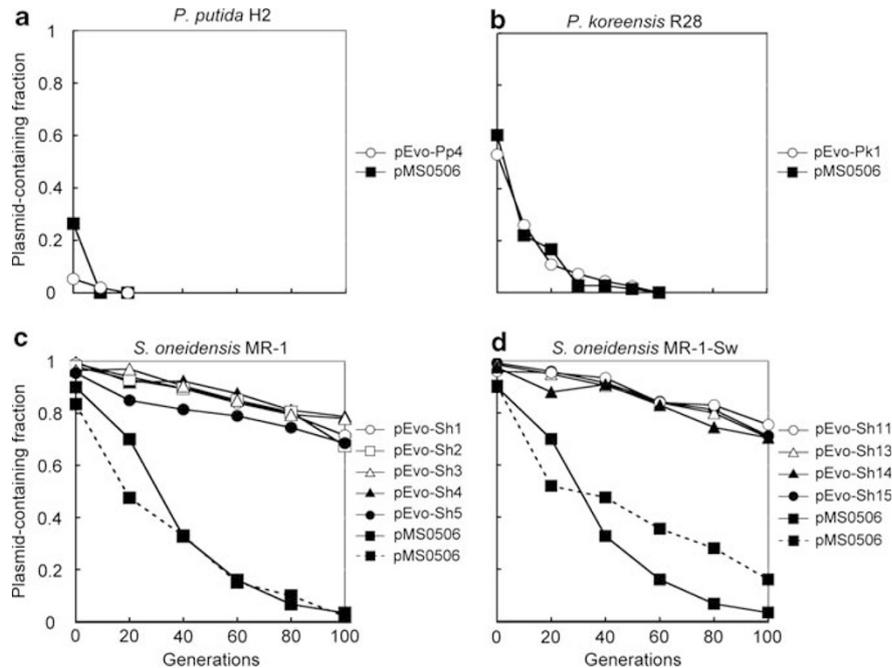


**Figure 2** Stability of evolved plasmids in their coevolved hosts in comparison with the ancestral plasmid pMS0506 in the ancestral host (ancestor). Panels **a–d** represent results from populations of four different strains containing pMS0506, evolved for 1000 generations using the non-host-switching protocol, whereas panel **e** represents populations of strain MR-1 that were evolved using the host-switching protocol.

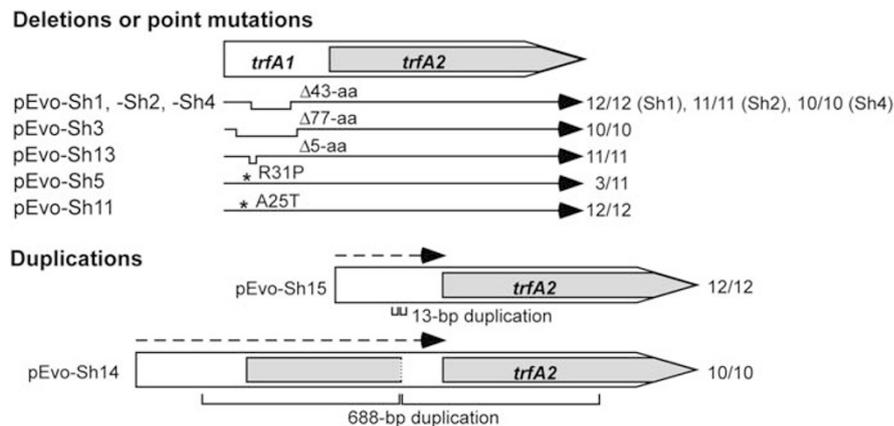
least in part, due to plasmid evolution, the stability of the evolved plasmids was determined in the corresponding ancestral host (except for strain 19606). The evolved plasmids from strains H2 and R28 were nearly as unstable as the ancestral plasmid (Figure 3a–b), suggesting that their improved stability in the coevolved hosts (Figure 2b–c) was not due to genetic changes in the plasmid alone, but must be due to either host evolution or plasmid–host coevolution. In contrast, plasmids from strain MR-1 evolved by either of the two protocols were much more stable in the ancestral host than in the ancestral plasmid (Figure 3c–d), strongly indicating plasmid evolution. To confirm the latter finding, the stability of one randomly chosen evolved plasmid from each of the nine MR-1 lineages was tested. The stability patterns for all nine plasmids were very similar and there were no significant differences between clones or between protocols (Figure 3c–d). To examine whether host MR-1 had adapted to the plasmid at all, the stability of the ancestral plasmid was tested in plasmid-free segregants of two evolved clones. The ancestral plasmid was as unstable in these evolved hosts as in the ancestral host (see pMS0506<sup>b</sup> in Figure 3c–d). Together, these findings strongly suggest that the improved stability of the plasmids evolved in MR-1 was solely due to mutations in the plasmid.

#### *Mutations in the evolved plasmids that conferred improved stability*

To identify the genetic changes that had occurred in the nine stable plasmids from independently evolved lineages of strain MR-1, their complete genome sequences were determined and compared with that of pMS0506. Interestingly, each plasmid contained only a single genetic change that always occurred in the 5′-end region of *trfA*, which codes for the N-terminus of the long replication initiation protein TrfA1 (Fang and Helinski, 1991; Thorsted *et al.*, 1998). The reading frame for the shorter protein form, TrfA2, was never affected. Deletions were observed in five of the nine plasmids. Three plasmids showed an identical 129-bp deletion, one plasmid showed a 231-bp deletion and one a small 15-bp deletion (Table 1, Figure 4). In each case, these were in-frame deletions, and translation should result in TrfA1 proteins with deletions of 43, 77 and 5 amino acids (aa) in the N-terminal region ( $\Delta 43$ ,  $\Delta 77$ ,  $\Delta 5$  in Figure 4). The two larger deletions occurred between two short DNA sequence repeats (5′-GCCGCCG and 5′-GCAGGAAT, respectively). Single point mutations were found in two other plasmids: a G-to-C and G-to-A mutation at nucleotides 92 and 73 of the *trfA* gene, respectively, resulting in amino acid changes R31P and A25T in



**Figure 3** Stability of evolved and ancestral plasmids in their respective ancestral hosts (panels a–d). For strain MR-1, plasmids from all lineages were tested, whereas for strains H2 and R28, only one evolved plasmid from the lineage with the highest stability was tested. Later sequencing efforts showed that pEvo-Sh1, Evo-Sh2 and pEvo-Sh4 are identical. The stability of the ancestral plasmid in plasmid-free segregants from lineage 4 of both protocols was also tested (filled squares with dashed lines in panels c and d). Data represent averages from at least three replicate stability assays. Standard deviations are not shown for clarity.



**Figure 4** Effect of mutations found in the evolved plasmids on the TrfA1 protein. The pentagons represent the two *trfA* genes that are translated from different start codons, and the solid and dashed arrows represent the corresponding TrfA1 gene products. Solid arrows indicate putative TrfA1 proteins translated from the evolved *trfA* genes with deletions and point mutations. These proteins are expected to differ from the ancestral protein only in terms of the missing or changed amino acids that correspond to the deleted gene regions (indentations,  $\Delta 43$ ,  $\Delta 77$ ,  $\Delta 5$ ) and point mutations (asterisks, R31P and A25T). Dashed arrows show the putative altered TrfA1 products that result from frameshifts caused by duplications in the *trfA* gene. Brackets below the genes show the duplicated DNA sequences. For more details about the mutations, see the Results and Table 1. The fractions shown to the right of the TrfA proteins represent the ratios ( $x/y$ ) of the number of clones with the same DNA sequence ( $x$ ) and the total number of clones tested ( $y$ ).

the TrfA1 protein. Finally, in two of the nine evolved plasmids, both from the host-switching protocol, sequence duplications in the *trfA* gene had occurred (Table 1, Figure 4). These duplications resulted in a shift of the reading frame that introduced stop codons at positions 1060 and 376, which is expected to result in truncation of TrfA1. As the complete TrfA2 reading frame remained intact, the ancestral TrfA2 protein must now be the only Rep protein

encoded by these plasmids. Thus, all mutations that occurred in the nine evolved plasmids only affected the sequence of TrfA1 and not of TrfA2, suggesting that these different genetic changes in the TrfA1 N-terminus dramatically improved the stability of the plasmid.

To determine whether these sequenced plasmid genotypes were dominant in the populations, we determined the sequence of the 5' end of

*trfA* for 10–12 randomly chosen clones per lineage. In eight of the nine lineages, every clone exhibited the mutation observed in the corresponding plasmid that was completely sequenced (Figure 4). The one exception was lineage 5 from the nonhost-switching protocol. Here, only 3 of 11 clones had the observed R31P point mutation found in plasmid pEvo-Sh5, whereas the remaining eight showed the same 129-bp deletion seen in three other plasmids. Overall, these results indicate that with the exception of pEvo-Sh5, the completely sequenced plasmids represented the numerically dominant plasmids in the evolved populations. Thus, the corresponding *trfA* mutations were sufficiently beneficial to become dominant in the population within 1000 generations or less.

#### Host fitness effects and copy number of evolved and ancestral plasmids

As MR-1 clones with evolved pMS0506 variants outcompeted cells containing the ancestral plasmid, the evolved plasmids must confer a fitness advantage to their hosts. To investigate this, we compared one of the components of fitness, the maximum growth rate (Stevenson and Schmidt, 2004). This was carried out for ancestral host MR-1, harboring either one of the evolved plasmid variants or pMS0506 in LB (Km) medium. The relative fitness, represented by the growth rate ratio, was always higher than 1 (Table 2), indicating that even in the ancestral host evolved plasmids conferred a significant fitness advantage in the presence of Km. As all mutations were found in *trfA1*, the improved persistence and fitness effects of evolved plasmids in host MR-1 could be due to an increase in average plasmid copy number per cell. Indeed, all five pMS0506 variants with in-frame mutations showed

a higher copy number than ancestral pMS0506 (Table 2). Surprisingly, the copy numbers of the two plasmids with frame-shift mutations (duplications) were not different from that of the ancestral plasmid, suggesting a different molecular mechanism of stability improvement.

#### Host shifts of evolved pMS0506 variants

To determine whether the adaptation of pMS0506 to host MR-1 came with a trade-off and resulted in a host shift, we tested the ability of two evolved plasmids to replicate and be stably maintained in five diverse strains belonging to the  $\alpha$ -,  $\beta$ - and  $\gamma$ -proteobacteria (Figure 5, Table 1). The two plasmids tested were deletion mutant pEvo-Sh4 and plasmid pEvo-Sh11 with a point mutation (Table 1). There was a drastic effect of mutations on replication in *P. aeruginosa* KG2512, a Km-sensitive PAO1 (ATCC15692) derivative. Transformants harboring either of the two evolved plasmids could never be obtained, whereas the strain could be transformed with the ancestral plasmid. In contrast, evolved plasmids replicated in the other four strains and there was no effect on plasmid stability (Figure 5). The *P. aeruginosa* assay was further extended with all other deletion mutants, as well as with the second point mutation variant and one duplication mutant. Again, none of these evolved plasmids could replicate in this host. These findings suggest that the genetic changes in *TrfA1* that allowed improved stability of the plasmid in MR-1 did not affect stable plasmid replication in several phylogenetically distinct hosts, but prevented replication in *P. aeruginosa*. Thus, by adapting to one host, these plasmids lost their ability to replicate in a previously permissive host, a clear example of trade-off and shift in host range.

**Table 2** Plasmid copy number and relative fitness

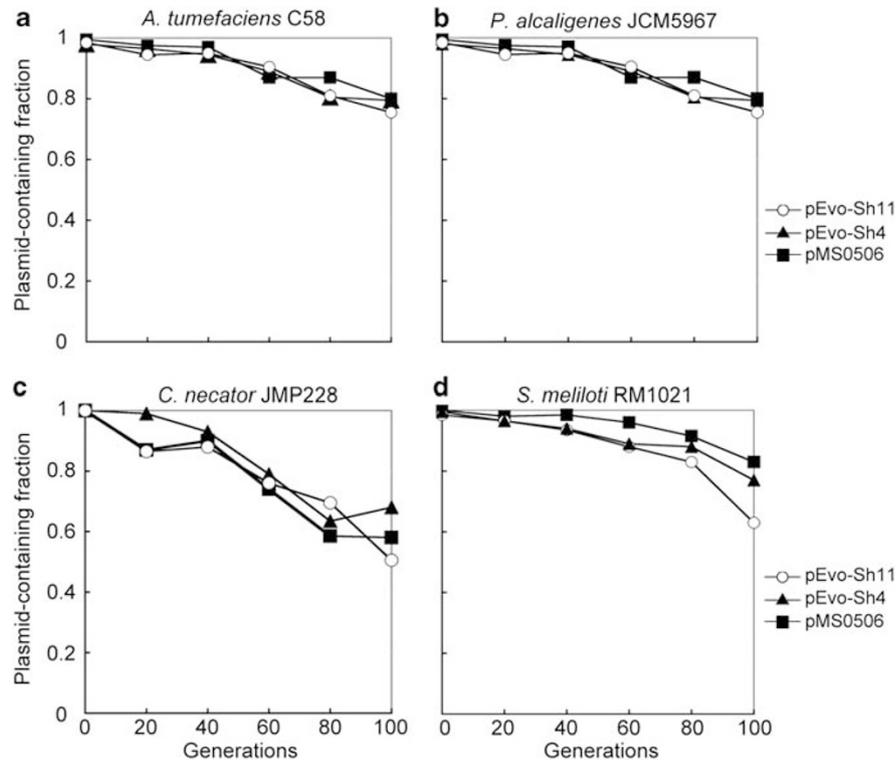
Plasmid	Plasmid copy number <sup>a</sup>	Relative fitness of host MR-1 <sup>b</sup>
pMS0506	3.35 ± 0.65	1.0000
pEvo-Sh1, -Sh2, -Sh4	4.69 ± 0.90*	1.0828***
pEvo-Sh3	4.82 ± 0.26*	1.0929***
pEvo-Sh5	5.66 ± 0.27***	1.0530***
pEvo-Sh11	6.46 ± 0.20***	1.0905***
pEvo-Sh13	4.99 ± 1.67**	1.0968***
pEvo-Sh14	3.38 ± 0.05	1.0593***
pEvo-Sh15	3.26 ± 0.19	1.0643***

<sup>a</sup>Plasmid copy number indicates the ratio of *kleE* to *atpB* in total DNA. Mean and standard deviation are shown. Five replicate samples were analyzed for pMS0506, pEvo-Sh1,-Sh2,-Sh4 and pEvo-Sh13, and three for pEvo-Sh3, pEvo-Sh5, pEvo-Sh11, pEvo-Sh14 and pEvo-Sh15.

<sup>b</sup>Relative fitness is represented by the ratio of the maximum growth rates in LB (Km):  $r_{MAX} [MR-1(pEvo-...)]/r_{MAX}[MR-1 (pMS0506)]$ . Absolute values for the evolved plasmids were compared to those for pMS0506 using Dunnett's test Dunnett (1964); \*, \*\*, \*\*\*: significantly different from the value obtained for pMS0506 ( $P < 0.10$ ,  $P < 0.05$ , and  $P < 0.01$ , respectively).

## Discussion

Adaptation of promiscuous plasmids to new hosts by improving their stability may be essential for their long-term persistence in bacterial communities of clinical and natural habitats. As this process may be detrimental to humans in the case of multidrug resistance or virulence plasmids, it is imperative that we improve our understanding of the mechanisms by which plasmids adapt to novel hosts, and thereby shift, contract or expand their host range. In this study, we showed that in one particular host, the 5'-end region of the *trfA1* gene of an IncP-1 plasmid, encoding the N-terminus of the plasmid replication protein *TrfA1*, was under strong selection during plasmid–host adaptation. Mutations in this region were the only evolutionary changes observed in the entire ~13-kb genomes of independently evolved plasmids after 1000 generations, and were responsible for the improved plasmid stability. We showed that most of the mutations affected



**Figure 5** Stability in four naive hosts (panels a–d) of plasmids evolved in *S. oneidensis* MR-1. Stability of pEvo-Sh4 (129-bp deletion in *trfA*) and pEvo-Sh11 (A25T point mutation) was compared with that of the ancestral plasmid pMS0506. The fraction of plasmid-containing cells at each time point is an average of four replicate stability assays. Standard deviations are not shown for clarity.

plasmid copy number and resulted in a clear shift in host range. Thus, under certain conditions, BHR plasmids can specialize to a long-term host and thereby experience trade-offs between stable maintenance in that host and the ability to replicate in others.

Our study showed that improved persistence of a particular plasmid can be achieved through different evolutionary pathways. First, host evolution or plasmid–host coevolution explained the improved stability of pMS0506 in two of the four strains examined. In contrast, the improved stability of the same plasmid in host MR-1 was explained by genetic changes in the plasmid alone, and more specifically in the same locus. Whereas the nine evolved plasmids showed similar stability patterns (phenotypes), seven of them had unique single mutations in the 5' region of *trfA1*. This indicates parallel evolution at the gene level (also referred to as parallel genotypic adaptation; Wood *et al.*, 2005), but through different evolutionary mechanisms. Parallel evolution at the gene, but not at the base pair, level was also recently shown for independently evolved *E. coli* lineages (Woods *et al.*, 2006), whereas in studies of bacteriophage evolution, more parallelism at the base pair level has been found (Wichman *et al.*, 1999). Ongoing studies on the underlying molecular mechanisms of improved stability of our evolved plasmids and on the evolutionary dynamics of the MR-1 populations will help explain how and why different mutations

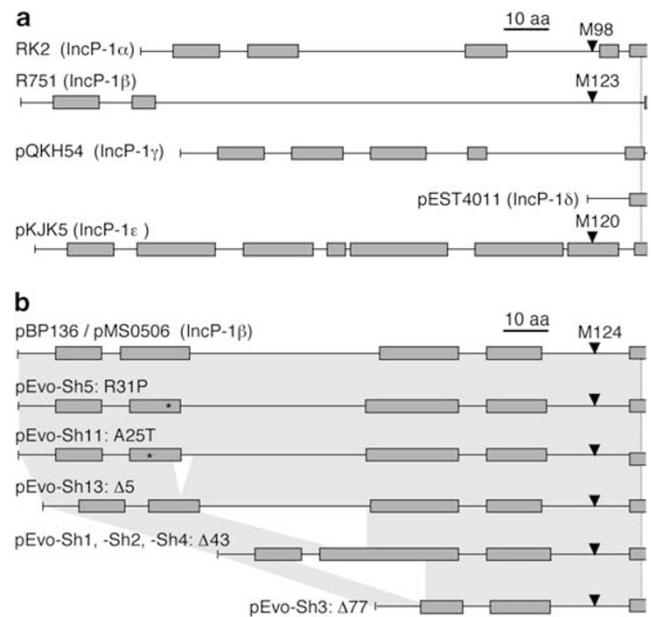
that led to similar phenotypes arose and fixed in different populations.

One of the intriguing findings of this study is the drastic improvement in plasmid stability in host MR-1 due to single changes in the region that encodes the N-terminal region of TrfA1. As described in the Introduction, there are at least five mechanisms by which a plasmid can increase its stability in the absence of positive selection. However, as the more stable plasmids evolved in host MR-1 only showed mutations in the replication initiation protein, we posit that only two of these mechanisms should be considered here: improved replication efficiency and decreased fitness cost. First, the mutations may positively affect plasmid replication by increasing or more tightly controlling the plasmid copy number, which may compensate for ineffective plasmid partitioning. The increased copy number of the five pMS0506 variants with in-frame mutations supports this hypothesis (Table 2). Whereas a higher plasmid copy number may negatively affect host fitness in the absence of selection, the evolution experiments reported here were performed with continuous selection for Km resistance. Under these conditions, these evolved plasmids significantly increased the growth rate of their host compared with pMS0506 (Table 2), probably because the higher copy number decreased the frequency of plasmid loss. Second, a decrease in plasmid cost may explain the two plasmids with the frameshift mutations that truncated the TrfA1

protein. The TrfA protein is known to directly interact with DnaB and DNA polymerase III, which are also required for chromosome replication (Pacek *et al.*, 2001; Jiang *et al.*, 2003; Kongsuwan *et al.*, 2006). Thus, elimination of TrfA1 may have increased the pool of essential host proteins, thus decreasing the plasmid's negative effect on host fitness, also referred to as 'interference cost' (Modi *et al.*, 1991). We did not determine the cost of plasmid through growth, competition or modeling of stability curves (De Gelder *et al.*, 2007), because the high instability of pMS0506 would confound the measurements, and the high stability of evolved plasmids would not allow estimating the cost parameter. Future studies will focus on accurately estimating the cost in the absence of antibiotics.

An important observation in this study was the shift in plasmid host range away from *P. aeruginosa*. We tested plasmid replication in this particular species, because it has been shown that the TrfA1 N-terminus encoded by the plasmid RK2 is essential for recruiting the *P. aeruginosa* DnaB helicase to the plasmid replication origin (Caspi *et al.*, 2001; Zhong *et al.*, 2003). More specifically, deletion of amino acids in the second of four putative helical regions (Figure 6, RK2) has resulted in loss of this function (Zhong *et al.*, 2003). To begin to understand the effect of the mutations in our evolved IncP-1 $\beta$  plasmids on the structure of TrfA1, we first analyzed the propensity of the ancestral TrfA1 N-terminus encoded by pMS0506 to form secondary and tertiary structures. The 148-amino-acid N-terminal region was predicted not to form a stable tertiary structure, but to form several  $\alpha$ -helices (Figure 6), just as in the RK2 study (Zhong *et al.*, 2003). The presence of a helical structure in such an intrinsically unstructured protein region is a key signature of molecular recognition elements, that is, short sequences that fold on binding to a protein or DNA partner (Fuxreiter *et al.*, 2004; Oldfield *et al.*, 2005). Interestingly, just as in the RK2 study, the mutations in our evolved IncP-1 $\beta$  plasmid were also located in and around the second helix of TrfA1, A25–A34. Moreover, when the helix sequence was shown on a helical wheel, the two point mutations, R31P and A25T, ended up on the same face of the helix (data not shown). Thus, while specializing to *S. oneidensis* MR-1, the plasmid likely shifted its host range away from *P. aeruginosa* because of a decreased DnaB loading capacity of the TrfA1 N-terminus in that host. Such a shift in host range due to single mutations is comparable to the observations made in studies with bacteriophage. Phage adaptation to a novel host due to a single mutation often caused a reduction in fitness in other hosts, a clear case of antagonistic pleiotropy (Crill *et al.*, 2000; Duffy *et al.*, 2006).

We compared the *trfA* mutations obtained in our study with the extant diversity of TrfA proteins encoded by more than 20 completely sequenced naturally occurring IncP-1 plasmids. Although



**Figure 6** Variation in the length and predicted secondary structure of the N-terminal region of replication initiation protein TrfA for (a) natural IncP-1 plasmids and (b) ancestral plasmid pMS0506 and five of its seven characterized variants evolved in *S. oneidensis* MR-1. The secondary structure was predicted by PredictProtein (Rost *et al.*, 2004);  $\alpha$ -helices are represented by boxes. Sequences were aligned on the dotted line at the right side of the figure, which is defined as the start point of the conserved C-terminus region (L110 in RK2). The positions of the start codon of the shorter version of TrfA (TrfA-33, TrfA2) are indicated with arrowheads, as well as position x of the methionine codon (Mx). IncP-1 $\gamma$  and -1 $\delta$  subgroup plasmids do not encode TrfA2. TrfA sequences of natural plasmids were obtained from the following accession numbers: AAA16606, RK2; AAA17040, R751; CAJ43307, pQKH54; CAK02642, pKJK5; AAS49458, pEST4011; BAF33454, pBP136. In panel (b), identical sequence regions are connected by gray shades and the positions of point mutations (R31P, A25T) are indicated by asterisks. Duplication mutants pEvo-Sh14 and pEvo-Sh15 are not shown here because of the difficulty of visualizing the alignments. For more details on plasmid variants, see Figure 4 and Table 1.

the structure of the TrfA1 N-terminal region with its typical  $\alpha$ -helices is conserved in most of the sequenced IncP-1 plasmids (Figure 6), there are notable exceptions. For example, the N-terminus is completely lacking in TrfA proteins encoded by three plasmids: IncP-1 $\delta$  plasmids pEST4011 (Figure 6) and pAKD4, and the IncP-1 $\beta$  plasmid pADP-1 (Martinez *et al.*, 2001; Vedler *et al.*, 2004; Sen *et al.*, 2010). Moreover, the TrfA1 N-terminus encoded by the IncP-1 $\gamma$  plasmid pQKH54 (Haines *et al.*, 2006) is shorter and more divergent than that of typical IncP-1 $\alpha$  and - $\beta$  plasmids (Figure 6). Whether these shorter and anomalous *trfA1* gene sequences explain why both IncP-1 $\gamma$  and IncP-1 $\delta$  plasmids have a more limited host range than the typical IncP-1 $\alpha$  and - $\beta$  plasmids needs to be further investigated. These observations strongly suggest that the evolutionary changes found in TrfA encoded by our evolved plasmids are not a result of artificial cultivation conditions, or specific to our

model mini-replicon or host, but reflect the natural evolution of IncP-1 plasmids.

Our results show that in a 13-kb plasmid, a single point mutation, deletion or duplication in the plasmid replication initiation gene is sufficient to change the host range. This is in agreement with previous observations from *in vitro* mutagenesis studies (Maestro *et al.*, 2003; Zhong *et al.*, 2003). Moreover, we also showed that different mutations in the same gene can arise in independently evolving populations, and that these are sufficiently beneficial to be fixed within 1000 generations. A more general understanding of the evolutionary patterns of plasmid–host adaptation will require experimental evolution and retrospective studies with several natural plasmids and hosts under different environmental conditions, followed by molecular analyses. This will also help us to understand whether plasticity in the 5' *trfA* region is a general mechanism of host range evolution for promiscuous IncP-1 plasmids.

## Acknowledgements

We thank Dr K Kamachi (National Institute for Infectious Diseases, Japan) for his kind gift of pBP136, and Prof N. Goto (Kyoto Pharm. University, Japan) for providing us strain KG2512. We also thank G. Deckert, L. Rogers, D. Sen, N. Packer, M. Bauer, S. Bassler and S. Sax for technical assistance. This project was supported by NIH Grant number P20 RR16448 from the COBRE program of the National Center for Research Resources (NCRP).

## References

- Agresti A. (2002). *Categorical Data Analysis*, 2 edn. John Wiley and Sons: Hoboken, NJ.
- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA *et al.* (1988). *Current Protocols in Molecular Biology*. Wiley: New York, USA.
- Bergstrom CT, Lipsitch M, Levin BR. (2000). Natural selection, infectious transfer and the existence conditions for bacterial plasmids. *Genetics* **155**: 1505–1519.
- Blasina A, Kittell BL, Toukdarian AE, Helinski DR. (1996). Copy-up mutants of the plasmid RK2 replication initiation protein are defective in coupling RK2 replication origins. *Proc Natl Acad Sci USA* **93**: 3559–3564.
- Bouma JE, Lenski RE. (1988). Evolution of a bacteria/plasmid association. *Nature* **335**: 351–352.
- Caspi R, Pacek M, Consiglieri G, Helinski DR, Toukdarian A, Konieczny I. (2001). A broad host range replicon with different requirements for replication initiation in three bacterial species. *EMBO J* **20**: 3262–3271.
- Chen J, Kadlubar FF, Chen JZ. (2007). DNA supercoiling suppresses real-time PCR: a new approach to the quantification of mitochondrial DNA damage and repair. *Nucleic Acids Res* **35**: 1377–1388.
- Cleveland WS, Grosse E, Shyu WM. (1992). Local regression models. In: JM Chambers & TJ Hastie (eds). *Statistical Models*. Wadsworth Brooks/Cole: Boca Raton, Florida.
- Crill WD, Wichman HA, Bull JJ. (2000). Evolutionary reversals during viral adaptation to alternating hosts. *Genetics* **154**: 27–37.
- Dahlberg C, Chao L. (2003). Amelioration of the cost of conjugative plasmid carriage in *Escherichia coli* K12. *Genetics* **165**: 1641–1649.
- De Gelder L, Ponciano JM, Joyce P, Top EM. (2007). Stability of a promiscuous plasmid in different hosts: no guarantee for a long-term relationship. *Microbiology* **153**: 452–463.
- De Gelder L, Williams J, Ponciano J, Sota M, Top EM. (2008). Adaptive plasmid evolution results in host-range expansion of a broad-host-range plasmid. *Genetics* **178**: 2179–2190.
- Diaz E, Munthali M, de Lorenzo V, Timmis KN. (1994). Universal barrier to lateral spread of specific genes among microorganisms. *Mol Microbiol* **13**: 855–861.
- Dionisio F, Conceicao IC, Marques AC, Fernandes L, Gordo I. (2005). The evolution of a conjugative plasmid and its ability to increase bacterial fitness. *Biol Lett* **1**: 250–252.
- Dosztanyi Z, Csizmok V, Tompa P, Simon I. (2005). IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content. *Bioinformatics* **21**: 3433–3434.
- Duffy S, Turner PE, Burch CL. (2006). Pleiotropic costs of niche expansion in the RNA bacteriophage  $\phi$ 6. *Genetics* **172**: 751–757.
- Dunnnett CW. (1964). New tables for multiple comparisons with a control. *Biometrics* **20**: 482–491.
- Durland RH, Helinski DR. (1987). The sequence encoding the 43-kilodalton TrfA protein is required for efficient replication or maintenance of minimal RK2 replicons in *Pseudomonas aeruginosa*. *Plasmid* **18**: 164–169.
- Durland RH, Toukdarian A, Fang F, Helinski DR. (1990). Mutations in the *trfA* replication gene of the broad-host-range plasmid RK2 result in elevated plasmid copy numbers. *J Bacteriol* **172**: 3859–3867.
- Fang FC, Helinski DR. (1991). Broad-host-range properties of plasmid RK2: importance of overlapping genes encoding the plasmid replication initiation protein TrfA. *J Bacteriol* **173**: 5861–5868.
- Frost LS, Leplae R, Summers AO, Toussaint A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat Rev Microbiol* **3**: 722–732.
- Fuxreiter M, Simon I, Friedrich P, Tompa P. (2004). Preformed structural elements feature in partner recognition by intrinsically unstructured proteins. *J Mol Biol* **338**: 1015–1026.
- Guiney DG. (1982). Host range of conjugation and replication functions of *Escherichia coli* sex plasmid F lac: comparison with the broad host range plasmid RK2. *J Mol Biol* **162**: 699–703.
- Guiney DG. (1984). Promiscuous transfer of drug resistance in gram-negative bacteria. *J Infect Dis* **149**: 320–329.
- Haines AS, Akhtar P, Stephens ER, Jones K, Thomas CM, Perkins CD *et al.* (2006). Plasmids from freshwater environments capable of IncQ retrotransfer are diverse and include pQKH54, a new IncP-1 subgroup archetype. *Microbiology* **152**: 2689–2701.
- Heuer H, Fox RE, Top EM. (2007). Frequent conjugative transfer accelerates adaptation of a broad-host-range plasmid to an unfavorable *Pseudomonas putida* host. *FEMS Microbiol Ecol* **59**: 738–748.
- Jiang Y, Pacek M, Helinski DR, Konieczny I, Toukdarian A. (2003). A multifunctional plasmid-encoded replication

- initiation protein both recruits and positions an active helicase at the replication origin. *Proc Natl Acad Sci USA* **100**: 8692–8697.
- Kamachi K, Sota M, Tamai Y, Nagata N, Konda T, Inoue T *et al.* (2006). Plasmid pBP136 from *Bordetella pertussis* represents an ancestral form of IncP-1 $\beta$  plasmids without accessory mobile elements. *Microbiology* **152**: 3477–3484.
- Kneller DG, Cohen FE, Langridge R. (1990). Improvements in protein secondary structure prediction by an enhanced neural network. *J Mol Biol* **214**: 171–182.
- Kongsuwan K, Josh P, Picault MJ, Wijffels G, Dalrymple B. (2006). The plasmid RK2 replication initiator protein (TrfA) binds to the sliding clamp  $\beta$  subunit of DNA polymerase III: implication for the toxicity of a peptide derived from the amino-terminal portion of 33-kilodalton TrfA. *J Bacteriol* **188**: 5501–5509.
- Konieczny I, Helinski DR. (1997). Helicase delivery and activation by DnaA and TrfA proteins during the initiation of replication of the broad host range plasmid RK2. *J Biol Chem* **272**: 33312–33318.
- Lenski RE, Simpson SC, Nguyen TT. (1994). Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. *J Bacteriol* **176**: 3140–3147.
- Levy SB, Marshall B. (2004). Antibacterial resistance worldwide: causes, challenges and responses. *Nat Med* **10**: S122–S129.
- Lin J, Helinski DR. (1992). Analysis of mutations in trfA, the replication initiation gene of the broad-host-range plasmid RK2. *J Bacteriol* **174**: 4110–4119.
- Maestro B, Sanz JM, Diaz-Orejas R, Fernandez-Tresguerres E. (2003). Modulation of pPS10 host range by plasmid-encoded RepA initiator protein. *J Bacteriol* **185**: 1367–1375.
- Maestro B, Sanz JM, Faalen M, Couturier M, Diaz-Orejas R, Fernandez-Tresguerres E. (2002). Modulation of pPS10 host range by DnaA. *Mol Microbiol* **46**: 223–234.
- Martinez B, Tomkins J, Wackett LP, Wing R, Sadowsky MJ. (2001). Complete nucleotide sequence and organization of the atrazine catabolic plasmid pADP-1 from *Pseudomonas* sp. strain ADP. *J Bacteriol* **183**: 5684–5697.
- Modi RI, Adams J. (1991). Coevolution in bacterial-plasmid populations. *Evolution* **45**: 656–667.
- Modi RI, Wilke CM, Rosenzweig RF, Adams J. (1991). Plasmid macro-evolution: selection of deletions during adaptation in a nutrient-limited environment. *Genetica* **84**: 195–202.
- Oldfield CJ, Cheng Y, Cortese MS, Romero P, Uversky VN, Dunker AK. (2005). Coupled folding and binding with alpha-helix-forming molecular recognition elements. *Biochemistry* **44**: 12454–12470.
- Pacek M, Konopa G, Konieczny I. (2001). DnaA box sequences as the site for helicase delivery during plasmid RK2 replication initiation in *Escherichia coli*. *J Biol Chem* **276**: 23639–23644.
- Pansegrau W, Lanka E, Barth PT, Figurski DH, Guiney DG, Haas D *et al.* (1994). Complete nucleotide sequence of Birmingham IncP $\alpha$  plasmids. Compilation and comparative analysis. *J Mol Biol* **239**: 623–663.
- Providenti MA, O'Brien JM, Ewing RJ, Paterson ES, Smith ML. (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.
- R-development-core-team (2009). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria, <http://www.R-project.org>.
- Rost B, Yachdav G, Liu J. (2004). The PredictProtein server. *Nucleic Acids Res* **32**: W321–W326.
- Schlüter A, Szczepanowski R, Pühler A, Top EM. (2007). Genomics of IncP-1 antibiotic resistance plasmids isolated from wastewater treatment plants provides evidence for a widely accessible drug resistance gene pool. *FEMS Microbiol Rev* **31**: 449–477.
- Sen D, Yano H, Suzuki H, Król JE, Rogers L, Brown CJ *et al.* (2010). Comparative genomics of pAKD4, the prototype IncP-1 $\delta$  plasmid with a complete backbone. *Plasmid* **63**: 98–107.
- Shingler V, Thomas CM. (1984). Analysis of the trfA region of broad-host-range plasmid RK2 by transposon mutagenesis and identification of polypeptide products. *J Mol Biol* **175**: 229–249.
- Slater FR, Bailey MJ, Tett AJ, Turner SL. (2008). Progress towards understanding the fate of plasmids in bacterial communities. *FEMS Microbiol Ecol* **66**: 3–13.
- Stevenson BS, Schmidt TM. (2004). Life history implications of rRNA gene copy number in *Escherichia coli*. *Appl Environ Microbiol* **70**: 6670–6677.
- Strahilevitz J, Jacoby GA, Hooper DC, Robicsek A. (2009). Plasmid-mediated quinolone resistance: a multifaceted threat. *Clin Microbiol Rev* **22**: 664–689.
- Taylor LA, Rose RE. (1988). A correction in the nucleotide sequence of the Tn903 kanamycin resistance determinant in pUC4K. *Nucleic Acids Res* **16**: 358.
- Thomas CM, Helinski DR. (1989). Vegetative replication and stable inheritance of IncP plasmids. In: CM Thomas (ed.). *Promiscuous Plasmids of Gram-negative Bacteria*. Academic Press: London.
- Thomas CM, Nielsen KM. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat Rev Microbiol* **3**: 711–721.
- Thorsted PB, Macartney DP, Akhtar P, Haines AS, Ali N, Davidson P *et al.* (1998). Complete sequence of the IncP $\beta$  plasmid R751: implications for evolution and organisation of the IncP backbone. *J Mol Biol* **282**: 969–990.
- Top EM, Holben WE, Forney LJ. (1995). Characterization of diverse 2,4-dichlorophenoxyacetic acid-degradative plasmids isolated from soil by complementation. *Appl Environ Microbiol* **61**: 1691–1698.
- Toukdarian A. (2004). Plasmid strategies for broad-host-range replication in gram-negative bacteria. In: BE Funnell & GJ Phillips (eds). *Plasmid Biology*. ASM press: Washington, D.C., USA, pp 259–270.
- Turner PE, Cooper VS, Lenski RE. (1998). Tradeoff between horizontal and vertical modes of transmission in bacterial plasmids. *Evolution* **52**: 315–329.
- Turner PE. (2004). Phenotypic plasticity in bacterial plasmids. *Genetics* **167**: 9–20.
- Vedler E, Vahter M, Heinaru A. (2004). The completely sequenced plasmid pEST4011 contains a novel IncP1 backbone and a catabolic transposon harboring tfd genes for 2,4-dichlorophenoxyacetic acid degradation. *J Bacteriol* **186**: 7161–7174.
- Wichman HA, Badgett MR, Scott LA, Boulianne CM, Bull JJ. (1999). Different trajectories of parallel evolution during viral adaptation. *Science* **285**: 422–424.
- Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, Okura VK *et al.* (2001). The genome of the natural

- genetic engineer *Agrobacterium tumefaciens* C58. *Science* **294**: 2317–2323.
- Wood TE, Burke JM, Rieseberg LH. (2005). Parallel genotypic adaptation: when evolution repeats itself. *Genetica* **123**: 157–170.
- Woods R, Schneider D, Winkworth CL, Riley MA, Lenski RE. (2006). Tests of parallel molecular evolution in a long-term experiment with *Escherichia coli*. *Proc Natl Acad Sci USA* **103**: 9107–9112.
- Zhong Z, Helinski D, Toukdarian A. (2003). A specific region in the N terminus of a replication initiation protein of plasmid RK2 is required for recruitment of *Pseudomonas aeruginosa* DnaB helicase to the plasmid origin. *J Biol Chem* **278**: 45305–45310.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)