Published in final edited form as:

Plasmid. 2009 March ; 61(2): 94–109. doi:10.1016/j.plasmid.2008.11.004.

The RepA_N replicons of Gram-positive bacteria: a family of broadly distributed but narrow host range plasmids

Keith E. Weaver^{1,*}, Stephen M. Kwong², Neville Firth², and Maria Victoria Francia³

¹ Division of Basic Biomedical Sciences, Sanford School of Medicine, University of South Dakota, Vermillion, SD 57069

² School of Biological Sciences, University of Sydney, New South Wales 2006, Australia

³ Servicio de Microbiología, Hospital Universitario Marques de Valdecilla-IFIMAV, Avda. de Valdecilla s/n, 39008 Santander, Cantabria, Spain

Abstract

The pheromone-responsive conjugative plasmids of *Enterococcus faecalis* and the multi-resistance plasmids pSK1 and pSK41 of *Staphylococcus aureus* are among the best studied plasmids native to Gram-positive bacteria. Although these plasmids seem largely restricted to their native hosts, protein sequence comparison of their replication initiator proteins indicates that they are clearly related. Homology searches indicate that these replicons are representatives of a large family of plasmids and a few phage that are widespread among the low G+C Gram-positive bacteria. We propose to name this family the RepA_N family of replicons after the annotated conserved domain that the initiator protein contains. Detailed sequence comparisons indicate that the initiator protein phylogeny is largely congruent with that of the host, suggesting that the replicons have evolved along with their current hosts and that intergeneric transfer has been rare. However, related proteins were identified on chromosomal regions bearing characteristics indicative of ICE elements, and the phylogeny of these proteins displayed evidence of more frequent intergeneric transfer. Comparison of stability determinants associated with the RepA_N replicons suggests that they have a modular evolution as has been observed in other plasmid families.

Introduction

Detailed studies of replication and stable inheritance of theta-replicating plasmids remains a pursuit largely restricted to Gram-negative bacteria, in spite of the obvious importance of such plasmids to the dissemination of antibiotic resistance in clinically important Gram-positive bacteria (Paulsen et al. 2003, Firth and Skurray 2006, Weaver 2006). While some important papers have addressed replication and stability of a few plasmids native to Gram-positive bacteria, particularly the closely-related plasmids of the Inc18 family (Le Chatelier et al. 2001, Dmowski et al. 2006, Liroy et al. 2006, Heidrich and Brantl 2007), only studies of the rolling circle replicating plasmids rival those of Gram-negative bacterial plasmids in breadth of plasmid coverage and level of detail (Khan 2005). Clearly, much remains to be discovered concerning the mobile genomes of Gram-positive bacteria.

Corresponding author: Keith E. Weaver, Ph.D., Division of Basic Biomedical Sciences, Sanford School of Medicine, University of South Dakota, Vermillion, SD 57069, Phone: 605-677-5169, Fax: 605-677-6381, e-mail: kweaver@usd.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

One family of plasmids that has attracted considerable interest because of their unique conjugative mechanism is the pheromone-responsive plasmids of *Enterococcus faecalis* (Clewell 2007, Dunny 2007). These plasmids encode signal sensing systems that operate via the import of short, peptide pheromones secreted by potential recipients lacking the plasmid. Signal sensing and transduction leads to the production of plasmid-encoded proteins that contribute to intercellular aggregation and facilitate single-stranded DNA transfer to the recipient cell. Once a copy of the plasmid is acquired, other plasmid-encoded components contribute to the reduction of and/or interference with the pheromone signal.

While the first description of the pheromone response was published in 1978 (Dunny et al. 1978), it wasn't until 15 years later that the basic replicon of the prototype plasmid, pAD1, was identified and characterized (Weaver et al. 1993, Weaver and Tritle 1994). The described pAD1 replicon included three predicted proteins, RepA, RepB, and RepC and a locus, named *par*, which was later determined to function as a post-segregational killing system (Weaver et al. 1996, Weaver et al. 1998). Transposon insertions in the *repA* gene prevented replication in *E. faecalis* (Weaver et al. 1993), suggesting that the RepA protein functioned as the replication initiator protein for the replicon, a suggestion that was later confirmed (Francia et al. 2004). Transposon insertions in the *repB* and *repC* genes resulted in an increase in plasmid copy number and a loss of plasmid stability (Weaver et al. 1993), suggesting that their protein products functioned to control replication frequency and/or facilitate segregational stability of the plasmid. At the time, the only proteins showing significant homology to pAD1 RepB in the database were several proteins, designated RepA, present on plasmids found in *Agrobacterium tumefaciens*. Later analyses revealed that both pAD1 RepB and the *Agrobacterium* RepA proteins belonged to a family of ATPases that, along with a DNA binding protein and a centromere-like site, comprise an "active partition locus" which facilitates plasmid distribution at cell division (Gerdes et al. 2000). Such loci are now referred to as type I partition loci, and the pAD1 RepB/C system was assigned to the type Ib sub-group based on sequence homology and genetic organization. Recently it was shown that the pAD1 RepB and RepC proteins and adjacent RepC binding sites do indeed function as a partition locus for pAD1 (Francia et al. 2007).

The only protein in the GenBank database with significant homology to pAD1 RepA at the time the pAD1 replicon was described was the single open reading frame on the ~3.3 kb *Lactobacillus helveticus* plasmid pLJ1 (Takiguchi et al. 1989). This was a surprising observation since the pheromone-responsive plasmids were all relatively large plasmids and none were known to replicate outside of the enterococci. Nevertheless, RepA homologues were soon identified on other pheromone-responsive plasmids, including pCF10 (Hedberg et al. 1996), pPD1 (Fujimoto et al. 1995) and pAM373 (De Boever et al. 2000), unifying the replicons as well as the conjugative systems of these plasmids. Subsequent studies then revealed that RepA-type replicons are broadly distributed throughout the low G+C Gram-positive bacteria (Gering et al. 1996, Firth et al. 2000). The replication initiator proteins of the *Staphylococcus aureus* plasmids pSK1, pSK41, and pI9789, and the *S. xylois* plasmid pSX267, were found to be homologous to the *E. faecalis* plasmid RepA proteins. Furthermore, this homology extended to *Lactobacillus* plasmid pLH1 (Thompson et al. 1999), as well as the aforementioned pLJ1, and the *Bacillus subtilis* plasmid pLS32 (Tanaka and Ogura 1998). Despite the apparently broad distribution of the RepA-type replicons, none of these plasmids appear to have a broad host range, essentially being restricted to replication in their native hosts.

The staphylococcal plasmids that utilize a RepA-like initiator are particularly important from a medical perspective. Clinical *S. aureus* strains often harbor plasmids that encode resistance to multiple antimicrobial agents, and nearly all such multiresistance plasmids utilize a RepA-like replication protein; plasmid sequences containing *repA* homologues were present in six of the seven independent methicillin resistant *S. aureus* (MRSA) isolates for which genome

sequences are currently available (Kuroda et al. 2001, Baba et al. 2002, Holden et al. 2004, Gill et al. 2005, Diep et al. 2006, Mwangi et al. 2007).

There are now close to 120 proteins homologous to pAD1 RepA in the sequence databases (Table 1), about 70 of which are associated with plasmids or phage found in low G+C Gram positive bacteria; it should be noted that there are many more known plasmids that are presumed to encode a homologous protein based on structural similarity to one or more of the plasmids listed, but for which sequence data are lacking. Similarity is greatest in the N-terminal 100 amino acids where a conserved domain has been annotated by NCBI designated RepA_N and classified as pfam06970. Similarly, a search for pAD1 RepB homologs identified numerous related proteins, again mostly in low G+C Gram-positive bacteria. Interestingly, the phylogeny of the RepB proteins does not match that of their linked RepA_N proteins and, indeed, RepB homologs are frequently associated with unrelated replication initiator proteins, indicating a modular organization of the replication and partition systems. We henceforth designate plasmids that utilize a RepA_N-type initiation protein as RepA_N plasmids.

In this article, we will examine the phylogenies of the RepA_N and RepB families of proteins along with recent functional data on the replication and stability systems to provide a more complete picture of the current state of knowledge about this important group of plasmids. Based on our examination of the available evidence we will argue that i) the RepA_N replicons represent an ancient group of plasmids that have co-evolved with the low G+C Gram-positive bacteria, ii) the replication, partition and conjugative transfer components of the RepA_N plasmids have not evolved as a single unit but rather as separate modules that have been shuffled among various plasmids native to these organisms, and iii) the homologs of the pAD1 RepB protein comprise a third subgroup of partition modules, designated type Ic, that have centromere-like binding sites both upstream and downstream of the partition genes.

Phylogeny of RepA proteins

A routine BLAST search using the pAD1 RepA protein sequence as the query identified numerous annotated or putative proteins from a variety of mostly low G+C Gram-positive organisms. The majority of these proteins are associated with plasmids that have either been identified experimentally or were identified as genetic elements separate from the chromosome by sequence characteristics. Some have been experimentally determined to be essential for replication of their cognate plasmids (Gering et al. 1996, Tanaka and Ogura 1998, Francia et al. 2004, Kwong et al. 2004, Kwong et al. 2008). A few appear to be associated with phage or prophage elements. Others are simply listed as putative replication proteins without apparent association with extrachromosomal elements. While we have included some of these last proteins for comparison, the analyses discussed here focus on those proteins most clearly associated with extrachromosomal elements. A divergent set of chromosomally-encoded homologs will be described in a separate section below.

Fig. 1 shows an un-rooted phylogenetic tree of 43 RepA homologues that are most clearly associated with mobile genetic elements (MGE). Also included are two putative proteins not obviously associated with MGE from *Streptococcus agalactiae* and *Leuconostoc mesenteroides*. Chromosomal sequences were also identified in some of the other species, but those sequences clustered with known plasmid sequences and so were excluded. The RepA protein from the pTEF1 plasmid from *E. faecalis* V583 was also excluded because it was identical to pAD1 RepA. Similarly, the Rep proteins of the staphylococcal plasmids pLW1043 and pUSA03 are identical to that of pSK41.

The RepA homologs fell into five large clusters with a few outlying proteins that appeared not to be associated with any cluster. Generally, the RepA clusters correlated with the host genus. The staphylococcal plasmids were organized into two clusters that will be referred to as the

pSK1 and pSK41 clusters as these plasmids were previously determined to be representative of two separate groups based on non-sequence-related properties (Firth et al. 2000). While the pSK41 cluster contained only conjugative *S. aureus* plasmids, the pSK1 cluster included non-conjugative plasmids from several other staphylococcal species in addition to *S. aureus*. Another cluster included enterococcal plasmids, both pheromone-responsive conjugative plasmids from *E. faecalis* and non-pheromone-responsive plasmids from *Enterococcus faecium*. A third cluster included plasmids from two different genera, *Lactobacillus* and *Lactococcus*. However, the *Lactococcus* plasmids formed a sub-cluster separate from the *Lactobacillus* plasmids. The two *Lactobacillus helveticus* plasmids also appeared to be phylogenetically distinct from the rest of the *Lactobacillus* representatives. The final cluster included five sequences from three *Streptococcus* species. Interestingly, all of these sequences were either present in known phage genomes, *Streptococcus pneumoniae* phage EJ-1 (Romero et al. 2004) and *Streptococcus thermophilus* phage 7201 (Stanley et al. 2000), or were associated with homologs to phage genes on their host chromosomes suggesting that they are prophage. The single *Bacillus* representative, *B. subtilis* plasmid pLS32, was phylogenetically distinct from all of the other clusters. A BLAST search with the pLS32 sequence identified no further *Bacillus* plasmids, so pLS32 appears to be the only currently identified RepA-like replicon in *Bacillus*. The two genomic representatives from *S. agalactiae* and *L. mesenteroides* also appeared to be distinct from the other clusters. Unlike the *S. pyogenes* chromosomal loci, these homologs did not appear to be associated with phage genes.

There is considerable evidence for natural genetic exchange between the genera in which RepA_N plasmids are found. The inc18 plasmids, typified by pAM β 1 originally identified in *E. faecalis* (Dunny and Clewell 1975), are considered to be broad host range plasmids and can be introduced by conjugation into Gram-positive and Gram-negative hosts (Kurenbach et al. 2003). Related plasmids have been implicated in the transmission of the VanA-encoding vancomycin resistance transposon Tn1546 from *E. faecalis* to *S. aureus* (Flannagan et al. 2003, Weigel et al. 2003, Zhu et al. 2008). Similarly, remnants of staphylococcal plasmids have been found in the chromosome of an *E. faecalis* strain (Bonafede et al. 1997). In this context, and given the apparently broad distribution of RepA_N plasmids, the observed generic phylogenetic grouping of the RepA_N replicons (Fig. 1; (Firth et al. 2000)) was somewhat surprising, and may reflect an inability to replicate successfully outside of their cognate hosts. In the cases where this has been investigated, the plasmids have indeed displayed a relatively narrow host range (B. Buttaro, personal communication, N. Firth, unpublished data). The extent of sequence divergence observed between the RepA_N replicons is consistent with a significant evolutionary timescale. Taken together, these observations suggest that ancestral RepA_N plasmids either existed in the last common ancestor of the low G+C bacteria, or that they were introduced into these organisms not long after the genera began to diverge, and that co-evolution with their hosts may have resulted in host range restriction.

A multi-sequence alignment using representatives from each cluster, revealed that most sequence similarity was restricted to the first ~100 amino acids of each protein (Fig. 2). A more extensive comparison is shown in Fig. S1. This region encompasses the RepA_N conserved domain, the sequence of which is also shown in the alignment, although the conserved sequence in the proteins themselves appears to extend 10–20 amino acids farther. There appeared to be little overall conservation of sequence beyond the RepA_N domain. However, when plasmid sequences from the same genus were compared, the C-terminal ~120 amino acids showed sequence conservation at least as strong as that observed in the RepA_N domain. In the enterococcal (Fig. S2) and staphylococcal (Fig. S3) lineages, sequence conservation appeared to transcend species boundaries. For example, while there are clearly some species-specific substitutions in the *E. faecium* plasmids, e.g., Q at alignment position 314, E at 319, and K at 340, the majority of amino acids are conserved with *E. faecalis* plasmids. Indeed, the two *E. faecalis* plasmids pEJ97-1 and pAM373 are the most divergent in the C-terminal region,

consistent with the phylogenetic tree. Similarly, C-terminal sequence conservation is observed in the *Lactobacillus/Lactococcus* cluster (Fig. S4) and amino acid substitutions do not consistently correlate with host genus. For example, while genus-specific substitutions are apparent in the three lactococcal plasmids, e.g., T at alignment position 358 and Y at 360, such substitutions are also observed in the two *L. helveticus* plasmids, e.g., M at 349 and P at 360. In the case of the streptococcal phage EJ-1 and 7201, the C-terminus contains a conserved DnaD domain (pfam04271). DnaD is a primosome component, suggesting that like their plasmid counterparts, the phage RepA_N proteins might be involved in replication.

RepA_N origins of replication

In each alignment, there was a region of 30–50 amino acids between the conserved N- and C-terminal domains that was poorly conserved. With few exceptions, the encoding DNA sequence within this region contains a variable number of repeated sequences, usually in direct orientation, in both plasmid and phage elements. This repeat region is marked on Fig. S2 for all of the enterococcal plasmids and on Fig. S3 for select staphylococcal plasmids. This region has been implicated in origin function in several cases and has been clearly shown to function as an origin of replication in the case of pAD1 (Francia et al. 2004), pSK41 (Kwong et al. 2004), pLS32 (Tanaka and Ogura 1998) and pSK1 (Kwong et al. 2008). Several representative repeat regions are shown in Fig. 3. The repeat regions vary in size from ~90–150 nucleotides and are consistently smaller in the staphylococcal plasmids and streptococcal phage than in the *Lactobacillus/Lactococcus* and *Enterococcus* plasmids. All contain direct sequence repeats, frequently with partial repeats interspersed with full length repeats. In some plasmids the repeats are presented in tandem as in pSK41 and pCF10, while in others there are gaps. In the pSK1 plasmid family, one repeat is present in the inverse orientation. In some cases the repeats are organized such that they are translated into amino acid repeats in the resulting protein, in others they are not. For example, the pAM373 *repA* gene contains two complete and one partial 54 nucleotide repeat arranged in tandem, resulting in two 18 amino acid repeats and a 13 amino acid repeat segment. In contrast, pSK41 encodes three 23 nucleotide repeats and a fourth fragment which, of course, are not translated into amino acid repeats. Frequently, repetitive sequences are also found downstream of this region in more conserved sequence, but generally the repeats in the non-conserved region are more prominent.

In vitro studies with pAD1 RepA (Francia et al. 2004) and pCF10 PrgW (B. Buttaro, personal communication) have shown that the proteins bind specifically to small imperfect inverted repeats located within the direct repeat sequences. These inverted repeats converge on a poly-A tract as indicated in Fig. 3. Similar inverted repeats are present in pSK41 and examination of DNA binding of pSK41 RepA to the repeats reveals that a single A within the poly-A tract is left unprotected (Kwong et al. 2004). Other RepA_N-family members contain similar repeat structures as exemplified by pRUM and phage EJ-1 shown in Fig. 3. In other cases, such as pSAK1, poly-A tracts are apparent near the center of the direct repeats, but no obvious converging inverted repeats are present. In still other cases, like pSK1, terminal inverted repeats are apparent within the direct repeats but they do not converge on a poly-A tract. The relative roles of the inverted and direct repeats in replication have yet to be determined. pAD1 RepA (Francia et al. 2004) and pCF10 PrgW (B. Buttaro, personal communication) also have single-strand DNA binding activity. This activity may facilitate strand opening as has been observed for the initiator protein of the inc18 plasmid pAM β 1 (Le Chatelier et al. 2001). Like pAM β 1 RepE, pAD1 RepA appears to bind ssDNA non-specifically while PrgW appears to bind specifically to the inverted repeats as either ss- or dsDNA. Interestingly, recombination can occur between DR1a and DR1b of pAD1 *repA* resulting in an in-frame deletion of 35 codons (Francia et al. 2004). This deletion eliminates *oriV* function, but the altered RepA is still able to support replication of a vector containing the wild-type *oriV* *in trans*, suggesting that, at

most, only one copy of the amino acids encoded within the repeats is required for RepA function.

The *oriV* regions of pAD1 (Francia et al. 2004), pLS32 (Tanaka and Ogura 1998) and pSK1 (Kwong et al. 2008) have been shown to function as incompatibility determinants, suggesting that sequence variability within the repeats may be required to allow co-existence of related plasmids in the same host. The *E. faecalis* pheromone-responsive plasmids pCF10 and pPD1 are interesting in this context. Preliminary evidence indicates that the two plasmids are compatible (G. Dunny, personal communication), but the putative *oriV* repeats differ in only a single nucleotide in the first direct repeat. The RepA protein sequences of the two proteins differ at only 15 of 333 residues, with only three changes in the N-terminal domain which has been implicated in origin binding in pAD1 (Francia et al. 2004). Whether these amino acid changes correlate with a change in origin specificity and thus impart compatibility to the two plasmids has yet to be investigated. On the other hand, the staphylococcal plasmids pSK1 and p19879 have been assigned to the same incompatibility group, even though their replication initiation proteins and origin repeats display considerable sequence divergence, perhaps indicating a lack of specificity in Rep-origin interactions. However, it should be remembered that plasmid incompatibility can also be mediated by segregational stability determinants and negative regulators of copy number. The latter may be particularly relevant in this case since the antisense RNA that regulates pSK1 copy number has been found to exert potent incompatibility against its own replicon (Kwong et al. 2008).

Function and regulation of RepA_N proteins

The pattern of sequence conservation suggests a potential domain organization for the RepA_N proteins. The N-terminal domain defines the family and is conserved throughout all of the representatives, regardless of host. This suggests that the N-terminal domain performs some common essential function throughout all of the RepA_N proteins, e.g., strand separation. Interestingly, this domain is duplicated in the staphylococcal plasmids pPI-1 and pSSP1 (Aso et al. 2005), and the first copy is more similar to other staphylococcal RepA_N proteins whereas the second more closely resembles the RepA_N consensus. The central region of the RepA_N proteins is highly variable and is derived from repetitive DNA sequences implicated in RepA_N protein binding to the origin of replication, as described above. The sequence variability may provide compatibility for plasmids maintained in the same host (see above). It is possible that the coincident variability of the amino acid sequence provides the variable DNA binding specificity for origin binding. However, the fact that the DNA repeats are translated into amino acid repeats in some plasmids but not others suggests that this may not be the case. Furthermore, the N-terminal domain of pAD1 RepA has been shown to be sufficient for *oriV* binding while the C-terminal domain with two of the amino acid repeats corresponding to *oriV* did not bind the origin (Francia et al. 2004). It is interesting to note in this context that in the *S. aureus* plasmids regions of low sequence conservation are observed both upstream and downstream of the repeat region that could also provide the specificity for origin binding. The C-terminal domain is conserved between RepA_N plasmids found in the genus-specific clusters, but not in plasmids replicated in other genera. This suggests that the C-terminal domain performs some host-specific function. In the case of plasmids in Gram-negative bacteria, such host-specificity has been associated with interaction with host replication proteins, especially loading of the host replicative helicase, DnaB (Konieczny 2003). The C-terminal domain of the RepA_N plasmids could perform a similar function.

In most bacterial plasmids copy number is controlled by regulating the level or activity or the initiator protein. Little is known about copy number control in the RepA_N family plasmids except for the staphylococcal plasmids where an antisense RNA has been shown to be involved (Kwong et al. 2004, Kwong et al. 2006, Kwong et al. 2008). In the case of the *S. aureus* plasmid

pSK41 the antisense molecule, RNAI, interacts with the leader region of the *rep* mRNA, primarily resulting in repression of initiator translation (Kwong et al. 2006). How widespread this form of regulation is has yet to be determined, but an examination of the pAD1 *repA* gene shows no evidence of an antisense RNA suggesting that replication is regulated differently in this plasmid. In pLS32 of *B. subtilis*, the repeat region in the origin appears to be involved in copy number control, since mutations in either of the first three repeats increased copy number, whereas mutations in the fourth or fifth repeat abolished plasmid replication (Tanaka and Ogura 1998). In pCF10, mutation of two cysteine residues that are conserved in the pheromone-responsive plasmids results in a decrease in plasmid copy number. Binding of a precursor of the cCF10 pheromone increases the stability of a disulfide bond between the two cysteines (B. Buttaro, personal communication). Whether this is a regulatory mechanism or whether the disulfide bond is important for initiator function is not yet clear. Interestingly, expression of pre-cCF10 in a heterologous host allows replication in that host, suggesting that it plays a central role in initiator protein function.

Segregational stability modules of RepA_N plasmids

While the RepA_N replication initiator proteins constitute a phylogenetically coherent group, their associated segregational stability genes do not, providing evidence for a modular evolution of these plasmids as has been observed for other systems (Thomas 2000). For example all RepA_N replicons from the enterococci, with the exception of the *E. faecium* DO plasmid, are associated with a protein belonging to the ParA-family of ATPases (Gerdes et al. 2000). Proteins in this family function in cooperation with a poorly conserved DNA binding protein in active partition of plasmid copies at cell division. In the case of pAD1, the ParA homolog RepB and its downstream partner RepC have been shown to be required for plasmid stability (Francia et al. 2007). However, as shown in the phylogenetic tree in Fig. 4, when pAD1 RepB homologs retrieved from a BLAST search are compared, those associated with RepA_N replicons fail to cluster together. Furthermore, the RepB homologs fail to segregate by host genus as did the RepA_N proteins. Thus, interspersed between the Enterococcal RepB homologs are proteins from plasmids from *Lactobacillus*, *Leuconostoc*, *Lactococcus*, *Listeria*, *Bacillus*, *Clostridium*, *Pediococcus*, *Streptococcus*, and *Staphylococcus*. Even more indicative of a modular evolution of these plasmids is the observation that many of the RepB homologs are not associated with RepA_N-type replicons at all. Many of the RepB-associated replicons encode inc18-type replication initiator proteins which are evolutionarily unrelated to the RepA_N initiator proteins. The pRUM and pTEF2 RepB proteins cluster with the Soj family of chromosomal ParA-family proteins believed to be involved in chromosomal segregation (Lee and Grossman 2006).

In a previous review, ParA/ParB-family partition systems were divided into two groups based on the organization of protein binding sites around the partition module and on whether the ParA component had DNA binding properties (Gerdes et al. 2000). In type Ia systems the ParA protein binds upstream of the *parAB* operon and autorepresses transcription while the ParB protein binds downstream of the operon at the centromere-like binding site. In type Ib systems the ParA protein lacks DNA binding capacity and the centromere-like site is located upstream of the operon, overlapping the promoter. ParB binding simultaneously stimulates partition and autorepresses transcription. Many of the ParA-like RepB homologs shown in Fig. 4 are associated with small, poorly conserved, downstream open reading frames reminiscent of ParB proteins. These ParA/ParB-like operons are often flanked by repeats that may act as centromere-like sites. In the case of pAD1 the putative ParB-like protein, RepC, has been demonstrated to bind to the associated repeats and, with the RepB protein, stabilize a heterologous plasmid (Francia et al. 2007). The organization of the repeats is, however, different than in either type Ia or Ib systems. As shown in Fig. 5, in most cases, related repeats are located both upstream and downstream of the RepBC operon, presumably binding the ParB-

like component in both locations. We therefore suggest that such systems be designated as type Ic, characterized by a type Ib-like ParA homologue but with DNA binding sites for the ParB-like protein both upstream and downstream of the operon. The pattern and orientation of repeats differs significantly in different plasmids. In some cases the majority of repeats are upstream of the operon, in others the number of repeats is relatively evenly split, while in others the majority of repeats are downstream of the operon. The relative importance of the upstream and downstream repeats is as yet unclear. While operons containing only the upstream repeats of the pAD1 (Francia et al. 2007) and pGENT (Derome et al. 2008) systems have been shown to be competent in stabilizing heterologous plasmids, these test plasmids were relatively high copy vectors that may not be representative of the low copy native plasmids. The low copy native plasmids may require the downstream repeats to fine tune partition. Interestingly in this context, the *par2* locus of plasmid pB171 of *Escherichia coli* also contains upstream and downstream ParB-binding repeats (Fig. 5). In this case binding to both of the repeat regions, *repC1* and *repC2*, is important for partition and IHF binding to the *par2* DNA has been proposed to bring the two repeat sequences together in a partition complex (Ringgaard et al. 2007). The ParA protein of the pB171 *par2* locus is not closely related to the ParA proteins shown in Fig. 4. Furthermore, partition in this system is further complicated by the involvement of a type II partition system, *par1*, that is coordinately regulated with *par2*. Not all of the RepB homologs shown in Fig. 4 have both upstream and downstream repeats. The *L. reuteri* plasmid pTE15 contains a complex set of two different repeats located only upstream of the operon (Fig. 5). It seems apparent that a variety of DNA binding site organizations can constitute functional centromere-like sites for plasmid partition in type I systems.

As further evidence of modular evolution, plasmids encoding RepA_N-family initiator proteins outside of the enterococci encode stability determinants unrelated to the RepB-family of proteins. *B. subtilis* pLS32 and staphylococcal conjugative plasmids related to pSK41 possess type II partitioning systems that are distantly related to that of the archetypal system of the *E. coli* IncFII plasmid R1 (Møller-Jensen et al. 2002, Becker et al. 2006, Schumacher et al. 2007). These systems comprise a centromere-like site, *parC*, upstream of an operon encoding an actin-like ParM protein and a DNA binding protein, ParR, that initiates partitioning by binding to the centromere and additionally autoregulates *par* operon transcription. However, most non-conjugative RepA_N plasmids from *S. aureus* and coagulase-negative species lack one of the well established types of partitioning systems. Instead, they possess a homolog of a single gene that for pSK1 has been shown to enhance segregational stability (Simpson et al. 2003). In pSK1 and most staphylococcal plasmids, this gene is located immediately upstream of, and transcribed divergently from, the replication initiation gene. However, it is not required for replication and has been shown to stabilize heterologous replicons without influencing multimer resolution or the viability of plasmid-free segregants. Since a DNA fragment containing sequence repeats located upstream of the gene exerts incompatibility, as would be expected for a centromere-like site, it has been tentatively designated *par*. pSK1 Par-like proteins are predicted to possess a helix-turn-helix DNA binding domain and a coiled-coil region, which together might allow them to perform both centromere binding and cytoskeletal activities that are performed by separate proteins in other partitioning systems (types I-III); as such, they potentially define a new class of partitioning system, type IV (Schumacher 2008). As with pAD1-like type 1c systems, pSK1 *par*-like determinants are also not restricted to RepA_N-like replicons; homologs are found on plasmids native to a variety of low G+C Gram-positive bacteria that utilize at least six distinct families of replication genes (Firth et al. 2000). Finally, pPI-1 and pSSP1 from *S. warneri* and *S. saprophyticus*, respectively, contain type Ib partitioning systems. Thus, one of three distinct partitioning determinants are variously utilized by the staphylococcal RepA_N plasmids. Additionally, genes encoding resolvases of the serine recombinase family appear to be ubiquitous on staphylococcal RepA_N plasmids; the *res* gene of pSK41 has been confirmed to contribute to segregational stability by resolving plasmid multimers (LeBard et al. 2008). The sequences of these resolvases vary considerably

between plasmids, suggesting a complex phylogeny akin to that described above for pAD1 RepB. The stability determinants of other RepA_N replicons have yet to be determined, but it is worth noting that in pLJ1 from *L. helveticus*, the RepA_N protein is encoded by the only apparent open reading frame (Takiguchi et al. 1989), indicating that this plasmid either relies on a high copy number or some other unknown mechanism to ensure stability.

repA_N genes: not just for plasmids?

A BLAST search using only the RepA_N consensus sequence identified numerous chromosomally-encoded RepA_N family members, most of which have been recently identified in a large scale genome sequencing project characterizing the “Human Gut Microbiome” performed by the Genome Sequencing Center at the Washington University School of Medicine (http://genome.wustl.edu/sub_genome_group.cgi?GROUP=3&SUB_GROUP=4). A phylogenetic tree of representatives of this group of RepA_N proteins along with several other homologs that were identified in the NCBI data-base is shown in Fig. 6. These proteins fell into several clusters distinct from the plasmid-encoded RepA_N proteins and did not appear to segregate by genus. Indeed, although the majority of encoding organisms belong to the low G+C Gram-positive group, four RepA_N-encoding genes were identified in the Gram-negative genus *Bacteroides*, three from the same strain of *B. capillosus* ATCC29799. This pattern of multiple RepA_N homologs on a single genome was common among the human gut bacterial species with the highest number, six, being found in *Ruminococcus obeum* strain ATCC29174. In general, RepA_N proteins from a single strain did not cluster together but were scattered throughout the tree. Note also that not all of the RepA_N proteins in this group are present in gut microbes. Thus, the C cluster contains a protein from *S. epidermidis* and the B cluster contains proteins from *S. pyogenes* and the mollicute *Ureaplasma urealyticum*.

Interestingly, while the chromosomal RepA_N proteins do not segregate by genera, they do segregate by size. While the plasmid-encoded initiator proteins ranged from ~300–360 amino acids and the Streptococcal phage proteins were slightly smaller at ~250 amino acids, the chromosomal proteins ranged in size from 113 to 478 amino acids. Phylogenetically, the proteins segregate into three clusters; one cluster consisted of relatively small proteins of 113 to 181 residues, a second consisted of proteins similar in size to the streptococcal phage proteins, and a third consisted of proteins similar in size to the plasmid proteins and larger. The significance of this pattern is unclear, but suggests that different C-terminal domains may be associated with the RepA_N domain in these clusters.

The distinct phylogenetic distribution of the gut microbial RepA_N proteins suggests that they have a different evolutionary history than their plasmid-encoded counterparts. Interestingly, two of the proteins on the tree are associated with known conjugative transposons, CTnBST from *Bacteroides uniformis* (Gupta et al. 2003) and the VanB-encoding Tn1549 from *E. faecalis* (Garnier et al. 2000). Indeed, many of the RepA_N encoding genes present in the gut bacterial genomes are linked with genes encoding tyrosine or serine recombinases, commonly used as integrases by conjugative transposons (Burrus et al. 2002). An association with conjugative transposons could explain the lack of segregation of these proteins with their host genera, since these elements are generally considered to have a very broad host range and may be frequently transferred among the human gut microbiome. What function these RepA_N proteins are performing for their host element is unclear. Since conjugative transposons are not believed to replicate independently from the host chromosome, it seems unlikely that they perform the same function as the plasmid-encoded replication initiator proteins. Conjugative transposons are frequently mosaic elements composed of plasmid, phage, and transposon sequences suggesting frequent recombination events with other genetic elements. It is possible that these RepA_N proteins are non-functional remnants of a recombination event with a

plasmid, although their prevalence and apparent integrity argues against this. Alternatively, the protein may have been adapted to perform a different function for the conjugative transposon, e.g., replacement strand synthesis in either the donor or the recipient.

Conclusion

Two classes of plasmids have been relatively well-described in Gram-positive bacteria, the rolling circle replicating plasmids and the Inc18 plasmids, both of which have a broad host range within this group of bacteria. We propose the establishment of a third category of plasmids encompassing replicons encoding putative replication initiator proteins homologous to those present on the *E. faecalis* pheromone-responsive plasmids and the staphylococcal multiresistance plasmids, designated the RepA_N replicons. Unlike most plasmid groups described in Gram-negative bacteria, the RepA_N plasmids do not make up a family of mutually incompatible elements. In fact, many members of the group able to replicate in the same species are compatible with one another. Nevertheless, the sequence homology of the initiator proteins and the organization and location of repeat sequences that likely function as replication origins indicates their evolutionary relatedness. Another curious aspect of the biology of these plasmids is that, in spite of the broad distribution of family members in the low G+C Gram-positive bacteria, the well-studied members of the group appear to have a relatively narrow host range. This observation suggests that these plasmids pre-existed the divergence of their host organisms and adapted over time to their chosen hosts. Correlation with the segregational stability determinants carried by RepA_N plasmids suggests that, like other plasmids, the RepA_N-encoding plasmids have evolved by accretion of functional modules rather than as a single unit. The conjugation systems encoded by some of these plasmids further illustrates this point. The *tra* system encoded by staphylococcal pSK41-like plasmids is unrelated to the conjugation systems of the enterococcal pheromone-responsive plasmids, instead belonging to the type IV secretion system superfamily (Grohmann et al. 2003), which includes well-studied DNA transfer systems from Gram-negative organisms, such as the *Escherichia coli* F plasmid conjugation system and the *Agrobacterium tumefaciens* Ti and Tra DNA transfer systems. However, the pSK41 transfer system is most closely related to those of other Gram-positive hosts, such as from the non-RepA_N plasmid pMRC01 from *L. lactis* and the inc-18 broad host range plasmid pIP501, originally identified in *S. agalactiae* (Horodniceanu et al. 1976). Antibiotic resistance genes are another obvious example of modules that are variously represented in RepA_N plasmids and replicons of other types.

Much remains to be discovered concerning the replication of RepA_N replicons. The N-terminal region of the replication initiator proteins is most widely conserved while the C-terminal region is conserved predominantly within members native to a particular host. This pattern suggests that the N-terminal domain performs a function common to all of the replicons while the C-terminal domain performs a genus-specific function. What these functions are is yet to be determined. How the protein interacts with the origin sequence to accomplish strand separation and recruitment of the replisome is also a mystery. With the exception of the staphylococcal multiresistance plasmids, mechanisms of copy control are also poorly understood. Given the importance of these plasmids in the dissemination of antibiotic resistance and virulence factors, it is important to continue study of their mechanisms of replication and stable inheritance. Finally, the distribution of RepA_N module-containing proteins in human gut bacteria suggests a different pattern of inheritance from their plasmid-encoded counterparts. It seems likely that these proteins are associated with conjugative transposons. It will be interesting to determine what role these proteins play in the biology of their host elements and how they relate to the plasmid proteins.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

Research in the laboratory of KW is supported by Public Health Service grant GM55544. Research in the laboratory of SMK and NF is supported by the National Health and Medical Research Council of Australia. Research in the laboratory of MVF is supported by grants FIS 02/3029 and FIS PI04/0802 from the Spanish Fondo de Investigacion Sanitaria, Instituto de Salud Carlos III, grant LSHE-CT-2007-037410 from the European Union and grant REIPI RD06/0008 from the Ministerio de Sanidad y Consumo, Instituto de Salud Carlos III-FEDER, Spanish Network for Research in Infectious Diseases.

Literature Cited

- Allignet J, El Solh N. Comparative analysis of staphylococcal plasmids carrying three streptogramin-resistance genes: *vat-vgb-vga*. *Plasmid* 1999;42:134–138. [PubMed: 10489330]
- Anthonisen IL, Sunde M, Steinum TM, Sidhu MS, Sorum H. Organization of the antiseptic resistance gene *qacA* and Tn552-Related β -lactamase genes in multidrug-resistant *Staphylococcus haemolyticus* strains of animal and human origins. *Antimicrob Agents Chemother* 2002;46:3606–3612. [PubMed: 12384372]
- Aso Y, Koga H, Sashihara T, Nagao J-i, Kanemasa Y, Nakayama J, Sonomoto K. Description of complete DNA sequence of two plasmids from the nukacin ISK-1 producer, *Staphylococcus warneri* ISK-1. *Plasmid* 2005;53:164–178. [PubMed: 15737403]
- Baba T, Takeuchi F, Kuroda M, Yuzawa H, Aoki K-i, Oguchi A, Nagai Y, Iwama N, Asano K, Naimi T, Kuroda H, Cui L, Yamamoto K, Hiramatsu K. Genome and virulence determinants of high virulence community-acquired MRSA. *The Lancet* 2002;359:1819–1827.
- Becker E, Herrera NC, Gunderson FQ, Derman AI, Dance AL, Sims J, Larsen RA, Pogliano J. DNA segregation by the bacterial actin AlfA during *Bacillus subtilis* growth and development. *EMBO J* 2006;25:5919–5931. [PubMed: 17139259]
- Beres SB, Richter EW, Nagiec MJ, Sumby P, Porcella SF, DeLeo FR, Musser JM. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen group A Streptococcus. *Proceedings of the National Academy of Sciences* 2006;103:7059–7064.
- Beres SB, Sylva GL, Barbian KD, Lei B, Hoff JS, Mammarella ND, Liu MY, Smoot JC, Porcella SF, Parkins LD, Campbell DS, Smith TM, McCormick JK, Leung DYM, Schlievert PM, Musser JM. Genome sequence of a serotype M3 strain of group A Streptococcus: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. *Proceedings of the National Academy of Sciences* 2002;99:10078–10083.
- Berg T, Firth N, Apisiridej S, Hettiaratchi A, Leelaporn A, Skurray RA. Complete nucleotide sequence of pSK41: evolution of staphylococcal conjugative multiresistance plasmids. *J Bacteriol* 1998;180:4350–4359. [PubMed: 9721269]
- Bonafede M, Carias L, Rice L. Enterococcal transposon Tn5384: evolution of a composite transposon through cointegration of enterococcal and staphylococcal plasmids. *Antimicrob Agents Chemother* 1997;41:1854–1858. [PubMed: 9303373]
- Burrus V, Pavlovic G, Decaris B, Guédon G. Conjugative transposons: the tip of the iceberg. *Mol Microbiol* 2002;46:601–610. [PubMed: 12410819]
- Chen C, Tang J, Dong W, Wang C, Feng Y, Wang J, Zheng F, Pan X, Liu D, Li M, Song Y, Zhu X, Sun H, Feng T, Guo Z, Ju A, Ge J, Dong Y, Sun W, Jiang Y, Wang J, Yan J, Yang H, Wang X, Gao GF, Yang R, Wang J, Yu J. A glimpse of streptococcal toxic shock syndrome from comparative genomics of *S. suis* 2 chinese isolates. *PLoS ONE* 2007;2:e315. [PubMed: 17375201]
- Clewell DB. Properties of *Enterococcus faecalis* plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. *Plasmid* 2007;58:205–227. [PubMed: 17590438]
- De Boever EH, Clewell DB, Fraser CM. *Enterococcus faecalis* conjugative plasmid pAM373: complete nucleotide sequence and genetic analyses of sex pheromone response. *Mol Microbiol* 2000;37:1327–1341. [PubMed: 10998166]

- Derome A, Hoischen C, Bussiek M, Grady R, Adamczyk M, Keedzierska B, Diekmann S, Barilla D, Hayes F. Centromere anatomy in the multidrug-resistant pathogen *Enterococcus faecium*. *Proceedings of the National Academy of Sciences* 2008;105:2151–2156.
- Desmond C, Ross RP, Fitzgerald G, Stanton C. Sequence analysis of the plasmid genome of the probiotic strain *Lactobacillus paracasei* NFBC338 which includes the plasmids pCD01 and pCD02. *Plasmid* 2005;54:160–175. [PubMed: 15964627]
- Diep BA, Gill SR, Chang RF, Phan TH, Chen JH, Davidson MG, Lin F, Lin J, Carleton HA, Mongodin EF, Sensabaugh GF, Perdreau-Remington F. Complete genome sequence of USA300, an epidemic clone of community-acquired methicillin-resistant *Staphylococcus aureus*. *The Lancet* 2006;367:731–739.
- Dmowski M, Sitkiewicz I, Ceglowski P. Characterization of a novel partition system encoded by the δ and ω genes from the streptococcal plasmid pSM19035. *J Bacteriol* 2006;188:4362–4372. [PubMed: 16740943]
- Dunny GM. The peptide pheromone-inducible conjugation system of *Enterococcus faecalis* plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. *Philos Trans R Soc B Biol Sci* 2007;362:1185–1193.
- Dunny GM, Brown BL, Clewell DB. Induced cell aggregation and mating in *Streptococcus faecalis*: evidence for a bacterial sex pheromone. *Proc Natl Acad Sci USA* 1978;75:3479–3483. [PubMed: 98769]
- Dunny GM, Clewell DB. Transmissible toxin (hemolysin) plasmid in *Streptococcus faecalis* and its mobilization of a noninfectious drug resistance plasmid. *J Bacteriol* 1975;124:784–790. [PubMed: 810482]
- Ebersbach G, Gerdes K. The double par locus of virulence factor pB171: DNA segregation is correlated with oscillation of ParA 10.1073/pnas.261569598. *Proceedings of the National Academy of Sciences of the United States of America* 2001;98:15078–15083. [PubMed: 11752455]
- Firth N, Apisiridej S, Berg T, O'Rourke BA, Curnock S, Dyke KGH, Skurray RA. Replication of staphylococcal multiresistance plasmids. *J Bacteriol* 2000;182:2170–2178. [PubMed: 10735859]
- Firth, N.; Skurray, RA. The *Staphylococcus*-Genetics: Accessory elements and genetic exchange. In: Fischetti, VA.; Novick, RP.; Ferretti, JJ.; Portnoy, DA.; Rood, JL., editors. *Gram-Positive Pathogens*. Washington, DC: ASM Press; 2006. p. 413–426.
- Flannagan SE, Chow JW, Donabedian SM, Brown WJ, Perri MB, Zervos MJ, Ozawa Y, Clewell DB. Plasmid content of a vancomycin-resistant *Enterococcus faecalis* isolate from a patient also colonized by *Staphylococcus aureus* with a VanA phenotype. *Antimicrob Agents Chemother* 2003;47:3954–3959. [PubMed: 14638508]
- Francia MV, Fujimoto S, Tille P, Weaver KE, Clewell DB. Replication of *Enterococcus faecalis* pheromone-responding plasmid pAD1: location of the minimal replicon and *oriV* site and RepA involvement in initiation of replication. *J Bacteriol* 2004;186:5003–5016. [PubMed: 15262938]
- Francia MV, Weaver KE, Goicoechea P, Tille P, Clewell DB. Characterization of an active partition system for the *Enterococcus faecalis* pheromone-responding plasmid pAD1. *J Bacteriol* 2007;189:8546–8555. [PubMed: 17905984]
- Fujimoto S, Tomita H, Wakamatsu E, Tanimoto K, Ike Y. Physical mapping of the conjugative bacteriocin plasmid pPD1 of *Enterococcus faecalis* and identification of the determinant related to the pheromone response. *J Bacteriol* 1995;177:5574–5581. [PubMed: 7559345]
- Garnier F, Taourit S, Glaser P, Courvalin P, Galimand M. Characterization of transposon Tn1549, conferring VanB-type resistance in *Enterococcus* spp. *Microbiology* 2000;146:1481–1489. [PubMed: 10846226]
- Gerdes K, Moller-Jensen J, Bugge Jensen R. Plasmid and chromosome partitioning: surprises from phylogeny. *Mol Microbiol* 2000;37:455–466. [PubMed: 10931339]
- Gering M, Gotz F, Bruckner R. Sequence and analysis of the replication region of the *Staphylococcus xylosus* plasmid pSX267. *Gene* 1996;182:117–122. [PubMed: 8982076]
- Gill SR, Fouts DE, Archer GL, Mongodin EF, DeBoy RT, Ravel J, Paulsen IT, Kolonay JF, Brinkac L, Beanan M, Dodson RJ, Daugherty SC, Madupu R, Angiuoli SV, Durkin AS, Haft DH, Vamathevan J, Khouri H, Utterback T, Lee C, Dimitrov G, Jiang L, Qin H, Weidman J, Tran K, Kang K, Hance IR, Nelson KE, Fraser CM. Insights on evolution of virulence and resistance from the complete

genome analysis of an early methicillin-resistant *Staphylococcus aureus* strain and a biofilm-producing methicillin-resistant *Staphylococcus epidermidis* strain. *J Bacteriol* 2005;187:2426–2438. [PubMed: 15774886]

- Glaser P, Rusniok C, Buchrieser C, Chevalier F, Frangeul L, Msadek T, Zouine M, Couvé E, Lalioui L, Poyart C, Trieu-Cuot P, Kunst F. Genome sequence of *Streptococcus agalactiae*, a pathogen causing invasive neonatal disease. *Mol Microbiol* 2002;45:1499–1513. [PubMed: 12354221]
- Goto T, Yamashita A, Hirakawa H, Matsutani M, Todo K, Ohshima K, Toh H, Miyamoto K, Kuhara S, Hattori M, Shimizu T, Akimoto S. Complete genome sequence of *Finegoldia magna*, an anaerobic opportunistic pathogen. *DNA Res* 2008;15:39–47. [PubMed: 18263572]
- Grady R, Hayes F. Axe-Txe, a broad-spectrum proteic toxin-antitoxin system specified by a multidrug-resistant, clinical isolate of *Enterococcus faecium*. *Mol Microbiol* 2003;47:1419–1432. [PubMed: 12603745]
- Grohmann E, Muth G, Espinosa M. Conjugative plasmid transfer in Gram-positive bacteria. *Microbiol Mol Biol Rev* 2003;67:277–301. [PubMed: 12794193]
- Gupta A, Vlamakis H, Shoemaker N, Salyers AA. A new *Bacteroides* conjugative transposon that carries an *ermB* gene. *Appl Environ Microbiol* 2003;69:6455–6463. [PubMed: 14602600]
- Hedberg PJ, Leonard BAB, Ruhfel RE, Dunny GM. Identification and characterization of the genes of *Enterococcus faecalis* plasmid pCF10 involved in replication and in negative control of pheromone-inducible conjugation. *Plasmid* 1996;35:46–57. [PubMed: 8693026]
- Heidrich N, Brantl S. Antisense RNA-mediated transcriptional attenuation in plasmid pIP501: the simultaneous interaction between two complementary loop pairs is required for efficient inhibition by the antisense RNA. *Microbiology* 2007;153:420–427. [PubMed: 17259613]
- Highlander S, Hulten K, Qin X, Jiang H, Yerrapragada S, Mason E, Shang Y, Williams T, Fortunov R, Liu Y, Igboeli O, Petrosino J, Tirumalai M, Uzman A, Fox G, Cardenas A, Muzny D, Hemphill L, Ding Y, Dugan S, Blyth P, Buhay C, Dinh H, Hawes A, Holder M, Kovar C, Lee S, Liu W, Nazareth L, Wang Q, Zhou J, Kaplan S, Weinstock G. Subtle genetic changes enhance virulence of methicillin resistant and sensitive *Staphylococcus aureus*. *BMC Microbiology* 2007;7:99. [PubMed: 17986343]
- Holden MTG, Feil EJ, Lindsay JA, Peacock SJ, Day NPJ, Enright MC, Foster TJ, Moore CE, Hurst L, Atkin R, Barron A, Bason N, Bentley SD, Chillingworth C, Chillingworth T, Churcher C, Clark L, Corton C, Cronin A, Doggett J, Dowd L, Feltwell T, Hance Z, Harris B, Hauser H, Holroyd S, Jagels K, James KD, Lennard N, Line A, Mayes R, Moule S, Mungall K, Ormond D, Quail MA, Rabinowitsch E, Rutherford K, Sanders M, Sharp S, Simmonds M, Stevens K, Whitehead S, Barrell BG, Spratt BG, Parkhill J. Complete genomes of two clinical *Staphylococcus aureus* strains: Evidence for the rapid evolution of virulence and drug resistance. *Proceedings of the National Academy of Sciences* 2004;101:9786–9791.
- Holden MTG, Scott A, Cherevach I, Chillingworth T, Churcher C, Cronin A, Dowd L, Feltwell T, Hamlin N, Holroyd S, Jagels K, Moule S, Mungall K, Quail MA, Price C, Rabinowitsch E, Sharp S, Skelton J, Whitehead S, Barrell BG, Kehoe M, Parkhill J. Complete genome of acute rheumatic fever-associated serotype M5 *Streptococcus pyogenes* strain Manfredo. *J Bacteriol* 2007;189:1473–1477. [PubMed: 17012393]
- Horodniceanu T, Bouanchaud DH, Bieth G, Chabbert YA. R plasmids in *Streptococcus agalactiae* (Group B). *Antimicrob Agents Chemother* 1976;10:795–801. [PubMed: 795369]
- Kearney K, Fitzgerald GF, Seegers JFML. Identification and characterization of an active plasmid partition mechanism for the novel *Lactococcus lactis* plasmid pCI2000. *J Bacteriol* 2000;182:30–37. [PubMed: 10613859]
- Khan SA. Plasmid rolling-circle replication: highlights of two decades of research. *Plasmid* 2005;53:126–136. [PubMed: 15737400]
- Konieczny I. Strategies for helicase recruitment and loading in bacteria. *EMBO Reports* 2003;4:37–41. [PubMed: 12524518]
- Kurenbach B, Bohn J, Prabhu J, Abudukerim M, Szewzyk U, Grohmann E. Intergeneric transfer of the *Enterococcus faecalis* plasmid pIP501 to *Escherichia coli* and *Streptomyces lividans* and sequence analysis of its *tra* region. *Plasmid* 2003;50:86–93. [PubMed: 12826062]
- Kuroda M, Ohta T, Uchiyama I, Baba T, Yuzawa H, Kobayashi I, Cui L, Oguchi A, Aoki K-i, Nagai Y, Lian J, Ito T, Kanamori M, Matsumaru H, Maruyama A, Murakami H, Hosoyama A, Mizutani-Ui

- Y, Takahashi NK, Sawano T, Inoue R-i, Kaito C, Sekimizu K, Hirakawa H, Kuhara S, Goto S, Yabuzaki J, Kanehisa M, Yamashita A, Oshima K, Furuya K, Yoshino C, Shiba T, Hattori M, Ogasawara N, Hayashi H, Hiramatsu K. Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*. *The Lancet* 2001;357:1225–1240.
- Kuroda M, Yamashita A, Hirakawa H, Kumano M, Morikawa K, Higashide M, Maruyama A, Inose Y, Matoba K, Toh H, Kuhara S, Hattori M, Ohta T. Whole genome sequence of *Staphylococcus saprophyticus* reveals the pathogenesis of uncomplicated urinary tract infection. *Proceedings of the National Academy of Sciences* 2005;102:13272–13277.
- Kwong SM, Lim R, LeBard RJ, Skurray RA, Firth N. Analysis of the pSK1 replicon, a prototype from the staphylococcal multiresistance plasmid family. *Microbiology* 2008;154:3084–3094. [PubMed: 18832314]
- Kwong SM, Skurray RA, Firth N. *Staphylococcus aureus* multiresistance plasmid pSK41: analysis of the replication region, initiator protein binding and antisense RNA regulation. *Mol Microbiol* 2004;51:497–509. [PubMed: 14756789]
- Kwong SM, Skurray RA, Firth N. Replication control of staphylococcal multiresistance plasmid pSK41: an antisense RNA mediates dual-level regulation of Rep expression. *J Bacteriol* 2006;188:4404–4412. [PubMed: 16740947]
- Le Chatelier E, Janniere L, Ehrlich SD, Canceill D. The RepE initiator is a double-stranded and single-stranded DNA-binding protein that forms an atypical open complex at the onset of replication of plasmid pAMβ1 from Gram-positive bacteria. *J Biol Chem* 2001;276:10234–10246. [PubMed: 11124267]
- LeBard RJ, Jensen SO, Arnaiz IA, Skurray RA, Firth N. A multimer resolution system contributes to segregational stability of the prototypical staphylococcal conjugative multiresistance plasmid pSK41. *FEMS Microbiol Lett* 2008;284:58–67. [PubMed: 18492061]
- Lee PS, Grossman AD. The chromosome partitioning proteins Soj (ParA) and Spo0J (ParB) contribute to accurate chromosome partitioning, separation of replicated sister origins, and regulation of replication initiation in *Bacillus subtilis*. *Mol Microbiol* 2006;60:853–869. [PubMed: 16677298]
- Lin C-F, Chung T-C. Cloning of Erythromycin-Resistance Determinants and Replication Origins from Indigenous Plasmids of *Lactobacillus reuteri* for Potential Use in Construction of Cloning Vectors. *Plasmid* 1999;42:31–41. [PubMed: 10413663]
- Lioy VS, Martin MT, Camacho AG, Lurz R, Antelmann H, Hecker M, Hitchin E, Ridge Y, Wells JM, Alonso JC. pSM19035-encoded ζ toxin induces stasis followed by death in a subpopulation of cells. *Microbiology* 2006;152:2365–2379. [PubMed: 16849801]
- Makarova K, Slesarev A, Wolf Y, Sorokin A, Mirkin B, Koonin E, Pavlov A, Pavlova N, Karamychev V, Polouchine N, Shakhova V, Grigoriev I, Lou Y, Rohksar D, Lucas S, Huang K, Goodstein DM, Hawkins T, Plengvidhya V, Welker D, Hughes J, Goh Y, Benson A, Baldwin K, Lee JH, Diaz-Muniz I, Dosti B, Smeianov V, Wechter W, Barabote R, Lorca G, Altermann E, Barrangou R, Ganesan B, Xie Y, Rawsthorne H, Tamir D, Parker C, Breidt F, Broadbent J, Hutkins R, O'Sullivan D, Steele J, Unlu G, Saier M, Klaenhammer T, Richardson P, Kozyavkin S, Weimer B, Mills D. Comparative genomics of the lactic acid bacteria. *Proceedings of the National Academy of Sciences* 2006;103:15611–15616.
- Massidda O, Mingoa M, Fadda D, Whalen MB, Montanari MP, Valardo PE. Analysis of the β-lactamase plasmid of borderline methicillin-susceptible *Staphylococcus aureus*: Focus on *bla* complex genes and cadmium resistance determinants *cadD* and *cadX*. *Plasmid* 2006;55:114–127. [PubMed: 16229889]
- Møller-Jensen J, Jensen RB, Löwe J, Gerdes K. Prokaryotic DNA segregation by an actin-like filament. *EMBO J* 2002;21:3119–3127. [PubMed: 12065424]
- Mwangi MM, Wu SW, Zhou Y, Sieradzki K, de Lencastre H, Richardson P, Bruce D, Rubin E, Myers E, Siggia ED, Tomasz A. Tracking the *in vivo* evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proceedings of the National Academy of Sciences* 2007;104:9451–9456.
- Nakagawa I, Kurokawa K, Yamashita A, Nakata M, Tomiyasu Y, Okahashi N, Kawabata S, Yamazaki K, Shiba T, Yasunaga T, Hayashi H, Hattori M, Hamada S. Genome sequence of an M3 strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. *Genome Res* 2003;13:1042–1055. [PubMed: 12799345]

- Nakayama J, Yoshida K, Kobayashi H, Isogai A, Clewell D, Suzuki A. Cloning and characterization of a region of *Enterococcus faecalis* plasmid pPD1 encoding pheromone inhibitor (*ipd*), pheromone sensitivity (*traC*), and pheromone shutdown (*traB*) genes. *J Bacteriol* 1995;177:5567–5573. [PubMed: 7559344]
- O'Brien FG, Price C, Grubb WB, Gustafson JE. Genetic characterization of the fusidic acid and cadmium resistance determinants of *Staphylococcus aureus* plasmid pUB101. *J Antimicrob Chemother* 2002;50:313–321. [PubMed: 12205055]
- O'Driscoll J, Glynn F, Fitzgerald GF, Sinderen Dv. Sequence analysis of the lactococcal plasmid pNP40: a mobile replicon for coping with environmental hazards. *J Bacteriol* 2006;188:6629–6639. [PubMed: 16952955]
- Paulsen IT, Banerjee L, Myers GSA, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, Tettelin H, Dodson RJ, Umayam L, Brinkac L, Beanan M, Daugherty S, DeBoy RT, Durkin S, Kolonay J, Madupu R, Nelson W, Vamathevan J, Tran B, Upton J, Hansen T, Shetty J, Khouri H, Utterback T, Radune D, Ketchum KA, Dougherty BA, Fraser CM. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis*. *Science* 2003;299:2071–2074. [PubMed: 12663927]
- Ringgaard S, Ebersbach G, Borch J, Gerdes K. Regulatory cross-talk in the double *par* locus of plasmid pB171. *J Biol Chem* 2007;282:3134–3145. [PubMed: 17092933]
- Romero P, López R, García E. Genomic organization and molecular analysis of the inducible prophage EJ-1, a mosaic myovirus from an atypical pneumococcus. *Virology* 2004;322:239–252. [PubMed: 15110522]
- Sanchez-Hidalgo M, Maqueda M, Galvez A, Abriouel H, Valdivia E, Martinez-Bueno M. The genes coding for enterocin EJ97 production by *Enterococcus faecalis* EJ97 are located on a conjugative plasmid. *Appl Environ Microbiol* 2003;69:1633–1641. [PubMed: 12620853]
- Schumacher MA. Structural biology of plasmid partition: uncovering the molecular mechanisms of DNA segregation. *Biochem J* 2008;412:1–18. [PubMed: 18426389]
- Schumacher MA, Glover TC, Brzoska AJ, Jensen SO, Dunham TD, Skurray RA, Firth N. Segrosome structure revealed by a complex of ParR with centromere DNA. *Nature* 2007;450:1268–1271. [PubMed: 18097417]
- Sebahia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R, Thomson NR, Roberts AP, Cerdeño-Tárraga AM, Wang H, Holden MTG, Wright A, Churcher C, Quail MA, Baker S, Bason N, Brooks K, Chillingworth T, Cronin A, Davis P, Dowd L, Fraser A, Feltwell T, Hance Z, Holroyd S, Jagels K, Moule S, Mungall K, Price C, Rabinowitsch E, Sharp S, Simmonds M, Stevens K, Unwin L, Whithead S, Dupuy B, Dougan G, Barrell B, Parkhill J. The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. *Nature Genetics* 2006;38:786.
- Simpson AE, Skurray RA, Firth N. A single gene on the staphylococcal multiresistance plasmid pSK1 encodes a novel partitioning system. *J Bacteriol* 2003;185:2143–2152. [PubMed: 12644483]
- Stanley E, Walsh L, van der Zwet A, Fitzgerald GF, van Sinderen D. Identification of four loci isolated from two *Streptococcus thermophilus* phage genomes responsible for mediating bacteriophage resistance. *FEMS Microbiology Letters* 2000;182:271–277. [PubMed: 10620678]
- Takiguchi R, Hashiba H, Aoyama K, Ishii S. Complete nucleotide sequence and characterization of a cryptic plasmid from *Lactobacillus helveticus* subsp *jugurti*. *Appl Environ Microbiol* 1989;55:1653–1655. [PubMed: 2764571]
- Tanaka T, Ogura M. A novel *Bacillus natto* plasmid pLS32 capable of replication in *Bacillus subtilis*. *FEBS Letters* 1998;422:243–246. [PubMed: 9490016]
- Tettelin H, Maignani V, Cieslewicz MJ, Eisen JA, Peterson S, Wessels MR, Paulsen IT, Nelson KE, Margarit I, Read TD, Madoff LC, Wolf AM, Beanan MJ, Brinkac LM, Daugherty SC, DeBoy RT, Durkin AS, Kolonay JF, Madupu R, Lewis MR, Radune D, Fedorova NB, Scanlan D, Khouri H, Mulligan S, Carty HA, Cline RT, Van Aken SE, Gill J, Scarselli M, Mora M, Iacobini ET, Brettoni C, Galli G, Mariani M, Vegni F, Maione D, Rinaudo D, Rappuoli R, Telford JL, Kasper DL, Grandi G, Fraser CM. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *Proceedings of the National Academy of Sciences* 2002;99:12391–12396.
- Thomas C. Paradigms of plasmid organization. *Mol Microbiol* 2000;37:485–491. [PubMed: 10931342]

- Thompson JK, Foley S, McConville KJ, Nicholson C, Collins MA, Pridmore RD. Complete sequence of plasmid pLH1 from *Lactobacillus helveticus* ATCC15009: Analysis reveals the presence of regions homologous to other native plasmids from the host strain. *Plasmid* 1999;42:221–235. [PubMed: 10545264]
- Tinsley E, Naqvi A, Bourgogne A, Koehler TM, Khan SA. Isolation of a Minireplicon of the Virulence Plasmid pXO2 of *Bacillus anthracis* and Characterization of the Plasmid-Encoded RepS Replication Protein 10.1128/JB.186.9.2717–2723.2004. *J Bacteriol* 2004;186:2717–2723. [PubMed: 15090513]
- Weaver, KE. Enterococcal Genetics. In: Fischetti, VA.; Novick, RP.; Ferretti, JJ.; Portnoy, DA.; Rood, JI., editors. *Gram-Positive Pathogens*. Washington, DC: ASM Press; 2006. p. 312-331.
- Weaver KE, Clewell DB, An F. Identification, characterization, and nucleotide sequence of a region of *Enterococcus faecalis* pheromone-responsive plasmid pAD1 capable of autonomous replication. *J Bacteriol* 1993;175:1900–1909. [PubMed: 8384618]
- Weaver KE, Jensen KD, Colwell A, Sriram SI. Functional analysis of the *Enterococcus faecalis* plasmid pAD1-encoded stability determinant par. *Mol Microbiol* 1996;20:53–63. [PubMed: 8861204]
- Weaver KE, Tritle DJ. Identification and characterization of an *Enterococcus faecalis* plasmid pAD1-encoded stability determinant which produces two small RNA molecules necessary for its function. *Plasmid* 1994;32:168–181. [PubMed: 7531349]
- Weaver KE, Walz KD, Heine MS. Isolation of a derivative of *Escherichia coli*-*Enterococcus faecalis* shuttle vector pAM401 temperature sensitive for maintenance in *E. faecalis* and its use in evaluating the mechanism of pAD1 *par*-dependent plasmid stabilization. *Plasmid* 1998;40:225–232. [PubMed: 9806859]
- Weigel LM, Clewell DB, Gill SR, Clark NC, McDougal LK, Flannagan SE, Kolonay JF, Shetty J, Killgore GE, Tenover FC. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. *Science* 2003;302:1569–1571. [PubMed: 14645850]
- Williams LE, Dettler C, Barry K, Lapidus A, Summers AO. Facile recovery of individual high-molecular-weight, low-copy-number natural plasmids for genomic sequencing. *Appl Environ Microbiol* 2006;72:4899–4906. [PubMed: 16820486]
- Yamaguchi T, Hayashi T, Takami H, Ohnishi M, Murata T, Nakayama K, Asakawa K, Ohara M, Komatsuzawa H, Sugai M. Complete nucleotide sequence of a *Staphylococcus aureus* exfoliative toxin B plasmid and identification of a novel ADP-ribosyltransferase, EDIN-C. *Infect Immun* 2001;69:7760–7771. [PubMed: 11705958]
- Yazdankhah SP, Asli AW, Sorum H, Oppegaard H, Sunde M. Fusidic acid resistance, mediated by *fusB*, in bovine coagulase-negative staphylococci. *J Antimicrob Chemother* 2006;58:1254–1256. [PubMed: 17060340]
- Zhang YQ, Ren SX, Li HL, Wang YX, Fu G, Yang J, Qin ZQ, Miao YG, Wang WY, Chen RS, Shen Y, Chen Z, Yuan ZH, Zhao GP, Qu D, Danchin A, Wen YM. Genome-based analysis of virulence genes in a non-biofilm-forming *Staphylococcus epidermidis* strain (ATCC 12228). *Mol Microbiol* 2003;49:1577–1593. [PubMed: 12950922]
- Zhu W, Clark NC, McDougal LK, Hageman J, McDonald LC, Patel JB. Vancomycin-resistant *Staphylococcus aureus* isolates associated with Inc18-like *vanA* plasmids in Michigan. *Antimicrob Agents Chemother* 2008;52:452–457. [PubMed: 18056272]

Figure 1.

Phylogenetic tree of RepA homologs. The RepA homologs were retrieved from the NCBI database by a standard BLAST search using pAD1 RepA as query. The tree was drawn using the CLC Combined Workbench sequence analysis software at default settings; bootstrap values are indicated at nodes. Clusters described in the text are bracketed and labeled with the associated genera. The following species abbreviations are used: Bs, *Bacillus subtilis*; Efm, *Enterococcus faecium*; Efs, *Enterococcus faecalis*; Lbb, *Lactobacillus brevis*; Lbc, *Lactobacillus casei*; Lbh, *Lactobacillus helveticus*; Lbp, *Lactobacillus paracasei*; Lbsak, *Lactobacillus sakei*; Lbsal, *Lactobacillus salivarius*; Lcl, *Lactococcus lactis*; Lm chrom, *Leuconostoc mesenteroides* chromosomal locus; Sa, *Staphylococcus aureus*; Se, *Staphylococcus epidermidis*; Sag chrom, *Streptococcus agalactiae* chromosomal locus; Sh, *Staphylococcus haemolyticus*; Spn, *Streptococcus pneumoniae*; Spy, *Streptococcus pyogenes*; Ss, *Staphylococcus saprophyticus*; S sp, undefined *Staphylococcus* species strain 693-2; Sth, *Streptococcus thermophilus*; Sw, *Staphylococcus warneri*. The references and accession numbers for these sequences are provided in Table 1.

Figure 2.

Alignment of select RepA homologs from deeply branching groups. Representative plasmids were selected from each deeply branching group and aligned using CLC Combined Workbench at default settings. Abbreviations are as in Fig. 1 and plasmids are shown in the same order in which they appear in Fig. 1. The RepA_N consensus sequence was obtained from the NCBI website. Amino acids are colored according to a polarity coloring scheme (red = acidic, blue = basic, green = polar, black = non-polar). The final ~30 amino acids of the alignment were deleted because only very few of the proteins had C-terminal amino acids extending beyond those shown here.

Figure 3.

Organization of direct and inverted repeats in *repA* genes of select plasmids. Large primary repeats are shown as solid arrows above the sequence. Secondary inverted repeats internal to the primary repeats are shown as open arrows over the sequence. Note that in the case of pAD1, pCF10, pSK41, pRUM, EJ-1 and pSK1 the inverted repeats converge on a poly-A tract. Although no obvious inverted repeats are present in the pSAK1 sequence, a poly-A tract is present near the center of the primary repeats.

Figure 4.

Phylogenetic tree of RepB homologs. The RepB homologs were retrieved from the NCBI database by a standard BLAST search using pAD1 RepB as the query. The tree was drawn using the CLC Combined Workbench sequence analysis software at default settings; bootstrap values are indicated at nodes. The RepB homologs from Enterococcal RepA_N replicons are marked with thick arrows, from RepA_N replicons from other genera with thin arrows, and from inc18 replicons with stars. The pMP118 *Lactobacillus salivarius* plasmid actually contains both a Rep₃-like initiator protein and a divergent RepA_N protein that does not cluster with any of the homologs shown in Fig. 1. The following species abbreviations are used: Ba, *Bacillus anthracis*; Bc, *Bacillus cereus*; Bt, *Bacillus thuringiensis*; Cp, *Clostridium*

perfringens; Efm, *E. faecium*; Efs, *E. faecalis*; Lbb, *L. brevis*; Lbc, *L. casei*; Lbpl, *Lactobacillus plantarum*; Lbr, *Lactobacillus reuteri*; Lbsal, *L. salivarius*; Lcl, *L. lactis*; Lin, *Listeria innocua*; Lm, *L. mesenteroides*; Pea, *Pediococcus acidilactici*; Spy, *S. pyogenes*; Ss, *S. saprophyticus*; Sw, *S. warneri*.

Figure 5.

Organization of RepC binding sites in pAD1 and putative binding sites in other RepBC-containing plasmids. Vertically lined boxes depict repeats in the orientation shown on the right side of the figure. Horizontally lined boxes depict repeats present in the inverted orientation. Numbers in the boxes depict total number of repeats in the cluster. Boxes without numbers contain only a single repeat. Numbers between the boxes depict the size in nucleotides of non-repetitive spacer regions. pAD1 and pAM373 show two different repeat organizations in related pheromone-responsive *E. faecalis* plasmids. pXO2 encodes a RepBC (type Ic) partition module associated with an inc18-like initiator protein (Tinsley et al. 2004). pTE15 (Lin and Chung 1999) contains the most closely related homolog of pAD1 RepB but does not have the type Ic organization of upstream and downstream repeats. The ParAB locus (*par2*) of *E. coli* plasmid pB171 is only distantly related to the pAD1 ParBC partition locus, but also possesses upstream and downstream repeats (Ebersbach and Gerdes 2001).

Figure 6.

Phylogenetic tree of chromosomally-encoded RepA_N homologs. RepA_N homologs were identified using the consensus RepA_N sequence as query in a BLAST search at the NCBI database. The tree was drawn using the CLC Combined Workbench sequence analysis software at default settings; bootstrap values are indicated at nodes. Select representatives from the plasmid RepA_N replication initiators were included for reference. Two RepA_N proteins associated with known conjugative transposons are boxed. Chromosomal RepA_N proteins segregate neatly into size groups which are indicated with brackets and size range on the right.

Table 1
Genomes carrying *repA_N* homologues

Organism/Plasmid	GenBank Acc.		Plasmid Phenotype/comment [†]	References
	RepA_N protein	Core Nucleotide [*]		
<i>Staphylococcus aureus</i>				
EDINA plasmid	YP_001573906	NC_010077	As, Cd, bacteriocin, enterotoxin	unpublished
pBORa53	AAX99125	AY917098	Pn, Cd	(Massidda et al. 2006)
PETB	NP_478361	NC_003265	Cd, exotoxin, remnant <i>repA_N</i> gene	(Yamaguchi et al. 2001)
pI9789::Tn522	AAF63254	AF203377 (partial)	Pn, As, Cd, Hg	(Firth et al. 2000)
pIP680	AAF24089	AF117259 (partial)	Sg	(Allignet and ElSolh 1999)
pLEW6932	YP_001096314	NC_009130	As	(Williams et al. 2006)
pLW1043	NP_877980	NC_005054	Gm, Km, Tb, Tp, Vn, Smr, Tra	(Weigel et al. 2003)
VRSAp (pMu50)	NP_115290	NC_002774	Gm, Tb, Km, Qac	(Kuroda et al. 2001)
pMW2	NP_863271	NC_005011	Pn, Cd, remnant <i>repA_N</i> gene	(Baba et al. 2002)
pN315	NP_395537	NC_003140	Pn, Cd, As	(Kuroda et al. 2001)
pSA1379	YP_536846	NC_007931	Pn, Cd, As	unpublished
pSAS	none	NC_005951	Pn, Cd, bacteriocin, remnant <i>repA_N</i> gene	(Holden et al. 2004)
pSJH101	YP_001312295	NC_009619	Cd, Pn, Gm, Km, Tb	(Mwangi et al. 2007)
pSK1	AAF63252	AF203376 (partial)	Gm, Tb, Km, Tp, Qac	(Firth et al. 2000)
pSK41	NP_863615	NC_005024	Gm, Km, Tb, Nm, Smr, Tra	(Berg et al. 1998)
pSR1	AAF99572	AF167161 (partial)	Em	unpublished
pUB101	NP_932180	NC_005127	Pn, Cd, Fa	(O'Brien et al. 2002)
pUSA03	YP_492687	NC_007792	Em, Mp, Tra	(Diep et al. 2006)
pUSA300-HOU- MR	YP_001569048	NC_010063	Pn, Cd, Km, Nm, St, Bc	(Highlander et al. 2007)
pUSA300-HOU-MS	YP_001569077	NC_010066	Pn, Cd, bacteriocin	(Highlander et al. 2007)
<i>Staphylococcus epidermidis</i>				
pSE-12228-04	NP_863230	NC_005005		(Zhang et al. 2003)
pSE-12228-05	NP_863225	NC_005004	Pn	(Zhang et al. 2003)
pSERP	YP_187561	NC_006663	Pn, Nm, Km, Sm	(Gill et al. 2005)
<i>Staphylococcus haemolyticus</i>				
pNVH96	CAC16673	AJ302698	Pn, Fa	(Yazdankhah et al. 2006)

Organism/Plasmid	GenBank Acc.		Plasmid Phenotype/comment [†]	References
	RepA_N protein	Core Nucleotide [*]		
pNVH97A	CAB94811	AJ400722 (partial)	Pn, Qac	(Anthonisen et al. 2002)
<i>Staphylococcus saprophyticus</i>				
pSSP1	YP_302561	NC_007351	As	(Kuroda et al. 2005)
pSSP2	YP_302585	NC_007352	Cd	(Kuroda et al. 2005)
<i>Staphylococcus xylois</i>				
pSX267	CAA63141	X92404 (partial)	As	(Gering et al. 1996)
<i>Staphylococcus warneri</i>				
pPI-1	NP_940781	NC_005207	bacteriocin	(Aso et al. 2005)
<i>Enterococcus faecalis</i>				
pAD1	AAB00503	L01794 (partial)	exotoxin, pheromonerresponse, Tra	(Weaver et al. 1993)
pAM373	NP_071995	NC_002630	pheromone response, Tra	(De Boever et al. 2000)
pCF10	YP_195765	NC_006827	Tc, pheromonerresponse, Tra	(Hedberg et al. 1996)
pEJ97-1	CAD35304	AJ490170 (partial)	bacteriocin, pheromone response, Tra	(Sanchez-Hidalgo et al. 2003)
pPD1	BAA11194	D78016 (partial)	bacteriocin, pheromone response, Tra	(Nakayama et al. 1995)
pTEF1	NP_816932	NC_004669	Gm, Tm, Km, Smr, pheromone response, Tra	(Paulsen et al. 2003)
pTEF2	NP_817022	NC_004671	pheromone response, Tra	(Paulsen et al. 2003)
<i>Enterococcus faecium</i>				
DO plasmid	ZP_00603532	NZ_AAAK0 3000014	putative plasmid only	unpublished
DO plasmid 2	ZP_00602496	NZ_AAAK0 3000001	putative plasmid only	unpublished
pRUM	NP_863172	NC_005000	Cm, Em, Km, Sm	(Grady and Hayes 2003)
<i>Lactococcus lactis</i>				
pCI2000	AAF27302	AF154674 (partial)		(Kearney et al. 2000)
plasmid 3	YP_796521	NC_008505		(Makarova et al. 2006)
pNP40	ABG00341	DQ534432	Cd, Ns, Tra	(O'Driscoll et al. 2006)
<i>Lactobacillus brevis</i>				
plasmid 1	YP_796401	NC_008498		(Makarova et al. 2006)
<i>Lactobacillus casei</i>				
plasmid 1	YP_796447	NC_008502		(Makarova et al. 2006)
<i>Lactobacillus helveticus</i>				

Organism/Plasmid	GenBank Acc.		Plasmid Phenotype/comment [†]	References
	RepA_N protein	Core Nucleotide [*]		
pLH1	NP_052201	NC_002102		(Thompson et al. 1999)
pLJ1	NP_040419	NC_001379		(Takiguchi et al. 1989)
<i>Lactobacillus paracasei</i> plasmid	ABA12819	AY673957		(Desmond et al. 2005)
<i>Lactobacillus sakei</i> pSAK1	CAA90733	Z50862 (partial)		unpublished
<i>Lactobacillus salivarius</i> pSF118-44	YP_163778	NC_006530		unpublished
pSF118-20	YP_163743	NC_006529		unpublished
<i>Bacillus natto</i> pLS32	BAA24877	D49467 (partial)		(Tanaka and Ogura 1998)
Phage				
<i>Streptococcus pneumoniae</i> phage EJ-1	NP_945250	NC_005294		(Romero et al. 2004)
<i>Streptococcus thermophilus</i> phage 7201	NP_038304	NC_002185		(Stanley et al. 2000)
Conjugative transposon				
<i>Enterococcus faecalis</i> 268-10	AAF72340	AF192329	Carried by conjugative transposon Tn1549	(Garnier et al. 2000)
<i>Bacteroides uniformis</i> WH207	AAR05651	AY345595	Carried by conjugative transposon CTnBST	(Gupta et al. 2003)
Chromosomal loci				
<i>Alkaliphilus metalliredigens</i> QYMF	YP_001321733	NC_009633	MGE?	unpublished
<i>Anaerofustis stercorihominis</i> DSM 17244	ZP_02862671	NZ_ABIL02 000006 (partial)	MGE?	unpublished
<i>Anaerotruncus colihominis</i> DSM 17241	ZP_02443791	NZ_ABGD0 2000024, (partial)	MGE?	unpublished
	ZP_02442959	NZ_ABGD0 2000018, (partial)	MGE?	unpublished
	ZP_02442802	NZ_ABGD0 2000014, (partial)	MGE?	unpublished
<i>Bacteroides capillosus</i> ATCC 29799	ZP_02036899	NZ_AAXG0 2000015, (partial)	MGE?	unpublished
	ZP_02037049	NZ_AAXG0 2000016, (partial)	MGE?	unpublished
	ZP_02037346	NZ_AAXG0 2000028, (partial)	MGE?	unpublished
<i>Clostridium bolteae</i> ATCC BAA-613	EDP17002	ABCC02000 025, (partial)	MGE?	unpublished
	ZP_02089120	NZ_ABCC0 2000057, (partial)	MGE?	unpublished
	ZP_02086747	NZ_ABCC0 2000034, (partial)	MGE?	unpublished
<i>Clostridium difficile</i> 630	YP_001086875	NC_009089	MGE?	(Sebaihia et al. 2006)
<i>Clostridium leptum</i> DSM753	ZP_02081939	NZ_ABCB0 2000021, (partial)	MGE?	unpublished
<i>Clostridium scindens</i> ATCC 35704	ZP_02432545	NZ_ABFY0 2000037, (partial)	MGE?	unpublished
<i>Clostridium</i> sp. SS2/1	ZP_02438503	NZ_ABGC0 3000029, (partial)	MGE?	unpublished

Organism/Plasmid	GenBank Acc.		Plasmid Phenotype/comment [†]	References
	RepA_N protein	Core Nucleotide [*]		
<i>Coprococcus eutactus</i> ATCC 27759	ZP_02207641	NZ_ABEY0 2000028, (partial)	MGE?	unpublished
<i>Dorea formicigenerans</i> ATCC 27755	ZP_02235829	NZ_AAXA0 2000015, (partial)		unpublished
	ZP_02233650	NZ_AAXA0 2000008, (partial)	MGE?	unpublished
<i>Dorea longicatena</i> DSM 13814	ZP_01995573	NZ_AAXB0 2000007, (partial)		unpublished
	ZP_01994506	NZ_AAXB0 2000002, (partial)	MGE?	unpublished
	ZP_01994260	NZ_AAXB0 2000001, (partial)	MGE?	unpublished
<i>Enterococcus faecalis</i> V583	NP_815989	NC_004668	MGE?	(Paulsen et al. 2003)
<i>Eubacterium siraeum</i> DSM 15702	ZP_02421265	NZ_ABCA0 3000019 (partial)		unpublished
<i>Eubacterium ventriosum</i> ATCC 27560	ZP_02025677	NZ_AAVL0 2000030, (partial)	MGE?	unpublished
	ZP_02026446	NZ_AAVL0 2000035, (partial)		unpublished
<i>Faecalibacterium prausnitzii</i> M21/2	ZP_02090534	NZ_ABED0 2000020, (partial)	MGE?	unpublished
	ZP_02090896	NZ_ABED0 2000023, (partial)	MGE?	unpublished
	ZP_02091628	NZ_ABED0 2000027, (partial)	MGE?	unpublished
	ZP_02091715	NZ_ABED0 2000027, (partial)		unpublished
<i>Finegoldia magna</i>	YP_001692294	NC_010376	MGE?	(Goto et al. 2008)
<i>Leuconostoc mesenteroides</i> subsp. <i>mesenteroides</i> ATCC 8293	YP_818737	NC_008531		(Makarova et al. 2006)
<i>Peptostreptococcus micros</i> ATCC 33270	ZP_02093836	NZ_ABEE0 2000016, (partial)	MGE?	unpublished
<i>Ruminococcus gnavus</i> ATCC 29149	ZP_02041516	NZ_AAYG0 2000017, (partial)		unpublished
	ZP_02041518	NZ_AAYG0 2000017, (partial)		unpublished
	ZP_02042675	NZ_AAYG0 2000031, (partial)	MGE?	unpublished
<i>Ruminococcus obeum</i> ATCC 29174	ZP_01962661	NZ_AAVO0 2000001, (partial)	MGE?	unpublished
	ZP_01962946	NZ_AAVO0 2000002, (partial)	MGE?	unpublished
	ZP_01962958	NZ_AAVO0 2000002, (partial)	MGE?	unpublished
	ZP_01964360	NZ_AAVO0 2000008, (partial)		unpublished
	ZP_01964512	NZ_AAVO0 2000009, (partial)	MGE?	unpublished
	ZP_01965708	NZ_AAVO0 2000021, (partial)	MGE?	unpublished
<i>Ruminococcus torques</i> ATCC 27756	ZP_01968338	NZ_AAVP0 2000010, (partial)	MGE?	unpublished
<i>Staphylococcus aureus</i> MRSA252	none	BX571856	<i>repA_N</i> remnant on an integrated plasmid	(Holden et al. 2004)
<i>Staphylococcus epidermidis</i> ATCC 12228	NP_765060	NC_004461		(Zhang et al. 2003)
<i>Streptococcus agalactiae</i> 2603V/R	NP_688297	NC_004116		(Tettelin et al. 2002)
<i>Streptococcus agalactiae</i> NEM316	NP_735595	NC_004368	MGE?	(Glaser et al. 2002)
<i>Streptococcus dysgalactiae</i> subsp. <i>equisimilis</i>	ABV55455	EU142041	MGE?	unpublished
<i>Streptococcus pyogenes</i>	YP_603173	NC_008024	Prophage?	(Beres et al. 2006)

Organism/Plasmid	GenBank Acc.		Plasmid Phenotype/comment [†]	References
	RepA_N protein	Core Nucleotide [*]		
<i>Streptococcus pyogenes</i> MGAS315	NP_664495	NC_004070	Carried by an integrated prophage	(Beres et al. 2002)
<i>Streptococcus pyogenes</i> MGAS8232	NP_606602	NC_003485	Carried by an integrated prophage	(Beres et al. 2002)
<i>Streptococcus pyogenes</i> SSI-1	NP_802421	NC_004606	Carried by an integrated prophage	(Nakagawa et al. 2003)
<i>Streptococcus pyogenes</i> str. Manfredo	YP_001128066	NC_009332	Carried by an integrated prophage	(Holden et al. 2007)
<i>Streptococcus suis</i> 05ZYH33	YP_001198347	NC_009442		(Chen et al. 2007)
<i>Streptococcus suis</i> 89/1591	ZP_00876142	NZ_AAFA0 2000134, (partial)		unpublished
<i>Ureaplasma urealyticum</i>	ZP_02548168	NZ_AAYQ0 100000, (partial)	Prophage?	unpublished

* GenBank entries with incomplete plasmid or chromosome genome sequences are denoted (partial).

[†] MGE? indicates *repA_N* is found in close proximity to gene(s) with products that are associated with mobile, conjugative or integrative elements (such as recombinases/invertases/integrases, VirD2-like relaxases, VirD4-like type IV secretion proteins, replicative helicases and/or transposases) and therefore may constitute a mobile genetic element. Prophage? indicates that flanking gene(s) are annotated as phage-like.

As, arsenate resistance; Bc, bacitracin resistance; Cd, cadmium resistance; Cm, chloramphenicol resistance; Gm, gentamicin resistance; Em, erythromycin resistance; Fa, fusidic acid resistance; Hg, mercury resistance; Km, kanamycin resistance; MGE, mobile genetic element; Mp, mupirocin resistance; Nm, neomycin resistance; Ns, Nisin resistance; Pn, penicillin resistance; Qac, resistance to quarternary ammonium compounds; Sg, streptogramin resistance; Sm, streptomycin resistance; Smr, multidrug resistance; St, streptothricin resistance; Tb, tobramycin resistance; Tc, tetracycline resistance; Tp, trimethoprim resistance; Tra, conjugative transfer; Vn, vancomycin resistance.