Plasmid 61 (2009) 94–109

Contents lists available at ScienceDirect

Plasmid

journal homepage: www.elsevier.com/locate/yplas

Review

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

Keith E. Weaver [a,*], Stephen M. Kwong [b], Neville Firth [b], Maria Victoria Francia [c]

[a] Division of Basic Biomedical Sciences, Sanford School of Medicine, University of South Dakota, 414 East Clark Street, Vermillion, SD 57069, USA
[b] School of Biological Sciences, University of Sydney, NSW 2006, Australia
[c] Servicio de Micrbiologia, Hospital Universitario Marques de Valdecilla-IFIMAV, Avda. de Valdecilla s/n, 39008 Santander, Cantabria, Spain

ARTICLE INFO

Article history:
Received 6 August 2008
Revised 30 October 2008
Available online 6 January 2009

Communicated by Manuel Espinosa

Keywords:
Plasmid evolution
Gram-positive bacteria
Plasmid replication
Plasmid stability

ABSTRACT

The pheromone-responsive conjugative plasmids of *Enterococcus faecalis* and the multiresistance 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.

1. Introduction

Detailed studies of replication and stable inheritance of θ-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; Lioy 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.

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.

* Corresponding author. Fax: +1 605 677 6381.
E-mail address: kweaver@usd.edu (K.E. Weaver).

0147-619X/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.plasmid.2008.11.004

While the first description of the pheromone response was published in 1978 (Dunny et al., 1978), it was not 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, 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 subgroup 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 homologs 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 *Staphylococcus xylosis* 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* homologs 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.

2. 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, 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