Analysis of pMA67, a predicted rolling-circle replicating, mobilizable, tetracycline-resistance plasmid from the honey bee pathogen, *Paenibacillus larvae* ∗

K. Daniel Murray a,*, Katherine A. Aronstein a, Jesse H. de León b

a USDA-ARS, Honey Bee Research Unit, Kika de la Garza Subtropical Agricultural Center, 2413 E. Hwy 83, Weslaco, TX 78596, USA

b USDA-ARS, Beneficial Insects Research Unit, Kika de la Garza Subtropical Agricultural Center, 2413 E. Hwy 83, Weslaco, TX 78596, USA

Received 25 October 2005, revised 31 January 2007
Available online 23 March 2007
Communicated by Manuel Espinosa

Abstract

This work characterizes a recently discovered natural tetracycline-resistance plasmid called pMA67 from *Paenibacillus larvae*—a Gram-positive bacterial pathogen of honey bees. We provide evidence that pMA67 replicates by the rolling-circle mechanism, and sequence comparisons place it in the pMV158 family of rolling-circle replicons. The plasmid contains predicted *rep*, *cop*, and *rnlII* genes for control of replication initiating at a predicted double-strand origin. The plasmid has an *ssot* single-strand origin, which is efficient enough to allow only very small amounts of the single-stranded DNA intermediate to accumulate. The overall efficiency of replication is sufficient to render the plasmid segregationally stable without selection in *P. larvae* and in *Bacillus megaterium*, but not in *Escherichia coli*. The plasmid is expected to be mobilizable due to the presence of a *mob* gene and an *oriT* site. The plasmid contains a *tetL* gene, whose predicted amino acid sequence implies a relatively ancient divergence from all previously known plasmid-encoded *tetL* genes. We confirm that the *tetL* gene alone is sufficient for conferring resistance to tetracyclines. Sequence comparisons, mostly with the well-characterized pMV158, allow us to predict promoters, DNA and RNA secondary structures, DNA and protein motifs, and other elements.

Published by Elsevier Inc.

Keywords: *Paenibacillus larvae*; Rolling-circle replication; Plasmid; *tetL*; Tetracycline; American Foulbrood

1. Introduction

*Paenibacillus larvae* is a Gram-positive bacterial pathogen which causes American Foulbrood, the most serious infectious disease of honey bees. Commercial and hobbyist beekeepers have controlled this disease for decades with the antibiotic
oxytetracycline (OTC). However, *P. larvae* resistance to this antibiotic has become widespread in the past few years (Cox et al., 2005; Evans, 2003; Miyagi et al., 2000; Mussen, 2000). We recently discovered a plasmid in *P. larvae* conferring OTC-resistance, which we named pMA67, and found that among 36 strains tested from across North America, there was a perfect correlation between the presence of the plasmid and resistance to tetracyclines (Murray and Aronstein, 2006). The predicted OTC-resistance gene on this plasmid is *tetL*, and it is the first representative of the tetracycline resistance genes to be found in the *Paenibacillus* genus.

A preliminary analysis of pMA67 suggested that the plasmid replicates by the rolling-circle mechanism (Murray and Aronstein, 2006). Rolling circle replication (RCR) plasmids have a double-strand origin (*dso*), which is nicked by a plasmid-encoded Rep protein in order to initiate replication. The leading strand is extended from that nick, generating a full length single-stranded copy of the plasmid, which can often be detected in cells containing RCR plasmids. This single-stranded DNA (ssDNA) is then used as template in synthesis of the lagging strand, beginning at the single-strand origin (*sso*) (reviewed in Khan, 2005). There are no plasmid-encoded functions necessary for production of the lagging strand from the ssDNA intermediate.

Regulation of replication of RCR plasmids occurs mainly at initiation of leading strand synthesis at the *dso*, such that Rep protein concentration controls plasmid replication. The Rep concentration is regulated by countertranscribed RNAs (ctRNA) alone or in combination with a protein (del Solar et al., 1998). RCR plasmids seem to lack active partitioning systems, so that segregational stability is dependent on random distribution of plasmids to daughter cells.

In this work, we analyzed the DNA sequence of plasmid pMA67, and identified conserved sequences and putative secondary structures suspected to be important for various pMA67 functions. All genes on pMA67 are predicted by sequence to be involved with either plasmid replication, plasmid mobilization, or antibiotic resistance. In addition, we partially characterized the plasmid with respect to its *sso* function, physiological stability, and host range, and we demonstrated the functionality of the *tetL* gene. We also report the uniqueness of this *tetL* in terms of its relatively ancient divergence from other plasmid-encoded *tetL* genes.

2. Materials and methods

2.1. Southern hybridization

DNA was isolated from stationary phase cultures of *P. larvae* strain 67E (pMA67-containing; Murray and Aronstein, 2006) by the method of O’Sullivan and Klaenhammer (1993). DNA samples were run on a 1% TAE agarose gel at 60 V for 90 min in duplicate lanes, which were then separated. One was kept in the neutralization buffer (1 M Tris, pH 7.4, 1.5 M NaCl) while the other was treated with alkali (0.5 M NaOH, 1.5 M NaCl) for 45 min before neutralization. DNA was transferred by capillary action to a BrightStar Plus Nylon membrane (Ambion Inc., Austin, TX) in 10x SSC. The transfer procedure, hybridizations, and washes were done according to standard procedures (Sambrook and Russell, 2001). The probe used was a digoxigenin-labeled fragment of pMA67 made by PCR using primers 1776-F (GTG GTTGGGAAGCAAAACATAT) and 2371-R (GCTTTC CATATAGAGCTGTT) (Fig. 1), and detection was performed according to the manufacturer’s instructions (Roche Diagnostics, Mannheim, Germany).

2.2. Segregational stability of pMA67

Serial cultures of *P. larvae* strain SD1 (pMA67-containing; Murray and Aronstein, 2006) were grown for 24 h at 35 °C with shaking in liquid J media (St. Julian et al., 1963) without antibiotics, except for the first culture in the series, which contained 5 µg/ml tetracycline. Dilutions of each serial culture were spread on J plates without antibiotics, except for the first culture in the series, which contained 5 µg/ml tetracycline. Dilutions of each serial culture were spread on J plates without antibiotics, except for the first culture in the series, which contained 5 µg/ml tetracycline. Dilutions of each serial culture were spread on J plates with and without tetracycline (10 µg/ml), and these were also incubated under the same conditions. These plates were scored after three days incubation to obtain the fraction of tetracycline-resistant (Tet') colonies.

For *Escherichia coli* segregational stability experiments, serial cultures of strain YMC9 (Backman et al., 1981) transformed with pMA67 (or pBR322) were grown for 9–16 h at 37 °C with shaking in LB media without antibiotics, except for the first culture in the series, which contained 10 µg/ml tetracycline. Dilutions of each serial culture were spread on LB plates with or without tetracycline (10 µg/ml) and incubated at 37 °C. Plates were scored after 24 h of growth to obtain the fraction of Tet' colonies.

2.3. Bacterial transformations with pMA67 and screening

Commercially prepared protoplasts of *Bacillus megaterium* strain WH320 (MoBiTech, Goettingen, Germany) were used in transformation experiments according to...
the manufacturer’s instructions. The DNA used was 1 \mu g of pMA67 or pWH1520 (positive control; MoBiTech). For *E. coli*, chemical transformation of strain YMC9 was done using a standard protocol (Sambrook and Russell, 2001). The DNA used was 10 ng of pMA67 or pBR322 (positive control). For either bacterium, selection was for tetracycline resistance at 10 \mu g/ml. Negative control transformation attempts done without addition of any DNA yielded no colonies.

Screening of potential pMA67-carrying transformants was done initially by colony PCR with primers 1776-F and 2371-R (Fig. 1). Cycling conditions were 95 °C for 5 min, then 35 cycles of 94, 60, and 72 °C for 1 min each, followed by a final extension of 72 °C for 10 min. Plasmid minipreps of potential pMA67-carrying *B. megaterium* transformants were done by the method of O’Sullivan and Klaenhammer (1993). Plasmid minipreps of *E. coli* were done using an Eppendorf Fast Plasmid Mini kit (Eppendorf, Hamburg, Germany) according to the manufacturer’s instructions.

### 2.4. Construction and cloning of tetL alleles derived from pMA67

The full length tetL gene for cloning was generated by PCR with primers 1776-F and 3664-R (CAGTTT AGCGAGAAATTA), and the 3' truncated allele with primers 1776-F and 805-R (CCACAA AGGACACCAATTAT) (Fig. 1). The PCR cycling conditions for the full length tetL were 94 °C for 4 min, then 35 cycles of 94 °C 1 min, 60 °C 1 min and 72 °C for 2 min, followed by a final extension of 72 °C for 10 min, and were the same for the truncated allele except that annealing temperature was 56 °C and extension time was 1 min.

We gel purified the PCR products using a QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA) according to the manufacturer’s instructions, ligated these two fragments into the pCR2.1-TOPO vector (Invitrogen Corp., Carlsbad, CA), and transformed TOP10 (Invitrogen) cells according to the manufacturer’s instructions. Tetracycline at 10 \mu g/ml was used for selection of transformants, and for the truncated tetL construct, a separate selection for ampicillin (100 \mu g/ml) resistance was also done (vector contains the *bla* gene).

Colonies confirmed by PCR to have pMA67 tetL sequences (primers 1776-F/2371-R) were tested for resistance to 10 \mu g/ml tetracycline on LB plates at 37 °C. Confirmation of cloning of correct tetL constructs was done by digestion of plasmid minipreps with *Bgl*II, and PCR amplification with primer pairs 1776-F/805-R (for 5' half of tetL) and 741-F/3664-R (for 3' half of tetL). The 741-F primer sequence is ACAGAACCTTTTGTTGAACC.

### 2.5. Bioinformatics

Genetic database searches for homologies to pMA67 sequences were done using BLAST (Altschul et al., 1990). Accession numbers for sequences related to pMA67 used in this study are shown in Table 1.

For the phylogenetic analysis, the alignment software ClustalX (Thompson et al., 1997) and the phylogenetic program PAUP version 4.0b10 for the Macintosh (PPC) (Swofford, 2002) were utilized for alignment, bootstrapping (as percentage of 1000 replications) (Felsenstein, 1995), and reconstruction of trees as described in de León et al. (2006). Phylogenetic trees were constructed using both distance and maximum parsimony methods. For the distance analysis, the neighbor-joining algorithmic
method was performed utilizing the uncorrected 'p' genetic distance parameter (Saitou and Nei, 1987). For the parsimony analysis, heuristic searches for the most parsimonious trees were conducted using closest step-wise addition and the branch swapping algorithm by tree bisec-tion–reconnection. Gaps were treated as missing data and constraints were not enforced.

3. Results and discussion

3.1. \textit{pMA67} is a Group II RCR plasmid

The definitive result that confirms a plasmid replicates by the RCR mechanism is the presence of ssDNA in cells that harbor the plasmid. However, several RCR plasmids do not produce detectable ssDNA (discussed below). We were able to detect a small amount of ssDNA by Southern hybridization from cultures of \textit{pMA67}-containing \textit{P. larvae} (Fig. 2). The 5.0 kb \textit{pMA67} supercoiled monomer migrates at about the rate of a 3 kb linear standard, as has been seen previously (Murray and Aronstein, 2006). The ssDNA does not significantly stain with ethidium bromide (Fig. 2, lane 1), but is detected by hybridization (Fig. 2, lanes 2 and 3). When DNA is not denatured by alkali before transfer, only the DNA which is already single-stranded binds to the membrane (Fig. 2, lane 3).

RCR plasmids are classified on the basis of amino acid sequence homologies among their Rep proteins. There are more than a dozen groups (or families) recognized, with four major groups called I–IV (www.essex.ac.uk/bs/staff/osborn/DPR_home.htm). Using this system, \textit{pMA67} is classified as a Group II (also called the \textit{pMV158} family) plasmid. Indeed, based on nucleotide and amino acid sequence analyses of the entire plasmid, \textit{pMA67} is quite similar in size and overall organization to the archetypal Group II plasmid, \textit{pMV158} from \textit{Streptococcus agalactiae}, the main difference being the presence of a second \textit{sso} in \textit{pMV158}, which \textit{pMA67} lacks (Fig. 1).

3.2. Rep protein and double-strand replication origin

The best match in NCBI databases to the predicted Rep protein of \textit{pMA67} is the Rep protein from \textit{Lactobacillus sakei} plasmid \textit{pLS141-1} with
which it shares 80% amino acid identity. Rep proteins across the four major RCR plasmid groups have conserved motifs. Two of the more important ones are called motif 2 (or HUH motif), thought to be a part of a metal binding domain, and motif 3, part of the catalytic (nicking) domain (del Solar et al., 1998). Motifs 2 and 3 have consensus sequences of KkxHYHUUU and gxUxYUtHxxxD respectively in Group II plasmids (del Solar et al., 1998), where ‘x’ is any amino acid, ‘U’ is a hydrophobic residue, and uppercase and lowercase letters represent strongly- and weakly-conserved residues, respectively. Both of these motifs are present in the predicted pMA67 Rep protein, with sequences KKAHYHIYIV and NIYLYLTHESKD (corresponding to nucleotide positions 653–679 and 785–820, respectively in the GenBank sequence) being perfect matches to consensus with the exception of the first weakly conserved glycine residue of the motif 3 consensus. The tyrosine residue in the motif 3 consensus is required for nicking the dso nick site (del Solar et al., 1998), and this tyrosine is also present in the predicted pMA67 Rep protein.

A host encoded helicase called PcrA, apparently ubiquitous in Gram-positive bacteria has been shown to be required for replication of a Group I RCR plasmid (Iordanescu, 1993; Iordanescu and Basheer, 1991), and is considered likely to be involved in replication of RCR plasmids of all groups (Khan, 2005). The PcrA helicase of Staphylococcus aureus was shown to physically interact with the Rep protein of an RCR plasmid (Chang et al., 2002). PcrA helicases may play a role in determining RCR plasmid host range (Ruiz-Maso et al., 2006). The P. larvae genome has recently been sequenced (Qin et al., 2006), and it contains a likely pcrA gene (accession CH981424; ORF is at coordinates 16,906–19,220) whose predicted protein is more than 50% identical in amino acid sequence to PcrA proteins from several Bacillus spp. (not shown).

The dso's of RCR plasmids contain a Rep protein binding site and nick site. Typically, the sequence of the nick site is well conserved in all plasmids belonging to the same RCR group, but Rep binding site sequences are not as well conserved (Khan, 2005). However, Fig. 3 shows that there is a high degree of sequence identity between both the nick and Rep binding site regions of the dso's of pMA67 and pLS1—a Δ(ssoU–mobM) derivative of pMV158 (Lacks et al., 1986). pMV158/pLS1 must be supercoiled in order for in vitro replication to initiate, suggesting that hairpin formation at the dso nick site is required for nicking by Rep protein (Puyet et al., 1988). The actual site of nicking is in the loop of the hairpin (Puyet et al., 1988), and the 9 bp sequence in the loop is identical between pMV158/pLS1 and pMA67 (Fig. 3). The Rep binding site in Group II plasmids is associated with two or three direct repeats (del Solar et al., 1998). In pMV158/pLS1 and pMA67, there are three 11 bp direct repeats (Fig. 3).

![Fig. 3. Double strand origins and Pcr promoters known in pMV158/pLS1, and predicted in pMA67. Bases in common are shaded. Solid arrows indicate inverted repeats (IR). Dashed arrows indicate an IR present in pMV158/pLS1 but absent in pMA67. The site of nicking by the Rep protein in pMV158/pLS1, and the Rep protein binding site (dotted underline) are identified in del Solar et al. (1993a). The Pcr promoter -35 and -10 sequences of pMV158/pLS1 as well as potential promoter hexamers in pMA67 are solid underlined. The bold nucleotides are short IRs. The +1 nucleotide for the pMV158/pLS1 copG–repB transcript is shown. Numbered bases at beginning and end of sequence refer to base numbers in the GenBank sequences. (Sequences shown are of pLS1; the first base of the -35 hexamer is an “A” in the pMV158 GenBank entry.)](https://example.com/fig3.png)
3.3. Cop and RNAII regulators of rep expression

In Group II plasmids, the cop and rnaII gene products are negative regulators of rep expression at the transcriptional and translational levels, respectively. Cop is a repressor of transcription from Pcr, the promoter for rep, and RNAII is a cRNA which is thought to interfere with translation initiation from the rep mRNA (del Solar et al., 1995). For Group II plasmids, the regulation of plasmid replication has been best characterized for plasmid pLS1 (del Solar et al., 1993a). The pMV158/pLS1 copG and repB genes are cotranscribed onto a single polycistronic mRNA from promoter Pcr (Fig. 3). There is a putative cop gene in pMA67 upstream of rep (Fig. 1), as well as two possible promoters which may correspond to the Pcr promoter of pLS1, and there is a short IR overlapping each of the potential pMA67 Pcr promoters (Fig. 3). A similar IR in pMV158/pLS1 constitutes the central element of the binding site for the dimeric CopG protein (del Solar et al., 1990).

The predicted pMA67 cop gene would produce a protein of 56 amino acids, similar in length to the pMV158/pLS1 CopG protein of 45 amino acids. The predicted Cop protein of pMA67 has little amino acid homology to known Cop proteins present in NCBI databases. It does have significant homology to unidentified proteins predicted from four other RCR plasmids, pLS141-1 from L. sakei, pLC2 from Lactobacillus curvatus, pPF107-3 from Lactobacillus lactis, and pJB01 from Enterococcus faecium, ranging from 50% to 57% identity, and 69% to 73% similarity when conserved substitutions are included. The homolog from pJB01 was recently tentatively identified as a Cop protein (Kim et al., 2006).

A thorough structural analysis of the pMV158/pLS1 Cop protein revealed that the protein forms a ribbon–helix–helix structure (Gomis-Rüth et al., 1998). An amino acid sequence alignment of this protein to many known and putative Cop proteins, including those unknown proteins from pLC2 and pPF107-3, revealed several conserved amino acids in crucial positions throughout the group, including a highly conserved glycine mediating a turn between two α-helices, a serine or threonine residue also in the turn, and conservation of hydrophobicity at seven isolated residues necessary for maintaining the hydrophobic core of the dimer protein (Gomis-Rüth et al., 1998). The six most highly conserved hydrophobic amino acid positions in that analysis are also present in pMA67, as well as the conserved glycine in the turn (Fig. 4a). Interestingly, in the pMA67-encoded Cop protein, as well as in each of those pMA67-like predicted Cop proteins from the four Lactobacillus and Enterococcus plasmids mentioned above, an isoleucine residue occupies the position in the turn of the conserved serine/threonine.

The sequence downstream of cop in pMA67 predicts that an RNAII is produced by this plasmid, as in pMV158/pLS1. The promoter and terminator of the rnaII gene from pLS1 (del Solar and Espinosa, 1992), and the corresponding hypothetical structures in pMA67 are shown in Fig. 4b. Like most...
Group II plasmids (but unlike pMV158 itself), the promoter for the pMA67 ctRNA has a so-called "extended +10" motif (del Solar and Espinosa, 2001). These motifs have a consensus of TGNTA TAAT, and are associated with strong promoters, at least in E. coli (deHaseth et al., 1998). The stretch of T's after the stem-loops of the terminators suggest that they are rho-independent.

### 3.4. Single strand origin

Most research groups recognize four major types of ssos—ssoA, ssoT, ssoU, and ssoW (Andrup et al., 2003; Khan, 2005). The ssos of a particular type have only minimal sequence homology, but instead are recognized mainly by their common secondary structures (Khan, 1997), and for some types of ssos, consensus sequences within motifs (Andrup et al., 2003).

The pMA67 sso contains all the motifs and predicted secondary structures that are typical of ssoT origins (Fig. 5). An efficient sso is expected to result in little accumulation of the ssDNA intermediate in the cell. The inability to detect ssDNA in bacteria containing RCR plasmids with ssoT origins is quite common, as seen for plasmids pBAA1 (Seery and Devine, 1993), pG12 (Hoflack et al., 1999), pTX14-3 (Boe et al., 1991), and pTA1060 (Meijer et al., 1995). Fig. 5 compares the ssoT regions of pMA67, which produces a small amount of ssDNA (Fig. 2), with those four plasmids with apparently highly efficient ssoTs, and the plasmid in NCBI databases with the highest sso sequence identity to pMA67, pUIBI-1 (78% identity over the 170 bp region encompassing pMA67 nucleotides 3185-3348), which produces abundant ssDNA that can be detected simply by ethidium bromide staining of DNA preparations in agarose gels (Lopez-Meza et al., 2003).

The alignments suggest that a possible reason for the relatively poor function of the pMA67 and especially the pUIBI-1 ssoTs is the makeup of the elements of the first stem-loop, i.e., motifs I and II, and IR1. In motif I of pMA67 or pUIBI-1, neither the spacing between conserved bases nor the bases themselves match consensus. For motif II, there is a suboptimal spacing between conserved bases for pMA67, and the consensus bases are not all present in pUIBI-1. Perhaps most important for sso efficiency, the bases making up the conserved stem of IR1, which are immediately adjacent to motifs I and II in the other plasmids, are a few bases removed from those motifs in pUIBI-1. Given that the first stem-loop has been shown to be important for sso function in pBAA1 (Seery and Devine, 1993), pTA1060 (Meijer et al., 1998), our speculation is that the non-consensus elements of that stem-loop in pMA67 and pUIBI-1 contribute to poorer
function for the sso, and subsequently the accumulation of ssDNA.

3.5. Segregational stability of pMA67 in P. larvae

As seen, each of the genetic elements—dso, rep, cop, and rnaII—involves in leading strand initiation in pMV158/pLS1 has also been predicted by sequence in pMA67. The functionality of the replication initiation system can be partially assessed by determining the segregational stability of a plasmid. Therefore, we sought to investigate maintenance of the pMA67 plasmid in P. larvae over many generations in the absence of selection. This effort was somewhat hindered by the difficulty in maintaining viable P. larvae during serial subculturing, a phenomenon which has been previously reported for P. larvae (Heyndrickx et al., 1996; Piccini and Zunino, 2001). We determined the fraction of cells maintaining pMA67 without selection in a liquid culture series by daily plating without selection to generate colonies, then testing these colonies (typically 100 each day) for tetracycline resistance along with tetracycline-resistant and sensitive control strains. The longest growing culture series lasted 66 generations (10 days), and every colony tested over the entire period was found to be tetracycline resistant. Other attempts which resulted in fewer generations before culture viability was lost, nevertheless had similar results with respect to plasmid maintenance. These results indicate that the replication initiation mechanisms encoded in pMA67 function quite efficiently in P. larvae, promoting a high enough copy number to ensure distribution of plasmids to daughter cells. We suspect that the segregational stability of pMA67 in P. larvae may also be facilitated by the relatively slow growth rate of P. larvae, since it has been shown that a plasmid with a negative replicon control system and without an active partitioning system is more segregationally stable at slower bacterial growth rates (Lin-Chao and Bremer, 1986).

3.6. Establishment of pMA67 in alternate hosts

The host range of an RCR plasmid may be determined by the host’s ability to replicate either the leading or the lagging strand (Khan, 2005). For leading strand replication, the interaction (or lack thereof) between the host PcrA helicase and the plasmid Rep protein can affect the maintenance of an RCR plasmid in the host (Anand et al., 2004; Ruiz-Masó et al., 2006). For lagging strand replication, host RNA polymerase recognition (or not) of the plasmid sso may play a role in plasmid host range. There is evidence that ssoA and ssoW are host-specific, while ssoT and ssoU are thought to be able to be used by a number of Gram-positive hosts (Khan, 2005).

To determine whether pMA67 replication is possible in another Gram-positive species, we attempted transformations of the plasmid into B. megaterium, and were successful (data not shown). Growth rates of the B. megaterium transformants in rich media were almost the same in the presence or absence of tetracycline (not shown), implying that plasmid replication is able to keep pace with cell division, and therefore that the B. megaterium PcrA helicase interacts well with the pMA67 Rep protein, and that the B. megaterium RNA polymerase recognizes the pMA67 ssoT.

Natural RCR plasmids are also known in Gram-negative hosts (www.essex.ac.uk/bs/staff/osborn/DPR_home.htm). In addition, although isolated from a Gram-positive host, pMV158 is able to replicate in some Gram-negative bacteria (del Solar et al., 1993b; Hernandez-Arriaga et al., 2000; Lacks et al., 1986). We attempted transformation of E. coli with pMA67, and found this also to be successful. Transformants of both B. megaterium and E. coli selected with tetracycline were confirmed as authentic both by pMA67-specific colony PCR and plasmid minipreps (not shown).

Unlike the B. megaterium transformants, our E. coli pMA67 transformants grew very slowly in rich media in the presence of tetracycline (not shown). To determine if this was due to poor replication of pMA67, we spread dilutions of an E. coli pMA67 transformant culture on plates with and without tetracycline at various intervals during serial growth in liquid cultures in the absence of selection (LB media at 37 °C), and found that the fraction of pMA67-containing cells dropped rapidly (Fig. 6). These results are similar to those seen in Hernandez-Arriaga et al. (2000) in experiments with pMV158 and pLS1 in E. coli. The instability of pMA67 in E. coli indicates that the slow growth previously observed in the presence of tetracycline is due to poor pMA67 replication, and suggests that either leading or lagging strand synthesis (or both) of pMA67 in E. coli is partially defective. Leading strand synthesis of pMV158 and pMA67 in E. coli probably requires interaction between the Rep protein and the host UvrD helicase. This helicase has
significant sequence homology to PcrA helicases from Gram-positive bacteria, and was shown to be required for in vivo leading strand synthesis of Group II plasmid pE194 in *E. coli* (Bruand and Ehrlich, 2000). Hernandez-Arriaga et al. (2000) showed that lagging strand synthesis of pMV158 in *E. coli* initiated at the ssoU, which had a low, but significant functionality. Since pMA67 can replicate in *E. coli*, presumably the *E. coli* RNA polymerase is also able to utilize the pMA67 ssoT to some degree. We have not performed experiments to determine whether leading or lagging strand synthesis is limiting for overall pMA67 replication in *E. coli*. The instability of pMA67 in *E. coli* does not necessarily imply that pMA67 or closely-related plasmids could have no significant role in Gram-negative species in natural environments, where the growth rates would typically be slower than those afforded by optimal laboratory growth conditions, and therefore plasmids would be more segregationally stable.

Attempts to transform an *E. coli* recA mutant strain (TOP10) with pMA67 was unsuccessful in multiple trials (data not shown), consistent with the result from Lacks et al. (1986) in which they showed that *E. coli* strains must be recA+ for successful transformation with plasmids pMV158 and pLS1.

### 3.7. Mob protein and oriT site

The *mob* genes of mobilizable plasmids encode DNA relaxases which nick supercoiled DNA in a site- and strand-specific fashion at the *oriT* site in preparation for conjugation (Francia et al., 2004). BLAST searches indicated that the predicted pMA67 Mob protein is most closely related to the Mob proteins from plasmids pUIBI-1 and pBMB1 from *Bacillus thuringiensis*, and pBMY1 from *Bacillus mycoides*, with 50–55% identities and up to 73% similarity when conserved substitutions are included. A classification scheme for mobilization regions of plasmids proposed by Francia et al. (2004) identifies four main superfamilies of mobilizable plasmids. Plasmids pUIBI-1 and pBMY1 are included in that analysis, and they are placed in the pMV158 Mob superfamily (these are different plasmid groupings than those based on the Rep proteins), indicating that pMA67 would belong in that family also. The superfamily is characterized by an N-terminus proximal motif with consensus ‘HxxR’, followed by a putative catalytic NY(D/E)L motif, and a downstream ‘HxDEXphxh’ motif (Francia et al., 2004; Guzmán and Espinosa, 1997). Each of these is present in the predicted amino acid sequence of the pMA67 Mob protein—HLDR, NYDL, and HNDTTPHM (corresponding to nucleotide positions 3567–3578, 3630–3641, and 3864–3893, respectively, in the GenBank sequence).

Guzmán and Espinosa (1997) identified the *oriT* in pMV158. It contained two overlapping IRs, and had a GTGTTG loop in hairpin 2 (Fig. 7). An analysis of plasmids in the pMV158 Mob superfamily has shown that there is a predicted 7–10 bp stem and usually a 6 bp loop at the presumed *oriT* sites (Francia et al., 2004). The overlapping IRs, the 7–10 bp stems, and the 6 bp loop are all present in pMA67 (Fig. 7). The first nucleotide of the hairpin loop in *oriT* sites ranges from 47 to 140 bp upstream of *mob* gene start codons in RCR plasmids from various superfamilies (Meijer et al., 1998), and this distance is 71 bp in pMA67.

![Fig. 6. Segregational stability of pMA67 in *Escherichia coli*.](image)

Percent of YMC9 cells with plasmid during growth without selection was determined by plating dilutions of serial cultures and counting colonies arising on plates with or without tetracycline. △, pMA67; ○, pBR322 (control). No Tet+ colonies from pMA67 cultures were obtained at 31 and 42 generations.

![Fig. 7. *oriT* regions of pMV158 and pMA67.](image)
3.8. TetL protein

Currently in the NCBI databases, there are tetL genes from ten different natural plasmids. These, as well as one chromosomal tetL from E. faecium, are predicted to produce TetL proteins with 98% or higher amino acid sequence identity. These are seen in the neighbor-joining phylogenetic tree of Fig. 8 as clade “A”, supported by a very strong bootstrap value. By contrast, the predicted pMA67 TetL was found to be unique among plasmid-encoded TetL proteins in that it has only 86–88% amino acid sequence identity to the others. As seen from the phylogram, pMA67 forms its own clade with strong support. Thus, the pMA67 tetL appears to have arisen from a relatively ancient divergence from the other plasmid encoded tetL genes. Two chromosomal tetL genes from different strains of Bacillus subtilis produce almost identical TetL proteins, which form another clade. The Fig. 8 results were further supported by a single most parsimonious tree with 87 parsimony-informative characters with a length of 299 steps, consistency index (CI) of 0.977, and a retention index (RI) of 0.946, which showed the same topology as the neighbor-joining tree with strong bootstrap values (data not shown).

In order to show that it is in fact the tetL gene of pMA67, which was identified only using genetic databases, that is responsible for conferring tetracycline resistance to the pMA67-containing bacteria in this study, we cloned in isolation the entire tetL gene of pMA67, as well as a mutated tetL (3’ truncated). The two tetL alleles were generated by PCR from the pMA67 regions indicated in Fig. 1 using primer pairs 1776-F/3664-R and 1776-F/805-R. We ligated each of the PCR products to a plasmid vector, and transformed E. coli. All transformants with the full-length tetL were found to be tetracycline resistant, and all with the truncated tetL were found to be tetracycline sensitive, as expected. This confirms that it is the tetL gene of pMA67 that confers the tetracycline-resistance phenotype.

The tetL gene of pMA67 possesses translational attenuation sequences (GenBank nucleotide positions 1450–1602) upstream of the coding region, containing the IRs and leader peptide ORF, which are typical of inducible tet genes (Hoshino et al., 1985). These intact regulatory sequences further support that the pMA67 tetL gene is expressed in P. larvae.

Although there have been plasmids found in P. larvae before (Benada et al., 1988; Bodorová-Urgosíková et al., 1992; Drobníková et al., 1994; Neuendorf et al., 2004) no phenotype was associated with any of them, no genes were identified, and no sequencing was done on them. Therefore, pMA67 is the first well characterized plasmid from P. larvae. Plasmid pMA67 also provides the first record of any tetracycline-resistance gene found in the entire Paenibacillus genus (Chopra and Roberts, 2001; Roberts, 2005). Since pMA67 is predicted to be mobilizable, then this plasmid, or a close relative, may be present in other bacteria sharing the same natural environment as P. larvae. We have, in fact, detected tetL by PCR in Paenibacillus alvei—another species which is closely associated with honey bees—so it may be that pMA67 or a closely related plasmid will be found in that species also.

Acknowledgments

We thank Anita Davelos Baines (University of Texas-Pan American, Edinburg, TX), Jay Evans
(USDA-ARS, Beltsville, MD), and Daniele Prov-enzano (University of Texas at Brownsville, Brownsville, TX) for their helpful critical reviews of the manuscript.

References


