CHAPTER III

Characterization of Plasmids from *Histophilus somni* and Construction of an *Escherichia coli–Histophilus somni* Shuttle Vector

[Published in part in *Plasmid*, volume 55**,** pages 227–234, under the title "Comparative analyses of two cryptic plasmids from *Haemophilus somnus* (*Histophilus somni*)"

ABSTRACT

Plasmid profiling has been used to gain insights into the genetic relatedness of different strains/species of bacteria. In this chapter, the characteristics of two *H. somni* plasmids have been described. The plasmid from a *H. somni* bovine abortion isolate has been designated pHS649 and has been shown to be closely related to a plasmid from an ovine isolate of *H. somni*. Analyses of the sequence of pHS649 identified a putative RepA protein that has 48% amino acid homology to the RepA protein of *Escherichia coli* plasmid pKL1. The plasmid from a *H. somni* bovine preputial isolate has been designated pHS129 and has been shown to be related to plasmids from members of the *Pasteurellaceae* and *Neisseriae*. Only pHS649 has been found to stably replicate in *E. coli* strains. Based on their sequence characteristics and similarity to plasmids with unobvious phenotypes, pHS649 and pHS129 have been categorized as cryptic plasmids unrelated to each other. Further characterization of these and other plasmids from members of the *Pasteurellaceae* should enhance understanding of plasmid biology within this family.

Introduction

As autonomously replicating extrachromosomal genetic elements, plasmids often confer several traits that could either enhance or reduce the fitness of their host species within a potentially competitive ecological niche (Lenski et al., 1994; Smith, 2001). Plasmids also constitute a simple machinery through which homologous and heterologous transfer of genetic material is accomplished both in the natural environment as well as in the laboratory (Smith, 2001; Sambrook and Russell, 2001). Conjugative plasmids can transfer genes between different species of bacteria and contribute to the evolutionary processes. Plasmids that do not replicate in a particular species of bacteria can be used as 'suicide vectors' in gene 'knock-out' mutation experiments. Because of their generally small size as well as comparatively fast rate of genetic variation and horizontal transfer, some plasmid-borne attributes can occasionally be used to gain insights into the evolution of related plasmids and their hosts (del Solar et al., 1998; Moreira, 2000).

Three fundamental categories of plasmid replication have been well characterized; replication by the theta type mechanism, replication by the rolling-circle mechanism, and replication by the strand displacement mechanism (del Solar et al., 1998). Several plasmids encode a replication protein that is essential for the synthesis and maintenance of new copies of the plasmid within their natural host. Although plasmid replication is not coupled with bacterial chromosome replication, several host proteins may be involved in the synthesis and partitioning of plasmids, besides maintaining their optimal copy number (Nordstrom and Dasgupta, 2006). Campbell et al. (1999) proposed that plasmids are viable in a given bacterial host only when their DNA sequence signatures (AA, AT, TT, CC, CG, and GG dinucleotide relative abundances) are sufficiently compatible with the host chromosomal DNA sequence signatures.

Plasmids from several species within the *Pasteurellaceae* have been identified and characterized with the aim of exploiting them as agents for genetic manipulation (Frey and MacInnes, 1995; Wright et al., 1997). Plasmid profiling has also been used to characterize field isolates of *H. somni* (Fussing and Wegener, 1993; Appuhamy et al., 1998). Since antibiotic resistance is relatively uncommon in *H. somni* (Aarestrup et al., 2004; Higgins et al., 1981), there have been few efforts to decipher and describe the complete nucleotide sequence of plasmids from these bacteria. The DNA sequences of two plasmids from *H. somni* have been available (GenBank accession numbers AF318175 and AF546882), but a systematic description of these and other *H. somni* plasmids in the context of their relationship to each other and to plasmids from other bacteria has also not been available. Furthermore, there have been no reports of attempts to test the replicability of *H. somni* plasmids in other bacteria or construct endogenous plasmid-based shuttle vectors that function in laboratory strains of *E. coli* and *H. somni*. In view of these facts and the importance of *H. somni* as a cattle pathogen, it is imperative to understand the plasmid biology within this genus and construct *E. coli-H. somni* shuttle vectors that can facilitate gene transfer.

In this study, the nucleotide sequences of plasmids from a clinical isolate and an asymptomatic carrier isolate of *H. somni* have been deciphered. The objectives were to complement the complete genome sequencing project as well as to characterize the plasmid-borne genes. The two *H. somni* strains were found to contain disparate circular plasmids that possibly originated by horizontal transfer from distinct sources during the evolution of *H. somni*. Evidence of genetic relationship between plasmids from two other *H. somni* strains has also been obtained. One of the two plasmids described in this study has been modified by inserting an antibiotic resistance gene and has been found to replicate in *E. coli* and *H. somni*.

Materials and Methods

Bacterial strains used in this study

Six pathogenic and five preputial isolates of *H. somni* were chosen for the study and are described in table 3.01.

Table 3.01: Bacterial strains used in this study

***H. somni* strains**	**Origin/Site of isolation**	**Source**	**Reference**
8025	Thrombotic meningo-encephalitis isolate	Dr. Lynette B. Corbeil UCSD Medical Center	Brown et al. (1972)
2336	Pneumonia isolate	Animal Dis. Diagnostic Lab. Washington State	Corbeil et al. (1985)
738	Respiratory challenge clonal isolate of strain 2336 from a calf	Animal Dis. Diagnostic Lab. Washington State	Inzana et al. (1992)
649	Isolated from the stomach contents of an aborted fetus	Animal Dis. Diagnostic Lab. Washington State	Widders et al. (1986)
93	Calf pneumonia/meningitis isolate	Vet. Inf. Dis. Org. Saskatchewan, Canada	Howard et al. (2000)
6519	Pneumonia isolate	Vet. Inf. Dis. Org. Saskatchewan, Canada	Unpublished
1P, 24P, 124P, 127P, 129Pt	Isolates from the prepuce of apparently healthy bulls	Animal Dis. Diagnostic Lab. Washington State	Corbeil et al. (1985)

Plasmid purification

H. somni was cultured on Columbia blood agar (Becton Dickinson, Sparks, MD) plates containing 5% sheep blood in a candle jar for 24-36 hours at 37^0C. The plate cultures were inoculated into 500 ml of Brain Heart Infusion (BHI) [Becton Dickinson, Sparks, MD] broth with 0.1% Trizma base (Sigma, St. Louis, MO) and 0.01% thiamine monophosphate (TMP) (Sigma, St. Louis, MO), and incubated for 24-36 hours at 37^0C, with shaking at 200 rotations per minute (rpm). Bacteria were pelleted by centrifugation and plasmids were purified using the standard alkaline lysis protocol (Sambrook and Russell, 2001). For large-scale plasmid preparation from 500 ml broth cultures, a Qiagen midiprep kit (Qiagen, Valencia, CA) was used.

Restriction enzyme digestion and electrophoretic analyses

All restriction enzymes were purchased from either New England Biolabs (NEB, Beverly, MA) or Promega (Promega Corporation, Madison, WI). Reaction mixtures were incubated at appropriate temperatures in a water bath for 2 hours. The purified plasmids (5 μl) and/or their restriction enzyme digested products (20 μl) were resuspended in 3 μl of 1x loading buffer [Buffer type III (6x): 0.25% bromophenol blue, 0.25% xylene cyanol FF and 30% glycerol in water] and were electrophoresed in a 1% agarose gel cast using 1x Tris Acetate EDTA (TAE) buffer and 0.5 μg/ml ethidium bromide.

Purification of restriction fragments and ligation

DNA bands were excised from the agarose gels using a clean scalpel and DNA was extracted using a gel purification kit (Qiagen, Valencia, CA). The gel-purified products were resuspended in 30 μl of sterile ddH_2O and cloned into pBluescript (Strategene, La Jolla, CA) by standard protocols (Sambrook and Russell, 2001). A 20 μl ligation reaction [containing 10 μl of the insert, 7 μl of the pBluescript vector, 2 μl of ligation buffer and 1 μl of T-4 DNA ligase (NEB)] was incubated overnight at room temperature on the bench.

Bacterial transformation

Chemically competent *E. coli* DH5α cells (Invitrogen, Carlsbad, CA) were transformed with 5 μl of the ligation reaction mixture following the protocol supplied by the manufacturer. The transformed cells were resuspended in 950 μl of Luria Bertani (LB) [Becton Dickinson, Sparks, MD] broth and were recovered at 37^0C with shaking at 200 rpm for one hour. The recovered cells were plated (100 μl) on LB agar plates containing 100 μg/ml ampicillin (Sigma,

St. Louis, MO) and 30 µl of 20 µg/ml X-gal (Sigma, St. Louis, MO). White colonies were selected, cultured, and screened (with few blue colonies as negative controls) using a rapid screening procedure [a loopful of cells resuspended in 100 µl sterile ddH_2O, add 100 µl saturated phenol, vortex gently, centrifuge at 15000 rpm for 3 min in a microcentirfuge, remove 20 µl of the aqueous supernatant for electrophoretic analysis]. Putative recombinants showing a phase shift upon agarose gel electrophoresis were cultured in 10 ml of LB broth with 100 µg/ml ampicillin and plasmid minipreparation was done using the Qiagen miniprep kit (Qiagen, Valencia, CA). The recombinant plasmids were confirmed by restriction enzyme digestion and DNA sequencing.

Sequencing, assembly, and blast analyses

The plasmid inserts in the vector (pBluescript) were sequenced using universal primers M13 and T7 (Invitrogen, Carlsbad, CA) to obtain at least 4x coverage of the plasmid sequence. Additional primer walking was done when necessary. Primers (listed in table 3.02) were purchased from Integrated DNA Technologies, Coralville, IA. Sequencing reactions [each 15 µl sequencing reaction consisted of 4 µl BigDye® Terminator chemistry (Applied Biosystems**,** Foster City, CA), 3.5 pmol of M13 or T7 primer, 10 ng plasmid template, and sterile ddH_2O q.s.] were set up on a Hybaid PCR Express thermocycler (Hybaid, Middlesex, UK) following the standard cycle sequencing protocol (30 cycles of 94^0C for 30sec-50^0C for 15sec-60^0C for 4 minutes). All sequencing reactions were performed at the Virginia Bioinformatics Institute. Chromatogram sequence results were analyzed using either the BioEdit (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) or Sequencher (Gene Codes Corporation, Ann Arbor, MI) program. Vector sequences from each of the sequence reads were identified using the

"vector contamination screening tool (VecScreen)" [available at National Center for Biotechnology Information (NCBI) web site]. After removing the vector contamination, the sequences were assembled into a linear contig using tools on the BioEdit and Sequencher programs. The assembled plasmid sequences were analyzed done using various blast algorithms available on the NCBI web site and plasmid maps were drawn using BioEdit.

Table 3.02: Primers used in this study

Primer Name	Primer Sequence
M13 Rev	5′-GGAAACAGCTATGACCATG-3′
T7	5′-TAATACGACTCACTATAGGG-3′
649F	5′-TTACTGACAAGAACGGTGAGAT-3′
129LT7ext	5′-GGCTTGCGAGATTTAAGGTGC-3′
129LM13ext	5′-TATGAGGATAGACATGAGACAG-3′
129extST7	5′-GCCGATAGGCTCTATATTTTTGTT-3′
129extSM13	5′-CGTGAATTTTACGAGTACCTTCG-3′

Sequence alignment and secondary structure prediction

Plasmid replication protein sequences of various bacteria were obtained from the NCBI database based on tblastx analyses. Protein and DNA sequences were aligned using the ClustalW multiple sequence alignment program (available at the European Bioinformatics Institute web site http://www.ebi.ac.uk/clustalw/). For protein sequence analyses, the ClustalW sequence submission form was used with default settings except the 'output format' set to 'pir'. The ClustalW 'pir' output was further analyzed using the program BOXSHADE (http://www.ch. embnet.org/software/BOX_form.html) with the following settings; output format: RTF_new, fraction of sequences that must agree for shading: 0.1, input sequence format: other. For secondary structure prediction, Mfold program (available at the Pasteur Institute web site http://bioweb.pasteur.fr/seqanal/interfaces/mfold-simple.html) was used with default settings.

Results

Plasmid profiles of *H. somni* isolates

Each of the five preputial isolates was found to carry at least one plasmid (Fig. 3.01). Of the six disease isolates that were examined, only strain 649 was found to carry a single predominant plasmid (data not shown). The plasmids found in strain 649 (designated as pHS649, ~1.3 kb) and strain 129Pt (designated as pHS129, ~5 kb) were selected for further analyses since these strains represent a pathogenic and asymptomatic carrier isolate from the urogenital tract, respectively.

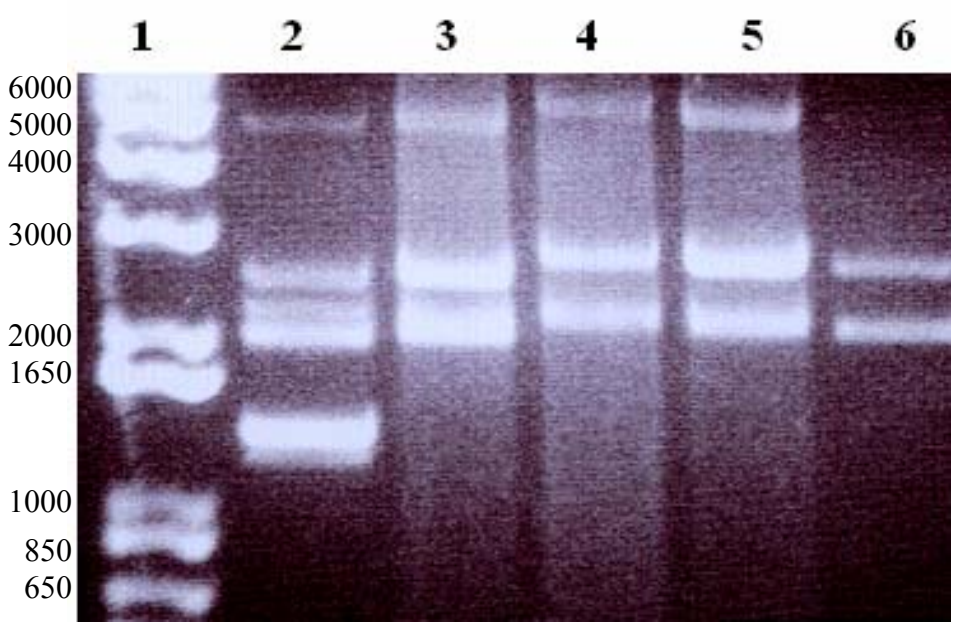


Fig. 3.01: Plasmid profiles of *H. somni* preputial strains

Lane 1; 1kb plus DNA ladder (only bands from 650 bp to 6,000 bp are shown), Lane 2; *H. somni* strain 1P plasmid extract, Lane 3; *H. somni* strain 24P plasmid extract, Lane 4; *H. somni* strain 124P plasmid extract, Lane 5; *H. somni* strain 127P plasmid extract, Lane 6; *H. somni* strain 129Pt plasmid extract. Lanes 3-6 probably contain different conformations of a single plasmid within each strain and lane 2 probably contains more than one plasmid each with different conformations.

Plasmid pHS649

The assembled length of pHS649 was 1,347 bp (GenBank accession number **AY288940**). The circular map of pHS649 is depicted in fig. 3.02. The size as well as the GC content (~43%) of pHS649 are similar to those of plasmids from bovine and ovine isolates of *H. somni*. However, the GC content of pHS649 is higher than the GC content of the chromosomes of *H. somni* strains 2336 and 129Pt (~37%). A blastn of pHS649 on the NCBI database revealed homology (Fig. 3.02) to plasmid p9L of *H. somni* (GenBank accession number AF546882) in six regions on the plus strand (90% identity within 225 nucleotides, 87% identity within 116 nucleotides, 91% identity within 78 nucleotides, 100% identity within 47 nucleotides, 86% identity within 76 nucleotides, and 100% identity within 32 nucleotides). pHS649 was also found to have homology (Figs. 3.02 and 3.03) to a short sequence of 67 nucleotides on the minus strand (98% identity; e-value: 4e-25) of plasmid p57/98 (GenBank accession number AF318175) from a Hungarian isolate of *H. somni* (Izadpanah et al., 2001). No homology was observed between pHS649 and the genome sequences from any of the members of the *Pasteurellaceae*.

Two open reading frames (ORFs) were identified within the sequence of pHS649: a 528 bp ORF (designated ORF1) encoding a putative RepA protein with homology to Rep proteins from several plasmids, and a 165 bp ORF (designated ORF2) whose translated protein product had no homologs within the sequence database (Fig. 3.02 and Table 3.03). The closest homologs of pHS649 RepA were from p9L (*H. somni*; 69% identity and 84% similarity within 163 amino acids), pKL1 (*E. coli*; 48% identity and 70% similarity within 140 amino acids), pONE429 (*Selenomonas ruminantium*; 26% identity and 45% similarity within 135 amino acids), pDTG1 (*Pseudomonas putida*; 31% identity and 43% similarity within 128 amino acids),

pBMYdx (*Bacillus mycoides*; 27% identity and 42% similarity within 118 amino acids), and pAM10.6 (*Pseudomonas fluorescens*; 22% identity and 46% similarity within 123 amino acids).

Figure 3.04 shows pHS649 RepA-related sequences from six different plasmids analyzed using ClustalW and BOXSHADE. The alignment revealed that the amino and carboxy terminal regions were less conserved, but a segment of ~100 amino acids between the amino and carboxy terminals had several conserved residues, notably alanine, leucine, tyrosine, and tryptophan. A putative leucine-zipper-like motif, which is characteristic to most proteins that interact with DNA, was also identified using the alignment (Fig. 3.04). The predicted molecular weight of pHS649 RepA is 20.11 kDa.

In some plasmids that replicate by the RC mode, the double strand origin (DSO) consists of a nick site flanked by Rep-binding inverted repeats (IRs) upstream of the *rep* genes. Some DSOs also contain two to three direct repeats (DRs) (Khan, 1997). The single strand origins (SSO) of plasmids that replicate by the RC mode, usually located upstream of the DSO, have complex secondary structures due to extensive IR sequences (Khan, 1997). In order to identify putative DSO and SSO of pHS649, we compared the sequences upstream of *repA* ORFs in pHS649 and p9L, as these plasmids show similar genetic organization and are from closely related bacteria (Table 3.04). A 13 bp IR flanking a 9 bp spacer region was identified 48 bp upstream, and four 6-10 bp perfect DRs were identified 95 bp upstream of the *repA* ORF in pHS649 (Fig. 3.05). This 128 bp region may constitute the DSO of pHS649. The region between this putative DSO and the *repA* stop codon, which also includes ORF 2, consists of at least seven 6-15 bp perfect IRs (Fig. 3.05) that potentially form convoluted secondary 'stem-loop' like structures (Fig. 3.06). Based on these features, it was predicted that this 583 bp region constitutes the SSO of pHS649.

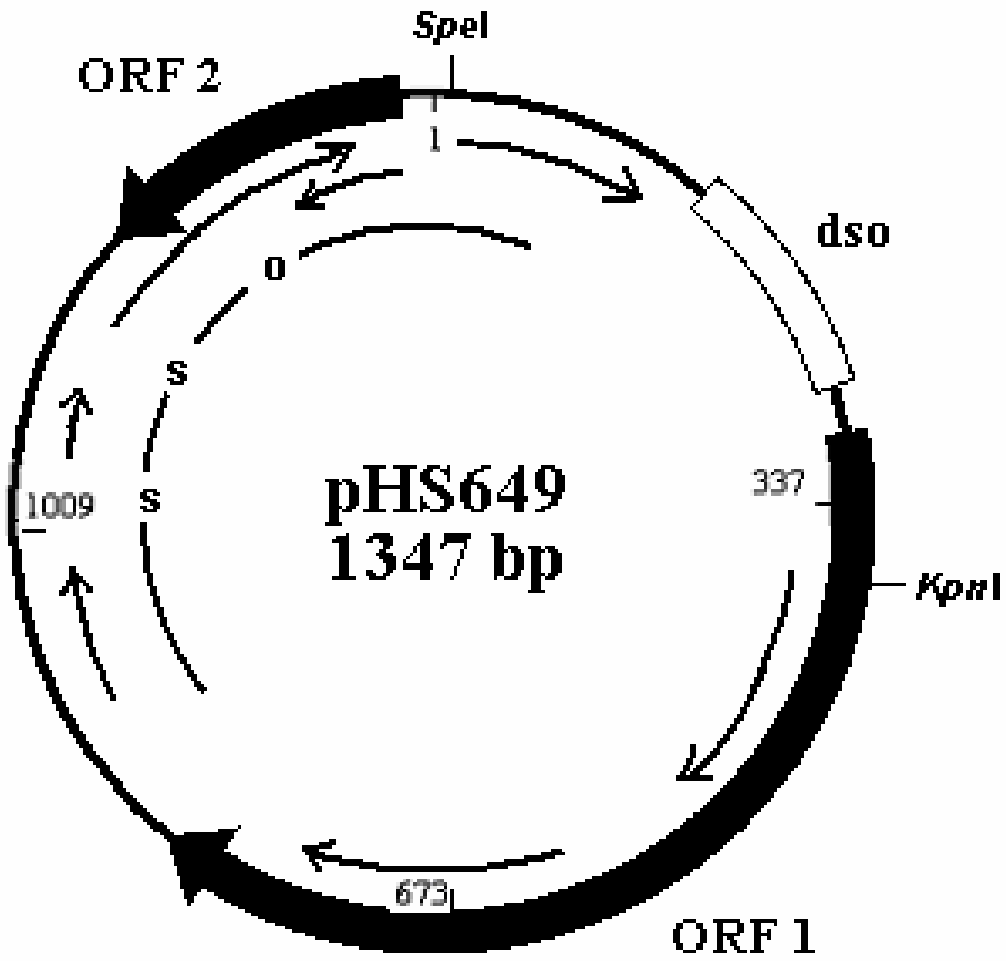


Fig. 3.02: Circular map of pHS649 **(AY288940)**

The figure shows putative ORFs, SSO, and DSO regions. ORF 1 encodes RepA. Unique restriction sites that were used for cloning and sequencing this plasmid (KpnI and SpeI) are also shown. Regions of homology to p9L (six thin clockwise arrows) and p57/98 (single thin anticlockwise arrow) are also indicated.

Fig. 3.03: ClustalW-BOXSHADE alignment of plasmids pHS649, p9L, and p57/98 showing a region of homology

Top Row; plasmid pHS649, sequence shown is from 1,260 bp to 1,347 bp

Center Row; plasmid p9L, sequence shown is from 189 bp to 276 bp

Bottom Row; plasmid p57/98, sequence shown is from 309 bp to 396 bp

Gray asterisks indicate the presence of two identical nucleotides among the three sequences.

Black asterisks indicate the presence of three identical nucleotides among the three sequences.

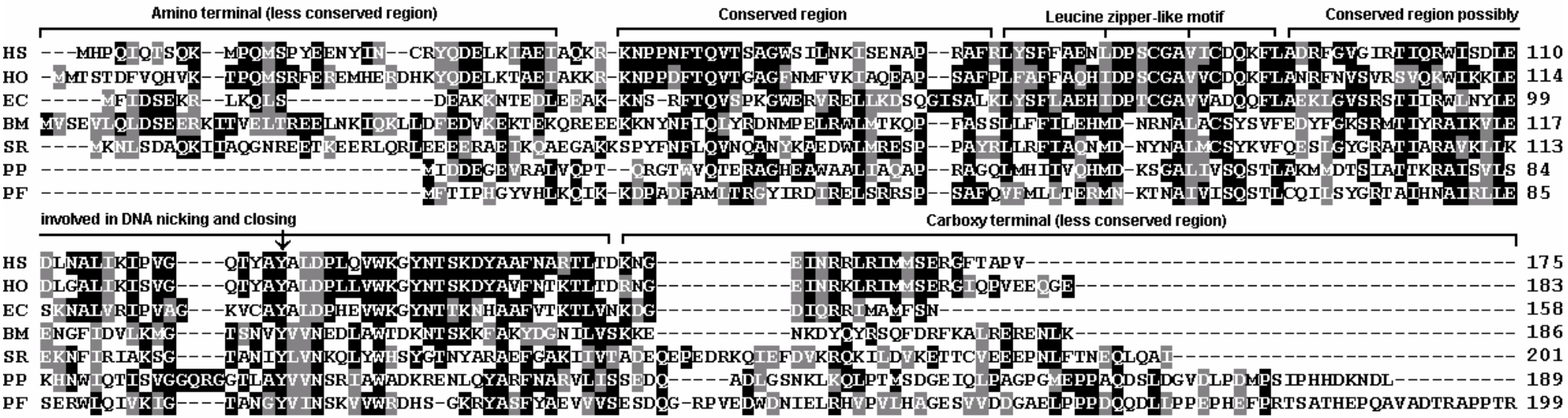


Fig. 3.04: ClustalW-BOXSHADE multiple sequence alignment of RepA homologs from seven bacterial plasmids

Plasmid names and GenBank accession numbers of protein sequences are given in parenthesis. HS; *H. somni* (pHS649: **AAP44497**), HO; *H. somni* (p9L: **NP_715625**), EC; *E. coli* (pKL1: **NP_053155**), BM; *Bacillus mycoides* (pBMYdx: **NP_981975**), SR; *Selenomonas ruminantium* (pONE429: **NP_862700**), PP; *Pseudomonas putida* (pDTG1: NP_863059), PF; *Pseudomonas fluorescens* (pAM10.6: **AAG23805**). Homologous regions are box-shaded dark (identical amino acid residues) and gray (conserved amino acid substitutions). The less conserved amino and carboxy terminal regions are shown. A putative leucine zipper-like motif is marked. A conserved region putatively involved in DNA nicking and closing is also marked. An arrow points to the conserved tyrosine residue (Y) that may participate in DNA nicking. Numbers denote amino acid positions.

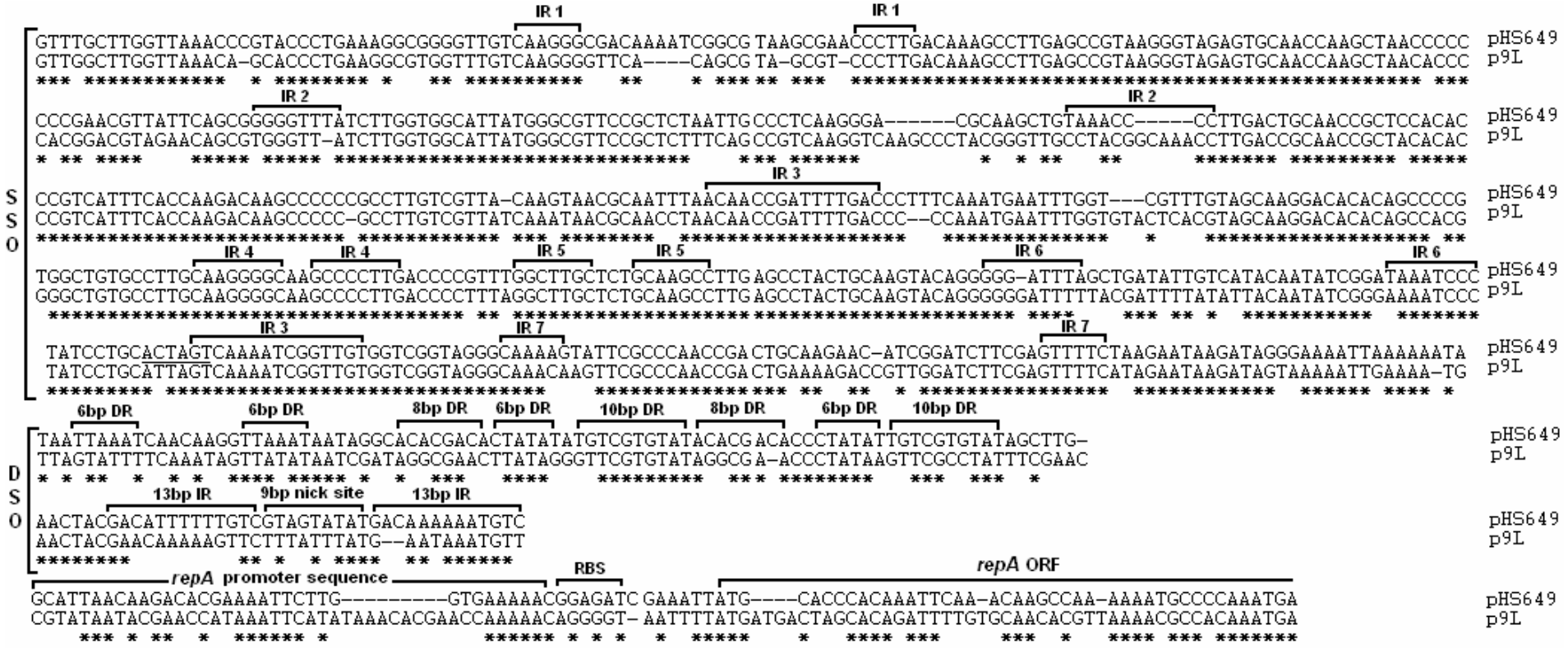


Fig. 3.05: ClustalW alignment of DNA sequence upstream of pHS649 *repA* with the DNA sequence upstream of p9L *repA*

Putative SSO (consisting of 7 pairs of IRs) and DSO (consisting of 4 pairs of DRs and one 13 bp IR pair flanking a 9 bp nick site) regions of pHS649 are marked. SpeI restriction site (ACTAGT) of pHS649 is underlined. A candidate *repA* promoter sequence and a potential ribosome binding site (RBS) from pHS649 are shown. Asterisks denote identical nucleotides.

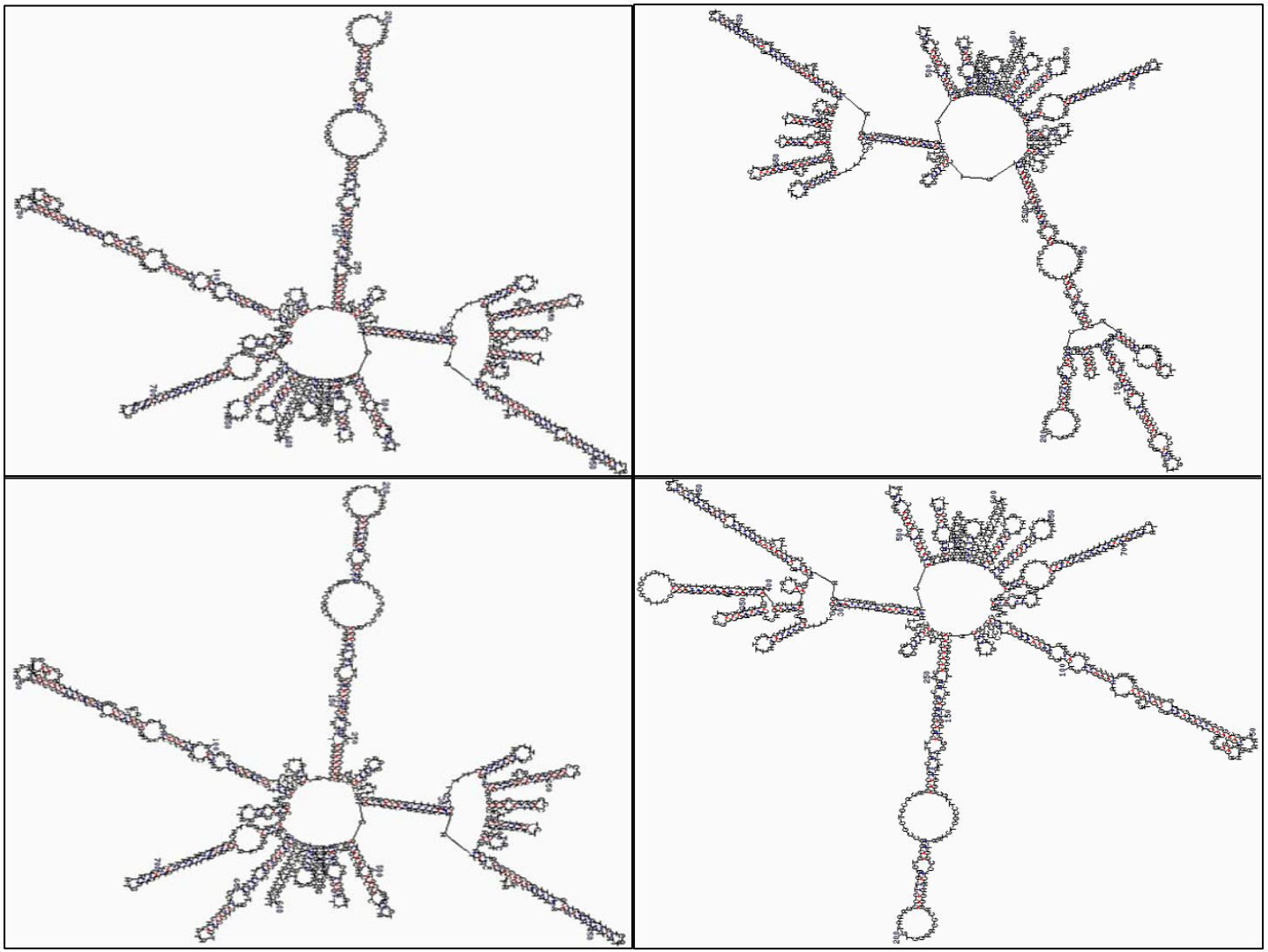
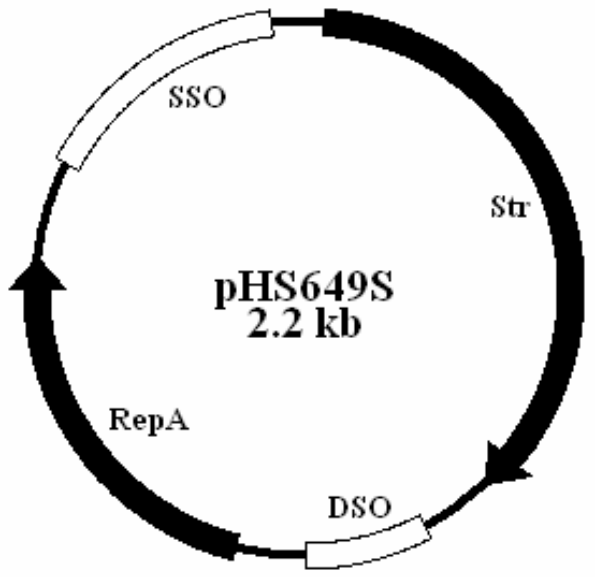


Fig. 3.06: Four predicted secondary structures (‘stem-loop’ like structures) formed by pHS649 DNA sequence

A derivative of pHS649 replicates in *E. coli* strains

The streptomycin gene from plasmid pLS88 (Dixon et al., 1994) was amplified by PCR and cloned into pCR2.1-TOPO (Invitrogen) to create plasmid pCR2.1TOPOS. The streptomycin gene from pCR2.1-TOPOS was excised by SpeI-XbaI digestion and was subcloned into the SpeI site of pHS649 to create plasmid pHS649S (Fig. 3.07). This plasmid was introduced into chemically competent cells of *E. coli* strains DH5α and DH10B (Invitrogen). Plasmid pHS649 was also found to be maintained in high copy numbers (Fig. 3.08) in these *E. coli* strains during subsequent subcultures without streptomycin selection on LB agar plates.



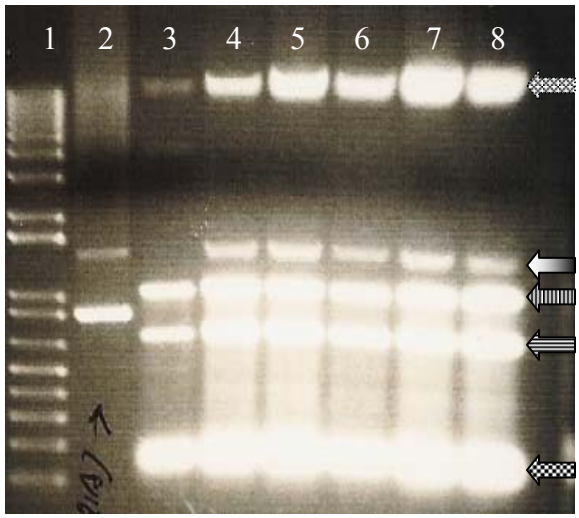


Fig. 3.07 (left): Circular map of pHS649S; Putative SSO and DSO as well as replication protein (RepA) and streptomycin (Str) coding regions are shown.

Fig. 3.08 (right): Rapid screening of *E. coli*; Lane 1; 1kb plus DNA ladder, Lane 2; pHS649 purified from *H. somni* strain 649 (contains both supercoiled and linear forms), Lane 3; *E. coli* DH5α negative control, Lanes 4-8; Five different clones of *E. coli* DH5α containing pHS649S.

Arrows: genomic DNA, pHS649S DNA, 23S rRNA, 16S rRNA, 5S rRNA.

Plasmid pHS649S contains single strand DNA intermediates

One of the characteristic features of plasmids replicating by the rolling circle mechanism is the presence of single strand DNA intermediates in cell and plasmid extracts (Khan, 1997). Single strand forms of DNA have also been observed during gel electrophoresis of pHS649S purified from *E. coli* DH5α. The single strand nature of this DNA was confirmed by digesting the pHS649S substrate with three different Type II restriction enzymes that are known to cleave double stranded DNA and not single stranded DNA (Fig. 3.09).

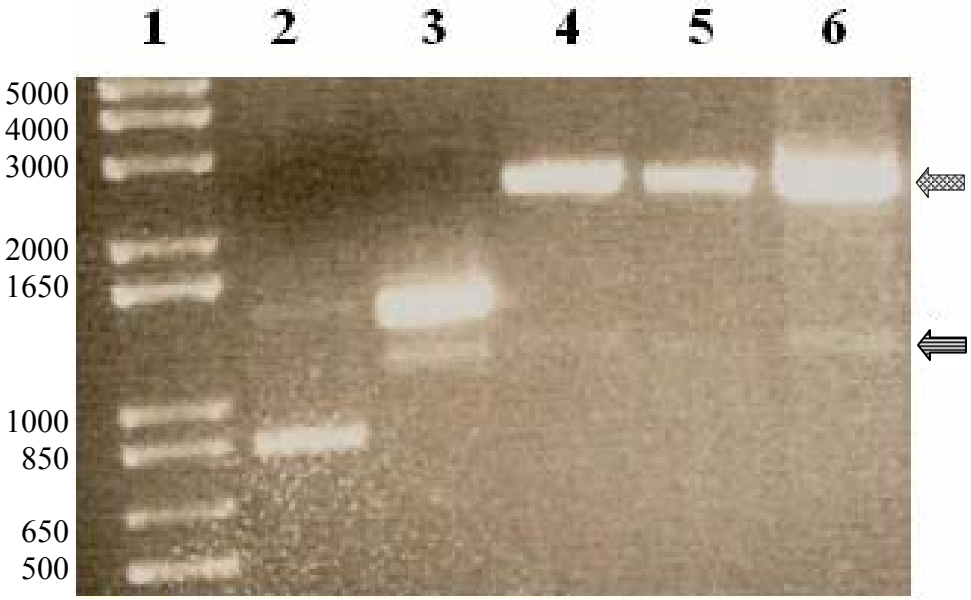
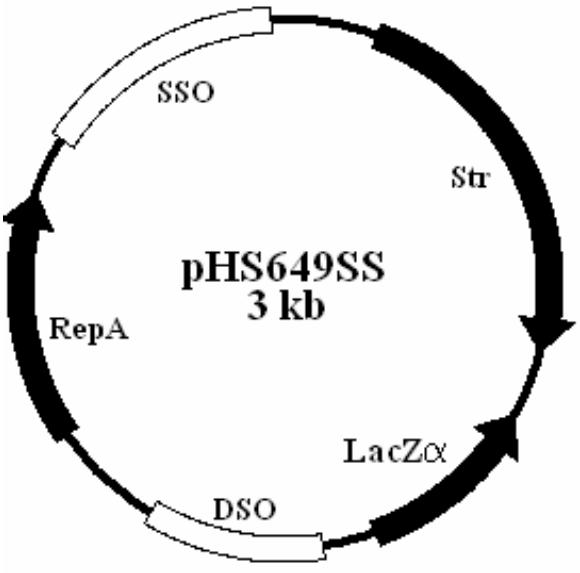


Fig. 3.09: Demonstration of single strand DNA intermediates in pHS649S

Lane 1; 1kb plus DNA ladder (bands shown are from 500 bp to 5,000 bp), Lane 2; pHS649 purified from *H. somni* strain 649, contains both supercoiled and linear double stranded DNA, Lane 3; pHS649S purified from *E. coli* DH5α, contains both double stranded supercoiled and single stranded DNA, Lanes 4, 5, and 6; pHS649S digested with EcoRV, SpeI, and XhoI, respectively, contain linearized double stranded DNA (marked with ⇦) and uncut single stranded DNA (marked with ⇦).

Modification of pHS649S for alpha complementation of LacZ in *E. coli*

Since preliminary tests showed that pHS649S replicates in *E. coli* (Fig. 3.08) as well as in *H. somni* (data not shown), this plasmid therefore had the potential to be developed into an *E. coli-H. somni* shuttle vector. To enhance the number of restriction sites for cloning DNA into the vector and to facilitate blue-white selection of recombinants in *E. coli*, a 713 bp fragment containing the *lacZα* gene from pBluescript II SK (-) was excised by SspI-AflIII double digestion and was treated with DNA polymerase I (Klenow) to blunt the ends. The vector pHS649S was digested with EcoRV-NsiI-SpeI and the fragments were similarly blunted. The DNA molecules were ligated together to create plasmid pHS649SS (Fig. 3.10). This plasmid was subsequently introduced into chemically competent cells of *E. coli* strain DH5α and alpha complementation was confirmed after addition of X-gal (Fig. 3.11).



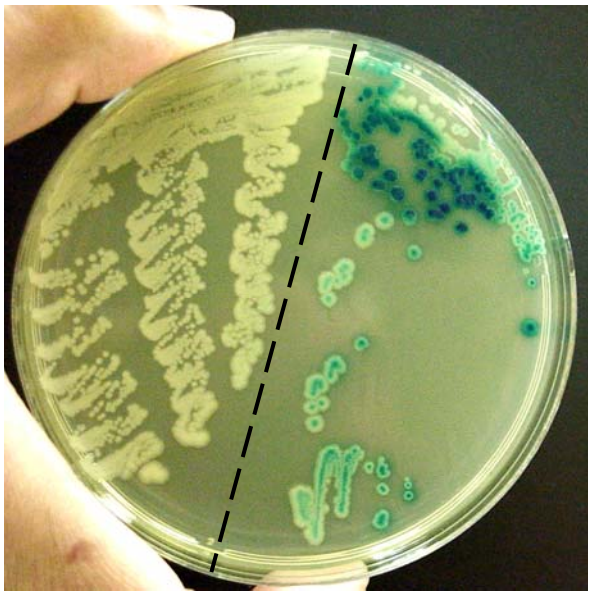


Fig. 3.10 (left): Circular map of pHS649SS; Putative SSO and DSO as well as replication protein (RepA), streptomycin (Str), and β-galactosidase α fragment (LacZα) coding regions are shown. Fig. 3.11 (right): LB agar plate containing *E. coli* DH5α transformed with pHS649SS; Only right half of the plate contains X-gal and confirms alpha complementation of LacZ by pHS649SS.

Plasmid pHS129

The assembled length of pHS129 was 5,178 bp (GenBank accession number **CP000019**). The GC content of this plasmid (~35%) is one of the lowest in its category, along with those of pHS-Rec (*Haemophilus parasuis*) and p250 (*Haemophilus paragallinarum*) (Table 3.04). Furthermore, the GC content of pHS129 is significantly less than the GC content of pHS649 (~43%). The GC content of pHS129 is also less than the GC content of the chromosome of its host; *H. somni* strain 129Pt (~37%).

This plasmid contains IRs of length 1,248 bp and 1,235 bp (77% identity), depicted in the circular map in fig. 3.12. Two putative *rep* ORFs, each encoding 323 amino acids, were identified within the plasmid sequence (*repB1* ATG is 276 bp downstream from the beginning of the 1,248 bp IR and *repB2* ATG is 263 bp downstream from the beginning of the 1,235 bp IR). Four additional ORFs were also identified within the sequence of pHS129 (Fig. 3.12 and Table 3.03). Two tandem repeat regions, a 4 bp 5'-$(TTTC)_8$-3' and a 22 bp 5'-$(AAAGGTGGTGT\ AATTTACGGTT)_4$-3', were identified within the 1,248 bp IR. Two tandem repeat regions, a 4 bp 5'-$(TTTC)_8$-3' and a 22 bp 5'-$(AAAGGTGCTATTTTTTGGTGTT)_4$-3', were also identified within the 1,235 bp IR. These characteristic, iteron-like direct repeat regions are thought to constitute putative origins of replication (*ori*) within pHS129 (Fig. 3.13). AT- and GC-rich regions were identified upstream of the 4 bp and 22 bp repeats. Based on sequence comparison to other iteron-containing plasmids, a putative 13 bp IHF binding site (CAAGCCTTGATAT) and a 9 bp *dnaA* box [TTA(C/T)(C/A)(A/G)C(T/A)(T/C)] were identified up- and down-stream of the 22 bp repeats, respectively. Sequences with partial homology to the putative iterons, which may interact with the Rep proteins to regulate *rep* transcription, were identified close to the presumptive promoter region of *repB1* (Fig. 3.13).

The two RepB proteins (predicted molecular weight 37.8 kDa each) putatively involved in plasmid replication are 87% similar (73% identity within 323 amino acids) to each other with more conserved residues being present on the C terminus of the predicted peptides, but are unrelated to the RepA protein that was identified in pHS649. Within the *Pasteurellaceae*, RepB1 has homology to RepB proteins from pJR2 (*Pasteurella multocida*; 39% identity and 62% similarity within 248 amino acids), pMVSCS1 (*Mannheimia varigena*; 38% identity and 62% similarity within 248 amino acids), p250 (*Haemophilus paragallinarum*; 35% identity and 62% similarity within 248 amino acids), pJR1 (*Pasteurella multocida*; 33% identity and 56% similarity within 312 amino acids), and pHS-Rec (*Haemophilus parasuis*; 29% identity and 56% similarity within 222 amino acids). Likewise, RepB2 has homology to RepB proteins from pMVSCS1 (*Mannheimia varigena*; 35% identity and 62% similarity within 250 amino acids), pJR2 (*Pasteurella multocida*; 34% identity and 62% similarity within 250 amino acids), p250 (*Haemophilus paragallinarum*; 37% identity and 58% similarity within 248 amino acids), pJR1 (*Pasteurella multocida*; 33% identity and 52% similarity within 311 amino acids), and pHS-Rec (*Haemophilus parasuis*; 29% identity and 57% similarity within 222 amino acids).

Figure 3.14 shows the ClustalW alignment of pHS129 RepB1 and RepB2 with the RepB homologs of plasmids from eight other species of bacteria. The alignment reveals that RepB is more highly conserved between this group of plasmids than RepA was among the plasmids shown in fig. 3.04. Several conserved amino acid residues that may form a domain and be involved in binding the *ori* to initiate a replication event are evident in the protein sequence alignment. Motifs that are characteristic to the replication initiation proteins of some plasmids, the leucine zipper hydrophobic motif and the helix-turn-helix motif (del Solar et al., 1998), were putatively identified within the sequences of both RepB1 and RepB2 (Fig. 3.14).

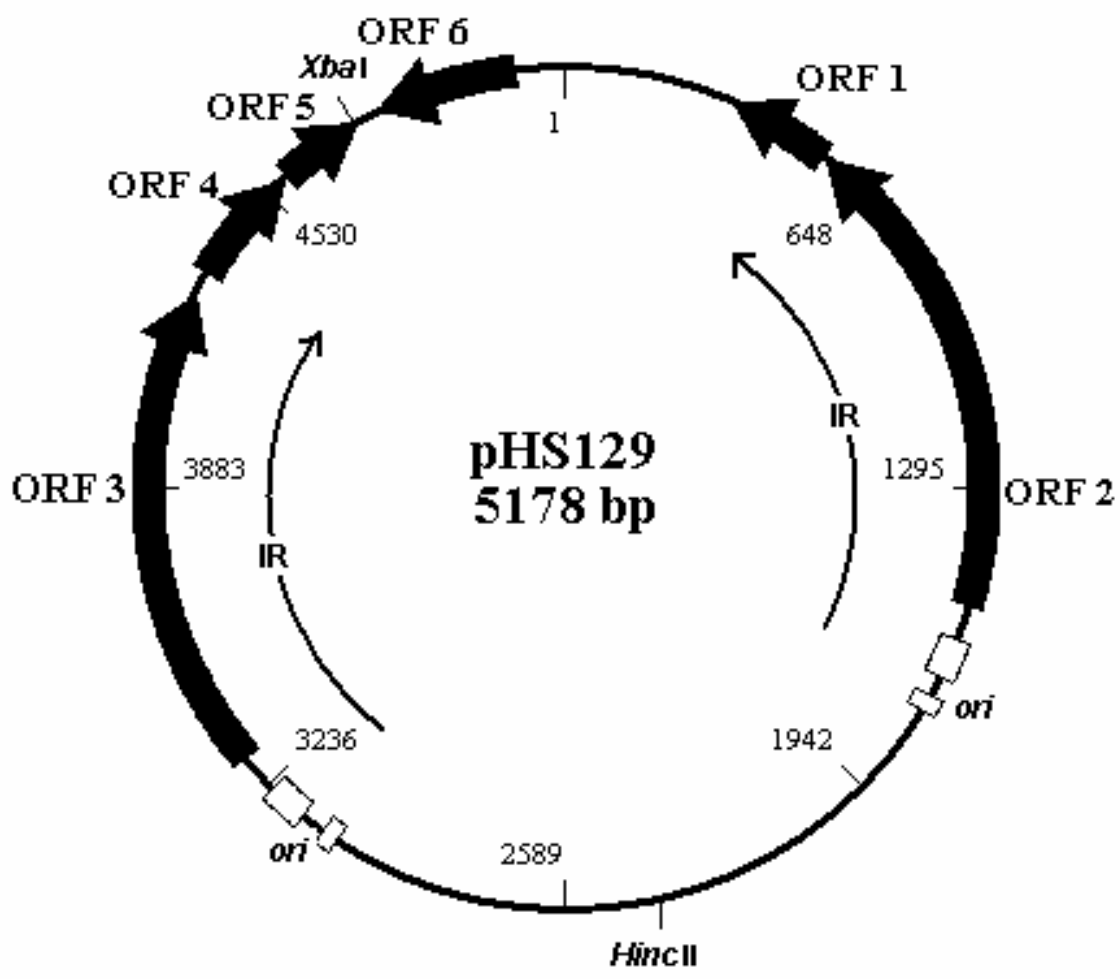


Fig. 3.12: Circular map of pHS129 **(CP000019)**

The figure shows putative ORFs and *ori* regions. ORF 2 encodes RepB1 and ORF 3 encodes RepB2. The rectangles represent iteron-like direct repeats. The IRs of this plasmid (indicated by thin arrows) begin close to the putative origins of replication (*ori*) and end with the stop codons of *repB1* and *repB2*. Unique restriction sites that were used for cloning and sequencing this plasmid (HincII and XbaI) are also marked.

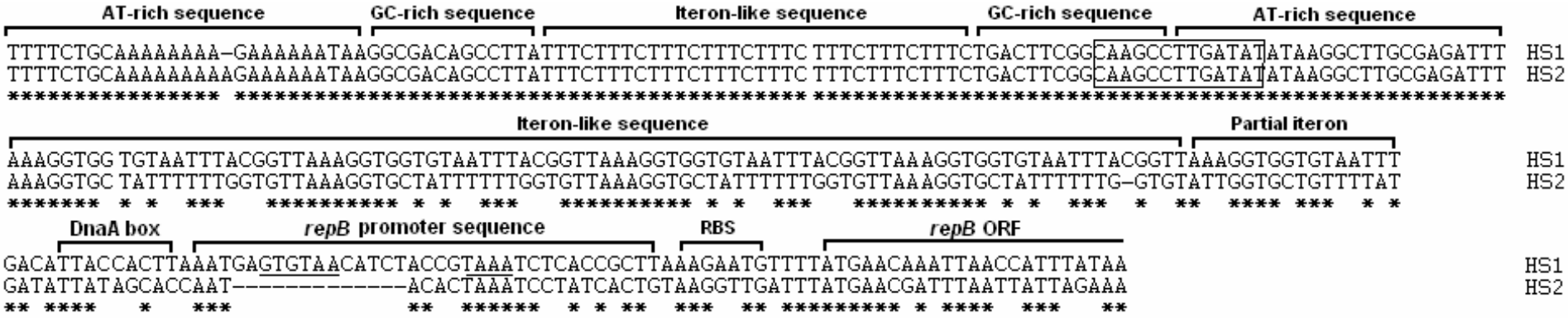


Fig. 3.13: ClustalW alignment of DNA sequences upstream of *repB1* (HS1) and *repB2* (HS2) ORFs of pHS129 (**CP000019**)

AT and GC rich sequences, iteron-like sequences (4 bp and 22 bp repeats), and a putative 9 bp DnaA box are marked. A putative 13 bp integration host factor-binding site is boxed. A candidate *repB* promoter sequence and a potential RBS are shown. Sequences with partial homology to the iterons (GTGTAA and TAAA) are underlined within the candidate *repB* promoter sequence. Asterisks denote identical nucleotides.

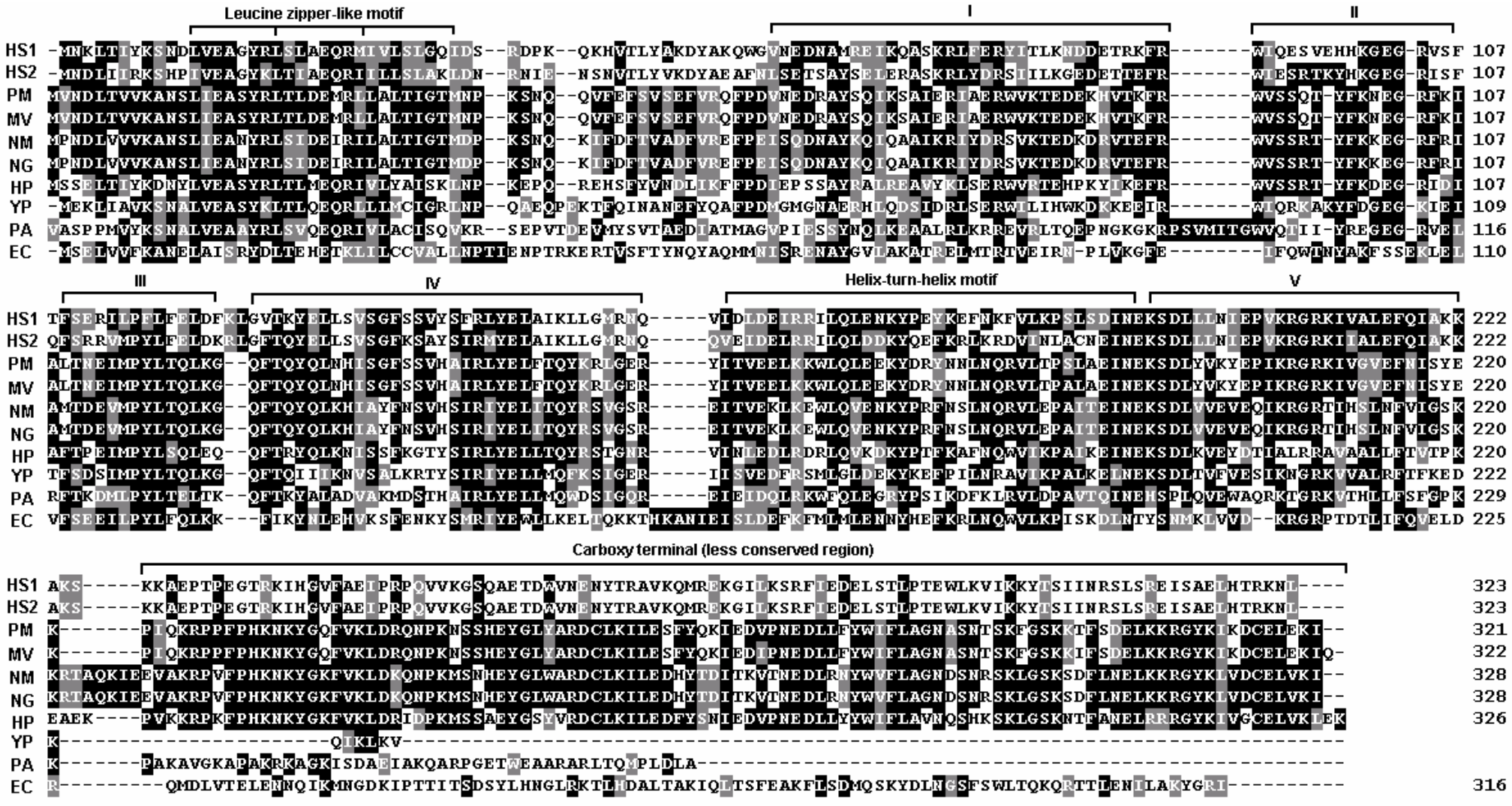


Fig. 3.14: ClustalW-BOXSHADE multiple sequence alignment of RepB sequences (detailed legend appears in the next page)

Fig. 3.14: ClustalW-BOXSHADE multiple sequence alignment of RepB1 and RepB2 of pHS129 with RepB homologs from plasmids of eight other bacterial species. Plasmid names and the GenBank accession numbers of protein sequences are given in parenthesis. HS1; *H. somni* RepB1 (pHS129: **AAU25833**), HS2; *H. somni* RepB2 (pHS129: **AAU25832**), PM; *Pasteurella multocida* (pJR2: **NP_848174**), MV; *Mannheimia varigena* (pMVSCS1: **NP_573540**), NM; *Neisseria meningitidis* (pAB6: **AAD31794**), NG; *Neisseria gonorrhoeae* (pFA3: **A35257**), HP; *Haemophilus paragallinarum* (p250: **NP_943193**), YP; *Yersinia pestis* (pYC: **NP_053154**), PA; *Pseudomonas aeruginosa* (pRO1614: **JC2565**), EC; *E. coli* (pSC101: **A43703**). Homologous regions are box-shaded dark (identical amino acid residues) and gray (conserved amino acid substitutions). Putative leucine zipper and helix-turn-helix motifs are marked. The less conserved carboxy terminal region is shown. I, II-III, IV, and V show conserved regions possibly involved in domain formation. Numbers denote amino acid positions.

Derivatives of pHS129 do not replicate in *E. coli* strains

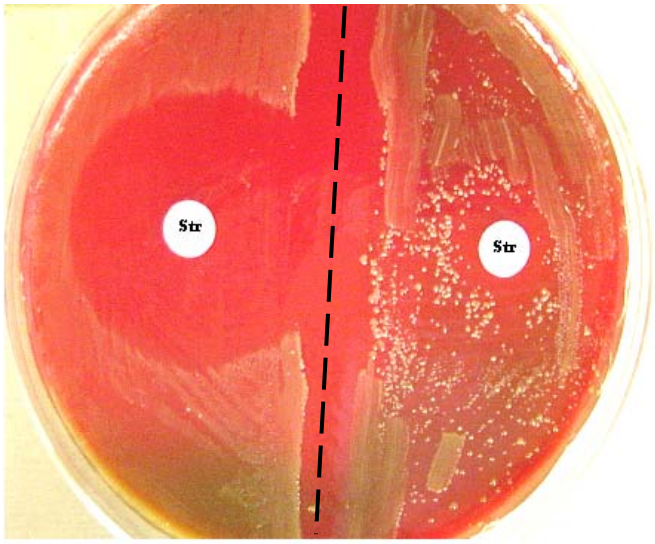
The streptomycin gene from pCR2.1-TOPOS was excised by SpeI-XbaI digestion and was ligated into the XbaI site of pHS129 by standard procedures. However, no streptomycin resistant colonies of *E. coli* DH5α or DH10B (Invitrogen) were obtained upon transformation with the ligated DNA molecules. This indicated that pHS129 may not function in *E. coli*, possibly due to host-plasmid incompatibility.

To test this hypothesis further, a Tn903 transposon insertion transposome kit (Epicentre, Madison, WI), which randomly transposes a selectable marker into a given piece of DNA, was used to insert the kanamycin resistance gene cassette into pHS129. Attempts to transform *E. coli* DH5α and DH10B with the transposome-treated pHS129 were unsuccessful. However, these *E. coli* strains were successfully transformed and became kanamycin resistant when the control plasmid pUC19 similarly treated with the transposome was used (data not shown). These results show that plasmid pHS129 with its own origin of replication may not be suitable for construction of *E. coli-H. somni* shuttle vectors.

Plasmid pHS129 is not incompatible with shuttle vector pHS649SS

H. somni strain 129Pt was transformed with a very high efficiency when plasmid pHS649SS was electroporated using procedures described by Sanders et al. (1997), confirming that the plasmid can replicate and function as a shuttle vector in *H. somni* (Fig. 3.15).

It was also found that the two disparate plasmids, pHS129 and pHS649SS, are compatible with each other and co-exist in high copy numbers within *H. somni* strain 129Pt with or without streptomycin selection during subsequent subcultures (Fig. 3.16).



Left side of the plate; streptomycin (Str) antibiotic disc inhibits the growth of *H. somni* strain 129Pt.

Right side of the plate; streptomycin (Str) antibiotic disc does not inhibit the growth of *H. somni* strain 129Pt transformed with pHS649SS.

Fig. 3.15: Transformation of *H. somni* strain 129Pt with pHS649SS

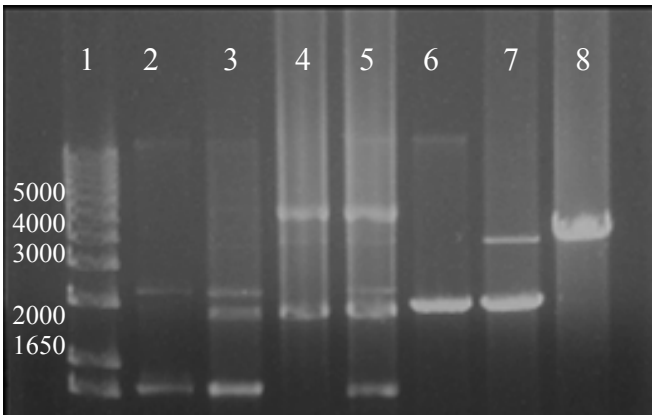


Fig. 3.16: Plasmids pHS129 and pHS649SS co-exist

Lane 1; 1 kb plus DNA ladder (only bands from 1,650 bp to 5,000 bp are marked), Lane 2; pHS649SS purified from *E. coli* DH5α, Lane 3; Plasmids purified from *H. somni* strain 129Pt transformed with pHS649SS, Lane 4; pHS129 purified from *H. somni* strain 129Pt, Lane 5; Plasmids pHS129 and pHS649SS combined for electrophoresis, Lane 6; Same as in lane 2, digested with XbaI, Lane 7; Same in lane 3, digested with XbaI, Lane 8; Same as in lane 4, digested with XbaI .

Table 3.03: Putative coding sequences of plasmids pHS649 and pHS129

Plasmid name	**Open reading frame (ORF)**	**Start**	**Stop**	**Frame**	**Number of amino acids**	**Description of the encoded protein**
pHS649	ORF 1	295	822	+1	175	Replication protein; possibly involved in plasmid replication
	ORF 2	1328	1164	-2	54	Hypothetical protein; has no homologs in the sequence database
pHS129	ORF 1	547	341	-3	68	Hypothetical protein; has homology to hypothetical protein Hsom02002512 from *H. somni* 129Pt
	ORF 2	1529	558	-2	323	Replication protein (RepB2); possibly involved in plasmid replication
	ORF 3	3309	4280	+3	323	Replication protein (RepB1); possibly involved in plasmid replication
	ORF 4	4327	4563	+1	78	Hypothetical protein; has homology to hypothetical proteins Hsom02002508 from *H. somni* 129Pt and Haso02001005 from *H. somni* 2336
	ORF 5	4574	4747	+2	107	Hypothetical protein; has homology to hypothetical protein Hsom02002509 from *H. somni* 129Pt.
	ORF 6	4798	5079	-1	93	Hypothetical protein; 67% similar to an hypothetical protein of plasmid pVT745 from *Actinobacillus actinomycetemcomitans*

Table 3.04: Comparison of plasmids from members of the *Pasteurellaceae* and *Neisseriae*

Plasmid name	**GenBank accession number**	**Species of origin**	**Size (bp)**	**Antibiotic resistance genes**	**Replication/ Mobilization protiens**	**G+C content (%)**
pHS649	AY288940	*H. somni*	1,347	None	*repA*	43.80
p9L	AF546882	*H. somni*	1,338	None	*repA*	43.80
p57/98	AF318175	*H. somni*	1,065	None	None	39.20
pHS129	CP000019	*H. somni*	5,184	None	*repB*	34.97
pHS-Rec	AY862436	*H. parasuis*	9,462	None	*repB*	32.78
p250	AY300023	*H. paragallinarum*	6,286	None	*repB*	33.50
pMVSCS1	AJ319822	*M. varigena*	5,621	$catA^R$, str^R, sul^R	*repB*	42.39
pJR1	AY232670	*P. multocida*	6,792	$catA^R$, sul^R, tet^R	*repB*	45.66
pJR2	AY232671	*P. multocida*	5,252	amp^R, str^R	*repB*	41.81
pMS260	AB109805	*Actinobacillus pleuropneumoniae*	8,124	str^R, sul^R	*repB, mobA*	60.48
pAB6	AF126482	*N. meningitidis*	5,597	pen^R	*repB, mobA*	39.65
pJD4	U20374	*N. gonorrhoeae*	7,426	pen^R	*repB, mobA*	38.34
pLS88	L23118	*Haemophilus ducreyi*	4,772	str^R, sul^R	None	50.50

Discussion

H. somni is a commensal of the bovine and ovine respiratory and reproductive tracts. Some strains of *H. somni* are resistant to host defense mechanisms and are capable of causing multisystemic diseases. This adaptation against host innate immunity by *H. somni* has been investigated at the biochemical and genetic levels. However, relatively little work was done previously on the plasmids of *H. somni*. There have been reports of *H. somni* plasmid profiles (Fussing and Wegener, 1993; Appuhamy et al., 1998) as well as deciphering the sequences of these plasmids (Izadpanah et al., 2001). Nevertheless, this is the first systematic effort to characterize the sequences of plasmids from *H. somni*, replication proteins encoded by these plasmids, and the utility of these plasmids as cloning vectors.

In this study, five bovine preputial isolates of *H. somni* that have been well studied by investigators in the USA (Corbeil et al., 1985; Howard et al., 2000) have been shown to carry a plasmid. Fussing and Wegener (1993) have also reported finding plasmids in several strains of *H. somni* isolated from the bovine genital tract or semen. These investigators found that only 15% of *H. somni* isolates from bovine pneumonia carry plasmids as opposed to 40% of urogenital isolates. Appuhamy et al. (1998) reported finding plasmids in 12 out of 19 (63%) isolates of *H. somni* from the ovine prepuce, semen, and vagina. In our study, only one of the *H. somni* pathogenic strains, abortion isolate 649, was found to carry a plasmid out of the 6 that were examined (16.6%). Fussing and Wegener (1993) have reported finding no plasmid in *H. somni* pneumonia strain 2336, which we confirmed. No plasmids were found in other pneumonia isolates in our collection. Although these and other strains might carry very low copy number plasmids, repeated purification efforts using methods developed for obtaining high copy number plasmids (Sambrook and Russell, 2001) failed to yield plasmids from any of these strains.

It is interesting to note that the size (1,347 bp) as well as the GC content (~43%) of pHS649 are similar to those of plasmid p9L. However, the GC content of pHS649 is greater than the GC content of the chromosomes of *H. somni* strains 2336 and 129Pt, indicating that it may have originated by horizontal transfer. Although conceptual translation assigned two ORFs to the sequence of pHS649, only *repA*, whose translated protein may be involved in plasmid replication (Burian et al., 1999), could be positively identified based on homology search. Analysis using blastn revealed that the sequence of pHS649 has significant homology within the noncoding regions of p9L and p57/98, but has no homology to any of the genome sequences from other members of the *Pasteurellaceae*. Furthermore, *H. somni* strain 649 was isolated from the stomach contents of an aborted bovine fetus (Widders et al., 1986), whereas the host strain of p9L was isolated from the ovine vagina (Walker et al., 1985). In view of this, it is noteworthy that the two plasmids have very high homology to each other. In addition, RepA is well conserved at both the nucleotide and amino acid levels (70% identical) within plasmids pHS649 and p9L. Based on these features, it is not unreasonable to suggest that these two plasmids, or their hosts, originated from a common ancestor. Several other bacteria including *E. coli*, *S. ruminantium*, and *B. mycoides* host such small, cryptic, RepA-containing plasmids. Burian et al. (1997) proposed that such genetic entities with no discernible phenotypes, generally termed small cryptic plasmids (SCPs), represent a class of stably replicating 'selfish DNA' that requires further investigation.

RepA of pHS649 also had homology to the Rep protein of plasmid pBMBt1 (data not shown). Plasmid pBMBt1 has been shown to possess ssDNA replication intermediates by Southern blotting (Loeza-Lara et al., 2005), indicating that it replicates by the RC mode. Plasmid pHS649S was constructed by inserting a streptomycin gene into pHS649. Although this cloning

strategy utilizing the SpeI restriction site apparently interrupted the putative SSO-DSO region of pHS649, pHS649S was found to stably replicate in *E. coli* strains. This indicates that interruption of the contiguity of SSO and DSO DNA sequence does not hinder the replication and maintenance of pHS649. However, pHS649S purified from *E. coli* was found to contain increased levels of ssDNA intermediates, indicating that interruption of the contiguity of SSO and DSO DNA sequence may lead to accumulation of ssDNA intermediates. In plasmids that replicate by the rolling circle (RC) mode, an active tyrosine residue within the Rep proteins catalyzes nicking-closing reactions during replication (Khan, 1997). It is possible that the conserved tyrosine residue in pHS649 RepA has a similar role. Therefore, it is proposed that pHS649 and its derivatives pHS649S and pHS649SS replicate by the RC mode. The high copy number shuttle vector pHS649SS stably replicates in strains of *E. coli* as well as *H. somni* strain 129Pt. With suitable modifications, pHS649SS could also be used to transform other strains of *H. somni*.

The 5,178 bp plasmid from preputial isolate strain 129Pt is thus far the largest *H. somni* plasmid that has been completely sequenced. The GC content of pHS129 is less than the GC content of the genomes of *H. somni* strains 2336 and 129Pt, which indicates that this plasmid may also have originated by horizontal transfer. Interestingly, this plasmid is comprised of two IRs, which make up about 50% of the total plasmid sequence. Although conceptual translation assigned several ORFs to the sequence of pHS129, only *repB*, whose translated protein may be involved in plasmid replication, could be positively identified in duplicate based on homology search. The occurrence of two *repB* genes, but two identical iteron-containing origins of replication, is a striking feature of pHS129. Some multicopy plasmids have been shown to undergo homologous recombination within their native hosts to form a dimeric plasmid

(Summers, 1998). It is possible that pHS129 originated by such a mechanism. Based on features within pHS129 that resemble plasmids replicating by the theta mechanism, it is proposed that pHS129 is also a cryptic plasmid that replicates by the theta mechanism.

Furthermore, most of the *Pasteurellaceae* and *Neisseriae* plasmids related to pHS129 by way of RepB are 'R plasmids' that encode for resistance to one or more antibiotics and *H. paragallinarum* plasmid p250 encodes a bacteriocin (Terry et al., 2003). Although all these plasmids appear to have had a common early ancestor, it is remarkable that during evolution some have acquired virulence genes while others have not. One of the causes of the evolution of antibiotic resistance in bacteria is selective pressure induced by antibiotic usage as part of nutrition or therapy. The most common therapeutic regimen against *H. somni* induced diseases involves the use of antibiotics (Inzana, 1999). In light of this, it is interesting to note that clinical and laboratory antibiotic resistance is uncommon in *H. somni* (Aarestrup et al., 2004) and that plasmids containing antibiotic resistance genes are infrequent in *H. somni*, if any.

From table 3.04, it is apparent that bacteria within the *Pasteurellaceae* contain a variety of plasmids. Some have been explored more extensively owing to their obvious role in virulence while others have been characterized as cryptic in view of their unknown functions. Whilst some of these plasmids encode proteins that modulate their replication, others conspicuously lack such proteins and probably are entirely dependent on host proteins for their own propagation. Some plasmids, like pLS88, have a very broad host range (Dixon et al., 1994) while others appear to have a narrow host range and have not been shown to replicate in *E. coli* strains. According to Campbell et al. (1999), DNA sequence signatures determine plasmid viability in a bacterial host. The viability of pHS649 as well as the compatibility of pHS649 and pHS129 with each other in *H. somni* may be explained once the dinucleotide relative abundances (Campbell et al., 1999)

within these plasmids and the chromosomes of *H. somni* strains are computed and compared. The inability of pHS129 to replicate in *E. coli* strains could also be explained by similar analyses.

Bakkali et al. (2004) examined the evolutionary stability of the *H. influenzae* DNA uptake signal sequence (USS) and reported that the 9 bp core sequence 5'-AAGTGCGGT-3' is well conserved in some genera (a total of 927, 1205, and 1760 copies of the USS were identified in the genomes of *P. multocida*, *H. somni*, and *A. actinomycetemcomitans*, respectively). Since an evolutionary association between all the diverse plasmids from members of the *Pasteurellaceae* shown in table 3.04 would be difficult to establish with the data and tools available, the sequences of these plasmids were examined for the presence of *H. influenzae* DNA USS. Of the thirteen plasmids from the *Pasteurellaceae* that were examined, only one, p57/98 from *H. somni* (Izadpanah et al., 2001), was found to carry a single copy of the USS 5'-AAGTGCGGT-3'. These results are not surprising since Bakkali et al. (2004) found only 41 copies of this USS in the 1.7 Mb genome of *H. ducreyi*. Nevertheless, the analysis indicates that the occurrence of the USS may not be a reliable indicator of the evolutionary relationship between plasmids. It would be interesting to see if the dinucleotide relative abundances within these plasmids correlate with each other and if they indicate any evolutionary relationship.

In conclusion, results presented in this study add to the dissimilarities documented between *H. somni* strains 649 and 129Pt. This study also shows that comparing plasmid DNA sequences and the proteins encoded by them can facilitate understanding the genetic diversity of bacteria within a genus/family. Such analyses will be more meaningful as the complete DNA sequences of more *Haemophilus/Histophilus* plasmids become available.

References

Aarestrup FM, Seyfarth AM, Angen O (2004). Antimicrobial susceptibility of *Haemophilus parasuis* and *Histophilus somni* from pigs and cattle in Denmark. *Veterinary Microbiology* **101**: 143-146.

Appuhamy S, Low JC, Coote JG, Parton R (1998). PCR methods and plasmid profile analysis for characterisation of *Histophilus ovis* strains. *Journal of Medical Microbiology* **47**: 987-992.

Brown LN, Eness PG, Self HL (1972). Application of the complement fixation test to the study of *Haemophilus somnus* infections of cattle. *Proceedings of the Annual Meeting of the U.S. Animal Health Association* **76**: 502–508.

Burian J, Guller L, Macor M, Kay WW (1997). Small cryptic plasmids of multiplasmid, clinical *Escherichia coli*. *Plasmid* **37**: 2-14.

Burian J, Stuchlik S, Kay WW (1999). Replication control of a small cryptic plasmid of *Escherichia coli*. *Journal of Molecular Biology* **294**: 49-65.

Campbell A, Mrazek J, Karlin S (1999). Genome signature comparisons among prokaryote, plasmid, and mitochondrial DNA. *Proceedings of the National Academy of Sciences of the U.S.A.* **96**: 9184-9189.

Corbeil LB, Blau K, Prieur DJ, Ward AC (1985). Serum susceptibility of *Haemophilus somnus* from bovine clinical cases and carriers. *Journal of Clinical Microbiology* **22**: 192-198.

del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R (1998). Replication and control of circular bacterial plasmids. *Microbiology and Molecular Biology Reviews* **62**: 434–464.

Frey J and MacInnes JI (1995). Genetic manipulation of members of the family *Pasteurellaceae*. In: Donachie W, Lainson FA, Hodgson JC (Editors). *Haemophilus, Actinobacillus, and Pasteurella.* New York, NY: Plenum Press, pp. 187–195.

Fussing V and Wegener HC (1993). Characterization of bovine *Haemophilus somnus* by biotyping, plasmid profiling, REA-patterns and ribotyping. *Zentralblatt fur Bakteriologie. International Journal of Medical Microbiology, Virology, Parasitology and Infectious Diseases* **279**: 60-74.

Higgins R, Godbout-DeLaSalle F, Messier S, Couture Y, Lamothe P (1981). Isolation of *Histophilus ovis* from vaginal discharge in ewes in Canada. *Canadian Veterinary Journal* **22**: 395-396.

Howard MD, Cox AD, Weiser JN, Schurig GG, Inzana TJ (2000). Antigenic diversity of *Haemophilus* somnus lipooligosaccharide: phase-variable accessibility of the phosphorylcholine epitope. *Journal of Clinical Microbiology* **38**: 4412-4419.

Inzana TJ, Gogolewski RP, Corbeil LB (1992). Phenotypic phase variation in *Haemophilus somnus* lipooligosaccharide during bovine pneumonia and after *in vitro* passage. *Infection and Immunity* **60**: 2943–2951.

Izadpanah R, Dan A, Benko M, Rusvai M, Fodor L, Harrach B (2001). DNA sequence of a small, unidentified plasmid isolated from a *Haemophilus somnus* strain: short communication. *Acta Veterinaria Hungarica* **49**: 11-16.

Khan SA (1997). Rolling-circle replication of bacterial plasmids. *Microbiology and Molecular Biology Reviews* **61**: 442–455.

Lenski RE, Simpson SC, Nguyen TT (1994). Genetic analysis of a plasmid-encoded, host genotype-specific enhancement of bacterial fitness. *Journal of Bacteriology* **176**: 3140-3147.

Moreira D (2000). Multiple independent horizontal transfers of informational genes from bacteria to plasmids and phages: implications for the origin of bacterial replication machinery. *Molecular Microbiology* **35**: 1-5.

Nordstrom K and Dasgupta S (2006). Copy-number control of the *Escherichia coli* chromosome: a plasmidologist's view. *EMBO Reports* **7**: 484-489.

Sambrook J and Russell DW (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Sanders JD, Tagawa Y, Briggs RE, Corbeil LB (1997). Transformation of a virulence associated gene of *Haemophilus somnus* into a strain lacking the gene. *FEMS Microbiology Letters* **154**: 251-258.

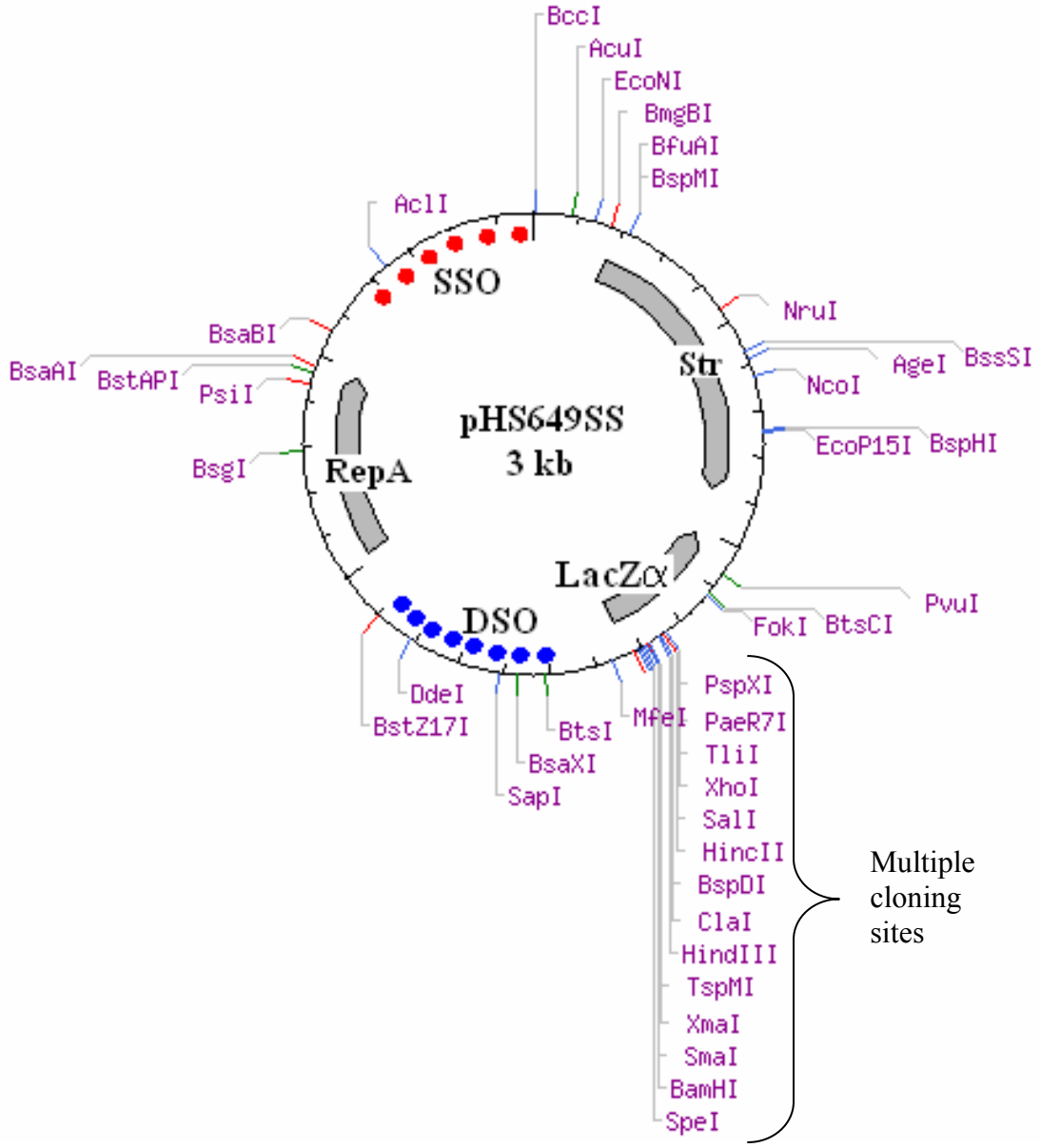
Smith J (2001). The social evolution of bacterial pathogenesis. *Proceedings of the Royal Society of London. Series B, Biological Sciences* **268**: 61-69.

Summers D (1998). Timing, self-control and a sense of direction are the secrets of multicopy plasmid stability. *Molecular Microbiology* **29**: 1137-1145.

Walker RL, Biberstein EL, Pritchett RF, Kirkham C (1985). Deoxyribonucleic acid relatedness among '*Haemophilus somnus*', '*Haemophilus agni*', '*Histophilus ovis*', '*Actinobacillus seminis*' and *Haemophilus influenzae*. *International Journal of Systematic Bacteriology* **35**: 46–49.

Widders PR, Paisley LG, Gogolewski RP, Evermann JF, Smith JW, Corbeil LB (1986). Experimental abortion and the systemic immune response to "*Haemophilus somnus*" in cattle. *Infection and Immunity* **54**: 555-560.

Wright CL, Strugnell RA, Hodgson ALM (1997). Characterization of a *Pasteurella multocida* plasmid and its use to express recombinant proteins in *P. multocida*. *Plasmid* **37**: 65-79.



Appendix 3.01: Detailed restriction enzyme map of pHS649SS

Restriction enzymes that have a single recognition site in pHS649SS and the multiple cloning site region are marked. Other restriction enzymes that have a single recognition site in the multiple cloning site region, but are not marked in the above map, are EcoRV, NotI, and XbaI. Putative SSO and DSO as well as replication protein (RepA), streptomycin (Str), and β-galactosidase α fragment (LacZα) coding regions are also shown.

Appendix 3.02: Table of restriction enzymes that have a single recognition site in pHS649SS

#	Restriction enzyme	Recognition sequence	Position in pHS649SS
1	AclI	AA^CG_TT	2740/2742
2	AcuI	CTGAAG(N)$_{14}$_NN^	88/86
3	AgeI	A^CCGG_T	588/592
4	AleI	CACNN_NNGTG	1318
5	BamHI	**G^GATC_C**	1288/1292
6	Bme1580I	G_KGCM^C	1232/1228
7	BmgBI	CAC_GTC	173
8	BsaAI	YAC_GTR	2480
9	BsaBI	GATNN_NNATC	2557
10	BsgI	GTGCAG(N)$_{14}$_NN^	2291/2289
11	BspDI	AT^CG_AT	1253/1255
12	BspHI	T^CATG_A	743/747
13	BspMI	ACCTGCNNNN^NNNN_	216/220
14	BspQI	GCTCTTCN^NNN_	1609/1612
15	BssSI	C^ACGA_G	569/573
16	BstAPI	GCAN_NNN^NTGC	2464/2461
17	BstZ17I	GTA_TAC	1897
18	BtsCI	GGATG_NN^	1115/1113
19	BtsI	GCAGTG_NN^	1511/1509
20	ClaI	**AT^CG_AT**	1253/1255
21	DdeI	C^TNA_G	1813/1816
22	EcoNI	CCTNN^N_NNAGG	137/138
23	EcoRV	**GAT_ATC**	1266
24	FokI	GGATG(N)$_9$^NNNN_	1122/1126
25	HincII	**GTY_RAC**	1245
26	HindIII	**A^AGCT_T**	1258/1262
27	MfeI	C^AATT_G	1367/1371
28	NcoI	C^CATG_G	630/634
29	NotI	**GC^GGCC_GC**	1307/1311
30	NruI	TCG_CGA	468
31	NsiI	A_TGCA^T	1730/1726
32	PaeR7I	C^TCGA_G	1237/1241
33	PsiI	TTA_TAA	2435
34	PspXI	VC^TCGA_GB	1237/1241
35	PvuI	CG_AT^CG	1067/1065
36	SalI	**G^TCGA_C**	1243/1247
37	SapI	GCTCTTCN^NNN_	1609/1612
38	SmaI	**CCC_GGG**	1284
39	SpeI	**A^CTAG_T**	1294/1298
40	TliI	C^TCGA_G	1237/1241
41	TspMI	C^CCGG_G	1282/1286
42	XbaI	**T^CTAG_A**	1300/1304
43	XhoI	**C^TCGA_G**	1237/1241
44	XmaI	**C^CCGG_G**	1282/1286

The table above was generated using NEBcutter on pHS649SS DNA sequence (http://tools.neb.com/NEBcutter2/index.php). Only recognition sites marked in bold have been confirmed by digestion of pHS649SS with the corresponding enzymes and gel electrophoresis.