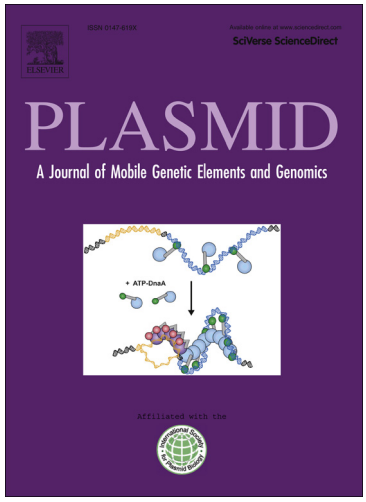
Accepted Manuscript



The Complete Plasmid Sequences of *Salmonella Enterica* Serovar Typhimurium U288

Steven P T Hooton, Andrew R Timms, Nicola J. Cummings, Joanna Moreton, Ray Wilson, Ian F Connerton

PII: S0147-619X(14)00061-4
DOI: http://dx.doi.org/10.1016/j.plasmid.2014.08.002
Reference: YPLAS 2221

To appear in: *Plasmid*

Received Date: 24 October 2013
Accepted Date: 21 August 2014

Please cite this article as: Hooton, S.P.T., Timms, A.R., Cummings, N.J., Moreton, J., Wilson, R., Connerton, I.F., The Complete Plasmid Sequences of *Salmonella Enterica* Serovar Typhimurium U288, *Plasmid* (2014), doi: http://dx.doi.org/10.1016/j.plasmid.2014.08.002

THE COMPLETE PLASMID SEQUENCES OF *SALMONELLA ENTERICA* SEROVAR TYPHIMURIUM U288

Steven P T Hooton[1], Andrew R Timms[1], Nicola J. Cummings[1], Joanna Moreton[2], Ray Wilson[3] & Ian F Connerton[1]*

[1]Division of Food Sciences, School of Biosciences, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD UK; [2]School of Veterinary Medicine and Science, University of Nottingham, Sutton Bonington Campus, Loughborough LE12 5RD UK; [3]DeepSeq, University of Nottingham, Queens Medical Centre, Nottingham NG7 2UH UK.

*Corresponding author: e-mail ian.connerton@nottingham.ac.uk Tel +115 9516119 Fax +115 9516162

Abstract

Salmonella enterica Serovar Typhimurium U288 is an emerging pathogen of pigs. The strain contains three plasmids of diverse origin that encode traits that are of concern for food security and safety, these include antibiotic resistant determinants, an array of functions that can modify cell physiology and permit genetic mobility. At 148,711 bp, pSTU288-1 appears to be a hybrid plasmid containing a conglomerate of genes found in pSLT of *S*. Typhimurium LT2, coupled with a mosaic of horizontally-acquired elements. Class I integron containing gene cassettes conferring resistance against clinically important antibiotics and compounds are present in pSTU288-1. A curious feature of the plasmid involves the deletion of two genes encoded in the *Salmonella* plasmid virulence operon (*spvR* and *spvA*) following the insertion of a *tnpA* IS26-like element coupled to a bla_{TEM} gene. The *spv* operon is considered to be a major plasmid-encoded *Salmonella* virulence factor that is essential for the intracellular lifecycle. The loss of the positive regulator SpvR may impact on the pathogenesis of *S*. Typhimurium U288. A second 11,067 bp plasmid designated pSTU288-2 contains further antibiotic resistance determinants, as well as replication and mobilization genes. Finally, a small 4,675 bp plasmid pSTU288-3 was identified containing mobilization genes and a *pleD*-like G-G-D/E-E-F conserved domain protein that modulate intracellular levels of cyclic di-GMP, and are associated with motile to sessile transitions in growth.

Keywords: *Salmonella* Typhimurium U288, hybrid virulence plasmid, antibiotic resistance plasmid, class I integron, transposable element, *spv* operon, G-G-D/E-E-F domain.

1. Introduction

The Gram negative bacterial genus *Salmonella* (part of the family *Enterobacteriaceae*) is composed of two distinct facultative anaerobic species: - *S. enterica* and *S. bongori* (Fluit, 2005). The six subspecies that define *S. enterica* (*enterica* [I], *salamae* [II]), *arizonae* [IIIA], *diarizonae* [IIIB], *houtenae* [IV], and *indica* [VI]) constitute in excess of 2,500 recognized serovars. *S. bongori* previously composed group V prior to recognition as an independent lineage (Fluit, 2005; Fookes et al., 2011; Lan et al., 2009). Whilst *S. bongori* is most often associated with infections of ectothermic animals such as lizards, there are limited reports of human infection (Fookes et al., 2011). In contrast, many of the serovars that form the group *S. enterica* are considered to be significant human and animal pathogens (Fookes et al., 2011), with *S. enterica* implicated as the causative agent in >99% of identified salmonellosis incidents in humans (Jacobsen et al., 2011). The majority of *S. enterica* that cause disease can be further grouped in terms of their host range, for instance, host-restricted serovars *S.* Typhi and *S.* Paratyphi A are limited to causing typhoid fever in humans (Baker and Dougan, 2007; Holt et al., 2008; Parkhill et al., 2001) whilst the causative agents of fowl typhoid - *S.* Gallinarum and pig paratyphoid - *S.* Choleraesuis are similarly host-restricted (Thomson et al., 2008). However, such serovars retain the ability to infect other animals, often with severe consequences in terms of disease-outcomes (Betancor et al., 2012; Ye et al., 2011). The broad host range serovars, notable examples include *S.* Enteritidis PT4 (Thomson et al., 2008) and *S.* Typhimurium DT104, are capable of causing gastrointestinal infections in a wide range of hosts with *S.* Typhimurium DT104 capable of infecting both humans and food-producing animals such as cattle, sheep, and pigs (Cooke et al., 2008; Gebreyes and Altier, 2002; Mulvey et al., 2006).

Over the last decade *S*. Typhimurium U288 has emerged as a significant pathogen of pigs in the UK and is now the serovar most often identified through surveillance programmes by the UK Veterinary Laboratories Agency (Mueller-Doblies et al., 2013). It has previously been reported that many *S*. Typhimurium U288 isolates of porcine origin possess a transmissible 150 kilobase (kb) plasmid, often found in tandem with smaller 14 and 3 kb plasmids. Analysis of the resistance phenotypes and associated genotypes of each *S*. Typhimurium U288 revealed the presence of determinants that provide protection against an array of antibiotics of clinical importance (Anjum et al., 2011) with the most commonly displayed antibiotic resistance profile being AmCSSuTTm (VLA, 2012). In order to further our understanding of *S*. Typhimurium U288, a whole genome sequencing approach was employed (Hooton et al., 2013). Following assembly of the sequence data obtained, three separate plasmids were constructed and analyzed and are presented in the course of this work.

2. Materials & methods

2. 1 Plasmid DNA sequencing and characterisation

The complete genome of *S*. Typhimurium U288 was sequenced from genomic DNA using the Roche 454 GS FLX sequencing system (Roche Diagnostics, USA) as described previously (Hooton et al., 2013). Following *de novo* assembly of the genomic DNA sequence reads, a single contig of 4,852,606 bp representing the *S*. Typhimurium U288 chromosome was constructed (Genbank Acc. No. CP003836). Sequence reads that did not map to the chromosome were then independently assembled into three circular plasmid DNAs ready for annotation (148,711 bp pSTU288-1, 11,067 bp pSTU288-2, and 4,675 bp pSTU288-3). Confirmation of the presence of three plasmids in *S*. Typhimurium U288 that correspond to those predicted from the assembled sequence reads was achieved by analysis of plasmid preparations. Briefly, an overnight culture of *S*. Typhimurium U288 was subjected

to GenElute™ Plasmid Miniprep extraction as per manufacturer's instructions (Sigma Aldrich, UK). Eluate from the miniprep was electrophoresed in a 0.7% agarose ethidium bromide-stained gel and visualized using a Biorad Image Capture (Biorad, USA). Plasmid preparations were used to transform competent *E. coli* DH5α that enabled the efficient recovery of the antibiotic resistance plasmid pSTU228-2 by selection on LB-agar plates containing either streptomycin (10 µg/mL) or tetracycline (10 µg/mL) or chloramphenicol (30 µg/mL) or neomycin (10 µg/mL).

2.2 DNA sequence analysis and annotation

Several different methods were employed in order to annotate the three plasmids including NCBI PGAAP (http://www.ncbi.nlm.nih.gov/genomes/static/Pipeline.html). For pSTU288-1, xBASE2 automatic annotation pipeline was used with pSLT of *S*. Typhimurium LT2 as the reference sequence (Chaudhuri et al., 2008). Regions of the pSTU288-1 genetic map which were insufficiently annotated through PGAAP and xBASE2 were manually curated using a combination of BLASTN searching (Altschul et al., 1990) and the Artemis sequence viewing tool (Rutherford et al., 2000). Due to the small sizes of pSTU288-2 and pSTU288-3 each plasmid was manually curated as described above. When necessary, translated amino acid sequences from open reading frames (ORFs) identified on each plasmid were subjected to protein domain searches using the Pfam database (Finn et al., 2010). All pairwise alignments of protein and DNA sequences were performed using ClustalW2 with default settings (Larkin et al., 2007). The Integrall database was accessed to assign integron families (Moura et al., 2009). The fully annotated plasmids were deposited in Genbank under the following accession numbers - pSTU288-1 (CP004058), pSTU228-2 (CP004059), and pSTU288-3 (CP004060).

2.3 Conjugation

Mating experiments were carried out by mixing exponential growth phase broth cultures of donor *S.* Typhimurium U288 and recipient *E. coli* DH5α at a ratio of approximately 1:3 on LB-agar plates and incubating for six hours at 37 ^{o}C. The mating mixture was resuspended in LB medium and plated as serial dilutions on LB-agar plates containing nalidixic acid (50 µg/mL) to counter-select *S.* Typhimurium U288 and streptomycin (10 µg/mL) for plasmids replicating in *E. coli* DH5α. LB-agar plates containing either ampicillin (25 µg/mL) or tetracycline (10 µg/mL) or trimethoprim (25 µg/mL) were used to discriminate the respective carriage of plasmids pSTU288-1 and pSTU228-2. Plasmid transfer frequencies are expressed as number of transconjugants per donor cell. Donor cells were enumerated by plating serial dilutions onto xylose lysine desoxycholate agar (XLD; Oxoid, Basingstoke UK).

3 Results

3.1 pSTU288-1 large virulence and antibiotic resistance plasmid

The largest of the plasmids identified following whole genome sequencing of *S.* Typhimurium U288 was the 148,711 bp large virulence and antibiotic resistance plasmid pSTU288-1 (Figure 1). The GC content of pSTU288-1 was observed to be 52.95% which is marginally above the 52.18% GC content determined for the *S.* Typhimurium U288 chromosome, but similar to other large *Salmonella* virulence plasmid sequences deposited in the database. BLASTN analysis of the nucleotide sequence of pSTU288-1 reveals that its closest relative is the 93,939 bp pSLT plasmid (GenBank Acc. No. AE006471) of *S.* Typhimurium LT2 (McClelland et al., 2001). It should be noted that many other *S.* Typhimurium plasmid sequences also align over various regions of the pSTU288-1 sequence. However, it is apparent that pSTU288-1 is a unique hybrid plasmid containing highly conserved regions comparable with pSLT, accompanied by multiple regions that have been

horizontally-acquired over time. The positions and sequence-related origins are indicated in the inner track of Figure 1. Based on nucleotide sequence alignments between pSTU288-1 and pSLT approximately 85 kb of sequence is highly conserved between the two plasmids. The conserved regions are dispersed throughout the plasmid with the largest continuous region of 56,732 bp showing 99% identity.

Numbering from a Transposase 31-like element insertion within the circular sequence, BLASTN analysis indicates the first 20,079 nucleotides of pSTU288-1 have 99% identity to a region of pSL476_91, a 91kbp plasmid from *S*. Heidelberg str. SL476 (GenBank Acc. No. CP001118). This sequence is notably duplicated in pSTU288-1 as it reoccurs between ~68 kb and 88 kb on the circular sequence of the plasmid. The genes found within the duplicated regions show a high degree of synteny when compared to each other but are clearly distinguishable by numerous insertions and deletions, accompanied by complete breakdown of homology in intergenic non-coding regions. Analysis of the genes found within the duplicated regions show that both are flanked by a Transposase 31-like element (pSTU288-1_0001 and pSTU288-1_0122). The amino acid sequences deduced for the transposases show a high degree of conservation over the first 200 amino acids that are recognizable as transposase 31 family protein domains (PF04754 PD(D/E)XK nuclease superfamily - CL0236), which encompasses a diverse range of nucleases and Holliday junction resolvases (Feder and Bujnicki, 2005). The presence of these transposases at the end of each duplicate sequence suggests a possible mechanism for transposition. Two further genes of known function present in the pSTU288-1 duplicated regions that are absent from the corresponding pSLT region are *ardA* (pSTU288-1_0005) and *ardB* (pSTU288-1_0015), which encode proteins belonging to the respective ArdA (PF07275) and ArdB (PF03230) anti-restriction families. The *ardA* and *ardB* genes show a high degree of nucleotide sequence conservation

153 with homologous genes identified in plasmids isolated from several *Salmonella* serovars, as
154 well as a large number of *E. coli* sequences.

155 Nucleotide sequence conservation between pSTU288-1 and pSLT begins at nucleotide
156 20,720 with a 12.5 kb region (12,445/12,500 bases) that contains several genes including a
157 replication initiation protein (*repA2* - pSTU288-1_0032), a conserved OB-fold protein (*ygiW*
158 - pSTU288-1_0033), as well as the virulence-associated *pef* operon (pSTU288-1_0034 to
159 pSTU288-1_0042). Immediately downstream of pSTU288-1_0045 a complete breakdown in
160 conservation with pSLT is observed due to the presence of a TnIS66 transposable element
161 (pSTU288-1_0043 to 0045). There is an identical TnIS66 element present at nucleotide
162 positions 145,640 to 148,255 of pSTU288-1, as well as a further example integrated in the *S.*
163 Typhimurium U288 chromosome (STU288_09455, STU288_09460, and STU288_09465)
164 that suggests the element continues to be active. Synteny between pSTU288-1 and pSLT is
165 restored after the transposable element resulting in 99% conservation over 64 kb of sequence
166 (bases 35,917-99,920) that features genes involved in virulence, plasmid replication and
167 conjugal transfer.

168 The region spanning nucleotides 121,821-132,269 in pSTU288-1 is conserved with pSLT,
169 which contains several recognizable genes including the *ccdA* and *ccdB* genes associated with
170 plasmid maintenance, a resolvase encoded by *rsdB*, an intimin-like protein encoded by *sinH*
171 and a truncated *spv* (*Salmonella* plasmid virulence) operon. The pSTU288-1 *spv* operon has
172 the configuration *spvBCDE* that departs from the *spvRABCDE* configuration observed
173 elsewhere (Gulig et al., 1993). Analysis of the sequence environment indicates that the *spvR*
174 and *spvA* genes have been deleted in pSTU288-1 following the insertion of a β-lactamase
175 gene (bla_{TEM}) coupled to a transposon. Figure 2 shows a comparison between the truncated
176 *spv* region of pSTU288-1 and the intact region found on pSLT. The transposable element
177 responsible for the deletion of *spvRA* encodes genes *tnpA* (pSTU288-1_0175), *tnpR*

(pSTU288-1_0176) and *bla*$_{TEM}$ (pSTU288-1_0177). TnpA is a 238 amino acid protein that contains a definable DDE_Tnp_IS240 protein domain (PF13610) of the RNase H-like Superfamily (CL0219) spanning residues 77-217. BLASTP analysis shows that the TnpA protein is conserved amongst many pathogenic bacteria but notably has 100% identity with the TnpA IS26 elements found in plasmids of *S*. Choleraesuis and *S*. Typhi CT18. The *tnpR* gene encodes a short 53 amino acid protein (molecular weight 5.9 kDa.) that contains a HTH_7 (helix-turn-helix) resolvase domain (PF02796 - HTH clan CL0123), which is highly conserved with several dozen *Klebsiella pneumoniae* and *E. coli* 0104:H4 TnpR sequences in public databases. The *bla*$_{TEM}$ gene encodes a 286 amino acid protein that belongs to the β-lactamase 2 enzyme family (PF13354), part of the Serine β-lactamase-like Superfamily (CL0013).

The sequences of the plasmid replication gene *repA* (pSTU288-1_0054) and its regulatory partner *repB* (pSTU288-1_0055) indicate that pSTU288-1 is an IncFII incompatibility group plasmid. The 293 amino acid RepA protein contains a definable IncFII RepA protein family domain (PF02387) from residues 10-276. The 88 amino acid RepB protein is a member of the RepB RCR regulatory protein family (PF10723) that regulate plasmid rolling circle replication (Khan, 1997). The genes encoding all the necessary functions for conjugal transfer (the *tra* genes) are located between nucleotides 42,034-77,017 (pSTU288-1_0056 to pSTU288-1_0096) on the genetic map.

3.2 The class I integron of pSTU288-1

A class I integron (In640) is located between nucleotides 99,672 and 114,690 within pSTU288-1 flanked by two large transposable elements. The 3′-CS (conserved sequence) of the class I integron is flanked by an IS26 element linked to genes encoding FabG (pSTU288-1_0134), SulIII (pSTU288-1_0137) and several conserved hypothetical proteins. The *tnpA*-

IS26 transposase gene (pSTU288-3_0131) encodes a 238 amino acid protein containing a DDE_Tnp_IS240 domain (PF13610) of the RNaseH-like superfamily (CL0219). Immediately downstream of *sulIII* is a *tnp_mut* gene (pSTU288-1_0138) that appears to have been inserted into the 3′-CS creating an atypical class I integron structure. Typical class I integrons are found to contain a *qac* gene fused to a *sulI* gene and it is this partnership that defines the 3′-CS region (Chuanchuen et al., 2010; Dawes et al., 2010). The atypical configuration of the 3′-CS in pSTU288-1 consists of a *qacEΔ1/tnp_mut/sulIII* combination. The *intI1* gene (pSTU288-1_0145) defines the 5′-CS region of the class I integron of pSTU288-1. The gene encodes a 337 amino acid protein containing two phage integrase domains that function in a similar manner to the site-specific tyrosine recombinase XerD. The first protein domain spans residues 16-100 and belongs to the Phage Integrase N2 N-terminal SAM-like domain family (PF13495 – λ-integrase N-terminal domain CL0469). Amino acid residues 115-319 encompass a second Phage integrase domain (PF00589) that is part of the DNA Breaking/Rejoining enzyme superfamily (CL0382). The *intI1* gene of pSTU288-1 is situated adjacent to three genes encoding a Tn21 element – *tnpM* (pSTU288-1_0146), *tnpR* (pSTU288-1_0147), and *tnpA* (pSTU288-1_0148). Encoded within the pSTU288-1 class I integron are a number of gene cassettes that confer resistance to several antibiotic classes detailed in Table 1.

Adjacent to the Tn21 element (pSTU288-1_0149 - pSTU288-1_0158) is a 7,097 bp region that aligns with plasmids pSal8934a (GenBank Acc. No. JF274933) of *S.* Typhimurium (100% identity) and pLM (GenBank Acc. No. JQ901381) of *E. coli* str. EC25 (99% identity). The final 16 kb of pSTU288-1 appears to be a hybrid of plasmid sequences of diverse origin but includes three genes associated with plasmid conjugal transfer – *trbA* (pSTU288-1_0181), *trbB* (pSTU288-1_0182), and *trbC* (pSTU288-1_0183), and the plasmid mobilization genes *nikA* (pSTU288-1_0185) and *nikB* (pSTU288-1_0186).

227

228 **3.3 Plasmid conjugal transfer**

229 Plasmid pSTU288-1 could be efficiently transferred by conjugation from *S*. Typhimurium
230 U288 to *E. coli* DH5α at a frequency of 2 (±0.8) $\times 10^{-4}$ transconjugants per donor cell by
231 selection for streptomycin resistance. The antibiotic resistance profiles of the transconjugants
232 confirmed the presence of functional resistance determinants against streptomycin,
233 ampicillin, chloramphenicol and trimethoprim for pSTU288-1. The presence of tetracycline
234 resistance in >90% of the transconjugants implied the efficient co-transfer of pSTU288-2.
235 Plasmid DNAs were recovered from tetracycline resistant transconjugants and used to
236 transform *E. coli* DH5α with selection for tetracycline resistance. Plasmid DNAs could be
237 recovered from the *E. coli* DH5α transformants that corresponded with the expected size of
238 pSTU288-2 upon agarose gel electrophoresis. The transformants exhibited resistance to the
239 antibiotics streptomycin, tetracycline, chloramphenicol and neomycin that is consistent with
240 the profile expected of pSTU288-2.

241

242 **3.4 pSTU288-2 antibiotic resistance plasmid**

243 Plasmid pSTU288-2 is 11,067 bp in size that is notable by its GC content of 61.76%, which
244 is well above 52.18% observed for the *S*. Typhimurium U288 chromosome. The plasmid
245 encodes a complement of genes that confer resistance to antibiotics of both medical and
246 veterinary importance (Table 1). Plasmid preparations of *S*. Typhimurium U288 were used to
247 transform *E. coli* DH5α with selection for either streptomycin or tetracycline resistance.
248 Plasmid DNAs of pSTU288-2 could be recovered from the *E. coli* DH5α transformants based
249 on DNA sequencing data. The antibiotic profiles conferred were similar to those noted above

for pSTU288-2 (streptomycin, tetracycline, chloramphenicol and neomycin), and confirm the potential of the plasmid to transfer functional antibiotic resistance determinants. The plasmid shares 100% identity over 6,435 bases with pTY474p3 of *S*. Typhimurium 4/74 (GenBank Acc. No. CP002490). The DNA sequence of pSTU288-2 is numbered from recognizable mobilization and replication open reading frames. Translations of *mobA* (pSTU288-2_002) and *mobC* (pSTU288-2_001) are identical with those from a broad host range multicopy IncQ plasmid pRSF1010 (GenBank Acc. No. NC_001740) isolated from *E. coli* (Scholz et al., 1989). The MobA protein contains a MobA/MobL domain (PF03389 - Rep-like domain CL0169) associated with plasmid transfer. Rep-like domain proteins are known to bind DNA and to contain a conserved central His-X-His motif, which is present in pSTU288-2 MobA as a His-Cys-His motif at residues 120-122. The following three genes on the plasmid are involved in replication of pSTU288-2 - *repF* (partial), *repA*, and *repC*.

3.5 pSTU288-3 plasmid

S. Typhimurium U288 was also found to possess a small plasmid of 4,675 bp designated pSTU288-3. The GC content of pSTU288-3 (50.99%) is slightly lower than that of the *S*. Typhimurium U288 chromosome. At the nucleotide level pSTU288-3 shares 99% identity with two other small plasmids in the database: pSD4.6 (GenBank Acc. No. JX566768) a small cryptic plasmid carried by an *S*. Derby isolate from pork meat (Bleicher et al., 2013) and pEC278 (GenBank Acc. No. AY589571.1) of *E. coli* 278B. Whilst the similarities between plasmids of equivalent size to pSTU288-3 are restricted to *E. coli* pEC278 and *S*. Derby pSD4.6, it is apparent that segments of the pSTU288-3 nucleotide sequence are present in many plasmid sequences deposited in the GenBank database. Several larger plasmids of approximately 10 kb in length contain the complete pSTU288-3 nucleotide sequence accompanied by an extra 5 kb region that encodes an IS2 transposable element and

a gene conferring resistance to quinolones (*qnrS1*). All of the 10 kb quinolone resistance plasmids are found in *S*. Typhimurium hosts of various origins: pHLR25 (GenBank Acc. No. HE652087) of *S*. Typhimurium HLR25 isolated from a human case of salmonellosis, pTPqnrS-1a plasmid (GenBank Acc. No. AM746977) from a multiple drug resistant *S*. Typhimurium DT193 strain (Kehrenberg et al., 2007), pQnrS1-cp17s (GenBank Acc. No. JN393220) of *S*. Typhimurium 484 isolated from frozen eel chunks, and pST728/06-2 (GenBank Acc. No. EU715253) of an *S*. Typhimurium strain isolated from a human case of salmonellosis (Wu et al., 2008).

The pSTU288-3 nucleotide sequence was aligned with the start position of *E. coli* pEC278 prior to opening up at base position 1. Analysis of all potential open reading frames on the plasmid reveals a total of five coding sequences. The first gene identified on pSTU288-3 is *mobC* (pSTU288-3_001), which encodes the signature residues (49-96) of the MobC family (PF05713) of bacterial mobilization proteins. Immediately downstream of *mobC* is the gene encoding the relaxase component of the relaxosome – *mobA* (pSTU288-3_002). The 499 amino acid MobA protein contains a relaxase/mobilization domain (PF03432) spanning residues 49-247 of the Rep-like domain clan (CL0169). For pSTU288-3 MobA, the conserved central motif His-Asp-His can be found at residues 337-339, which is conserved in the majority of relaxase/mobilization domain-containing MobA proteins present in the database. Located within the *mobA* gene, but +1 in terms of their reading frame, are two co-transcribed genes - *mobB* (pSTU288-3_003) and *mobD* (pSTU288-3_004). The protein products of these two genes form the remaining components of the relaxosome complex: MobB 161 amino acid and MobD 71 amino acid. The MobB protein of pSTU288-3 contains an MbeB-like conserved N-terminal domain (PF04837) situated between residues 1-52. The MobD protein contains an MbeD/MobD-like domain (PF04899) spanning the total length of the protein that is also associated with plasmid mobilization and transfer. The final gene

identified in pSTU288-3 designated *pleD*-like (pSTU288-3_005) encodes a 502 amino acid protein with an estimated molecular weight of 55.4 kDa. The PleD-like protein is found to belong to the G-G-D/E-E-F domain family (PF00990) of the nucleotidyl cyclase superfamily (CL0276). BLASTP analysis of the 502 amino acid PleD-like protein reveals only one true homologue in the database - the 520 amino acid hypothetical protein pSD4.6_002 of *S*. Derby plasmid pSD4.6 (496/502 identities - 99%) that also contains a conserved G-G-E-E-F domain motif.

4. Discussion

The three plasmids identified in *S*. Typhimurium U288 are of significant concern when viewed from clinical, veterinary and food security standpoints. *S*. Typhimurium U288 has the potential to acquire yet further antibiotic resistance genes as it appears to have efficiently integrated mobile DNA from many exogenous sources as indicated by the transposable elements containing bla_{TEM} found in pSTU288-1 and the presence of both plasmid- and chromosomally-integrated TnIS66 elements. *S*. Typhimurium U288 also has the potential to disseminate antibiotic resistance since it encodes the apparatus for conjugal transfer and the plasmids it harbours are mobilizable.

The transposition event that has disrupted the *spv* operon appears a unique occurrence to this *S*. Typhimurium U288 isolate, with no other *S. enterica spv*-containing plasmids possessing this configuration to date. This event could alter the course of infection and provide a reason as to why this particular strain has emerged as a successful and persistent pathogen within UK pig production units. The *spv* operon is considered to be a major virulence factor of *Salmonella* by performing a vital role during the infection process (Matsui et al., 2001). Two genes in particular (*spvB* and *spvC*) are recognized as being essential for the bacterium to survive during the intracellular stages of infection, and subsequent systemic dissemination

(Haneda et al., 2012; Mazurkiewicz et al., 2008). The co-ordinated expression of the *spv* genes is dependent on RpoS (stationary phase σ factor) and the LysR family transcriptional regulator SpvR. SpvR is recognized as being a positive regulator of the *spv* operon via transcription of itself and *spvA* (Fabrega and Vila, 2013; Heiskanen et al., 1994), and is expressed in a temporal fashion during infection. *In vitro* induction of *spvR* is observed as cells are entering the stationary phase of growth in liquid culture (Dodd and Aldsworth, 2002), and *in vivo* induction is associated with the transition from the gastrointestinal to intracellular stages of infection (Kappeli et al., 2011). The absence of *spvR* and *spvA* on pSTU288-1 is likely to impact on the coordination of these events. However, the acquisition of the bla_{TEM} gene is likely to be of significant benefit since it is considered to be a progenitor of the emerging extended-spectrum β-lactamase genes that are currently of global concern (Thomson, 2001). There are many reports in the literature of β-lactamase genes being linked to transposable elements. For example, *S* Choleraesuis SC-B67 has a $bla_{\text{CMY-2}}$ gene (conferring resistance to ceftriaxone) linked to a *tnpA* element in a similar fashion as that observed for pSTU228-1 (Chiu et al., 2006), as does *S*. Braenderup (Chiou et al., 2009). The selective benefit of carrying the bla_{TEM} gene would complement the resistance genes already found distributed on the pSTU288-1 and pSTU288-2 plasmids.

The presence of the PleD-like G-G-D/E-E-F domain-containing protein encoded by pSTU288-3 may play a role in allowing *S*. Typhimurium U288 to persist in the environment. Members of this protein family are associated with the synthesis of the intracellular signalling molecule cyclic di-GMP that has been found in the sequences of many bacterial genomes (Ryjenkov et al., 2005). Cyclic di-GMP is a ubiquitous secondary messenger of bacterial signalling and has been associated with many functions in pathogenic bacteria such as control of biofilm formation, cell cycle, cell differentiation and expression of virulence-associated traits (Romling et al., 2013). The PleD-like protein of pSTU288-3 contains a G-G-E-E-F

348 motif that has the potential to confer a regulatory function that would give *S*. Typhimurium
349 U288 the ability to persist in pig production units within biofilms.

350 It is apparent that a number of recombination and transposition events in the history of
351 pSTU288-1 have allowed development from an ancestral pSLT-like plasmid to take place.
352 Many of the genes in pSTU288-1 appear to have been acquired from *E. coli* strains, with
353 some of them showing little dispersal throughout *S. enterica*. An interesting example of this
354 in pSTU288-1 is the presence of the gene encoding ArdA (alleviation of restriction of DNA).
355 Members of this protein family are recognized as providing an anti-restriction mechanism
356 against Type I restriction-modification (R-M) systems. ArdA proteins have previously been
357 shown to have the ability to inactivate type I R-M systems in two different ways. ArdA has
358 been observed to have affinity to methyltransferase proteins associated with methylation of
359 DNA, as well as having the ability to bind intact R-M complexes. These two specific actions
360 of ArdA allows unmodified plasmid DNA (lacking the cognate modification pattern of the
361 target cell) to enter a recipient cell during conjugation, protected from restriction and
362 modification by type I R-M systems (Nekrasov et al., 2007; Thomas et al., 2003).

363 In summary, *S*. Typhimurium U288 has been demonstrated to possess an exceptional
364 complement of plasmid-borne antibiotic resistance and virulence genes, with the prospect that
365 this may be further enhanced in the future via a number of different mechanisms. The
366 presence of pathogenic microorganisms such as *S*. Typhimurium U288 that display resistance
367 to multiple antibiotic classes, especially within food production environments is of great
368 concern. Whilst it appears that this serovar is adapted to pigs, we have demonstrated that *S*.
369 Typhimurium U288 is capable disseminating antibiotic resistance genes to other
370 microorganisms.

371 **Acknowledgements**

SH was the recipient of a United Kingdom Food Standards Agency Postgraduate Scholarship Scheme grant PG101.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., 1990. Basic local alignment search tool. J Mol Biol. 215, 403-10.

Anjum, M. F., Choudhary, S., Morrison, V., Snow, L.C., Mafura, M., Slickers, P., Ehricht, R., Woodward, M.J., 2011. Identifying antimicrobial resistance genes of human clinical relevance within Salmonella isolated from food animals in Great Britain. J Antimicrob Chemother. 66, 550-559.

Baker, S., Dougan, G., 2007. The genome of Salmonella enterica serovar Typhi. Clin Infect Dis. 45 Suppl 1**,** S29-33.

Betancor, L., Yim, L., Martinez, A., Fookes, M., Sasias, S., Schelotto, F., Thomson, N., Maskell, D., Chabalgoity, J. A., 2012. Genomic Comparison of the Closely Related Salmonella enterica Serovars Enteritidis and Dublin. Open Microbiol J. 6**,** 5-13.

Bleicher, A., Schofl, G., Rodicio Mdel, R., Saluz, H. P., 2013. The plasmidome of a Salmonella enterica serovar Derby isolated from pork meat. Plasmid. 69**,** 202-10.

Chaudhuri, R. R., Loman, N. J., Snyder, L. A., Bailey, C. M., Stekel, D. J., Pallen, M. J., 2008. xBASE2: a comprehensive resource for comparative bacterial genomics. Nucleic Acids Res. 36**,** D543-6.

Chiou, C. S., Lin, J. M., Chiu, C. H., Chu, C. H., Chen, S. W., Chang, Y. F., Weng, B. C., Tsay, J. G., Chen, C. L., Liu, C. H., Chu, C., 2009. Clonal dissemination of the multi-drug resistant Salmonella enterica serovar Braenderup, but not the serovar Bareilly, of prevalent serogroup C1 Salmonella from Taiwan. BMC Microbiol. 9**,** 264.

Chiu, C. H., Chuang, C. H., Chiu, S., Su, L. H., Lin, T. Y., 2006. Salmonella enterica serotype Choleraesuis infections in pediatric patients. Pediatrics. 117**,** e1193-6.

Chuanchuen, R., Ajariyakhajorn, K., Koowatananukul, C., Wannaprasat, W., Khemtong, S., Samngamnim, S., 2010. Antimicrobial resistance and virulence genes in Salmonella enterica isolates from dairy cows. Foodborne Pathog Dis. 7**,** 63-9.

Cooke, F. J., Brown, D. J., Fookes, M., Pickard, D., Ivens, A., Wain, J., Roberts, M., Kingsley, R. A., Thomson, N. R., Dougan, G., 2008. Characterization of the genomes of a diverse collection of Salmonella enterica serovar Typhimurium definitive phage type 104. J Bacteriol. 190**,** 8155-62.

Dawes, F. E., Kuzevski, A., Bettelheim, K. A., Hornitzky, M. A., Djordjevic, S. P., Walker, M. J., 2010. Distribution of class 1 integrons with IS26-mediated deletions in their 3'-conserved segments in Escherichia coli of human and animal origin. PLoS One. 5**,** e12754.

Dodd, C. E., Aldsworth, T. G., 2002. The importance of RpoS in the survival of bacteria through food processing. Int J Food Microbiol. 74**,** 189-94.

Fabrega, A., Vila, J., 2013. Salmonella enterica serovar Typhimurium skills to succeed in the host: virulence and regulation. Clin Microbiol Rev. 26**,** 308-41.

Feder, M., Bujnicki, J. M., 2005. Identification of a new family of putative PD-(D/E)XK nucleases with unusual phylogenomic distribution and a new type of the active site. BMC Genomics. 6**,** 21.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran, P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E. L., Eddy, S. R., Bateman, A., 2010. The Pfam protein families database. Nucleic Acids Res. 38**,** D211-22.

Fluit, A. C., 2005. Towards more virulent and antibiotic-resistant Salmonella? FEMS Immunol Med Microbiol. 43**,** 1-11.

Fookes, M., Schroeder, G. N., Langridge, G. C., Blondel, C. J., Mammina, C., Connor, T. R., Seth-Smith, H., Vernikos, G. S., Robinson, K. S., Sanders, M., Petty, N. K., Kingsley, R. A., Baumler, A. J., Nuccio, S. P., Contreras, I., Santiviago, C. A., Maskell, D., Barrow, P., Humphrey, T., Nastasi, A., Roberts, M., Frankel, G., Parkhill, J., Dougan, G., Thomson, N. R., 2011. Salmonella bongori provides insights into the evolution of the Salmonellae. PLoS Pathog. 7**,** e1002191.

Gebreyes, W. A., Altier, C., 2002. Molecular characterization of multidrug-resistant Salmonella enterica subsp. enterica serovar Typhimurium isolates from swine. J Clin Microbiol. 40**,** 2813-22.

Gulig, P. A., Danbara, H., Guiney, D. G., Lax, A. J., Norel, F., Rhen, M., 1993. Molecular analysis of spv virulence genes of the Salmonella virulence plasmids. Mol Microbiol. 7**,** 825-30.

Haneda, T., Ishii, Y., Shimizu, H., Ohshima, K., Iida, N., Danbara, H., Okada, N., 2011. Salmonella type III effector SpvC, a phosphothreonine lyase, contributes to reduction in inflammatory response during intestinal phase of infection. Cell Microbiol. 14**,** 485-99.

Heiskanen, P., Taira, S., Rhen, M., 1994. Role of rpoS in the regulation of Salmonella plasmid virulence (spv) genes. FEMS Microbiol Lett. 123**,** 125-30.

Holt, K. E., Parkhill, J., Mazzoni, C. J., Roumagnac, P., Weill, F. X., Goodhead, I., Rance, R., Baker, S., Maskell, D. J., Wain, J., Dolecek, C., Achtman, M., Dougan, G., 2008. High-throughput sequencing provides insights into genome variation and evolution in Salmonella Typhi. Nat Genet. 40**,** 987-93.

Hooton, S. P., Timms, A. R., Moreton, J., Wilson, R., Connerton, I. F., 2013. Complete Genome Sequence of Salmonella enterica Serovar Typhimurium U288. Genome Announc. 1.

Jacobsen, A., Hendriksen, R. S., Aaresturp, F. M., Ussery, D. W., Friis, C., 2011. The Salmonella enterica pan-genome. Microb Ecol. 62**,** 487-504.

Kappeli, R., Kaiser, P., Stecher, B., Hardt, W. D., 2011. Roles of spvB and spvC in S. Typhimurium colitis via the alternative pathway. Int J Med Microbiol. 301**,** 117-24.

Kehrenberg, C., Hopkins, K. L., Threlfall, E. J., Schwarz, S., 2007. Complete nucleotide sequence of a small qnrS1-carrying plasmid from Salmonella enterica subsp. enterica Typhimurium DT193. J Antimicrob Chemother. 60**,** 903-5.

Khan, S. A., 1997. Rolling-circle replication of bacterial plasmids. Microbiol Mol Biol Rev. 61**,** 442-55.

Lan, R., Reeves, P. R., Octavia, S., 2009. Population structure, origins and evolution of major Salmonella enterica clones. Infect Genet Evol. 9**,** 996-1005.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., Higgins, D. G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics. 23**,** 2947-8.

Matsui, H., Bacot, C. M., Garlington, W. A., Doyle, T. J., Roberts, S., Gulig, P. A., 2001. Virulence plasmid-borne spvB and spvC genes can replace the 90-kilobase plasmid in conferring virulence to Salmonella enterica serovar Typhimurium in subcutaneously inoculated mice. J Bacteriol. 183**,** 4652-8.

Mazurkiewicz, P., Thomas, J., Thompson, J. A., Liu, M., Arbibe, L., Sansonetti, P., Holden, D. W., 2008. SpvC is a Salmonella effector with phosphothreonine lyase activity on host mitogen-activated protein kinases. Mol Microbiol. 67**,** 1371-83.

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., Wilson, R. K., 2001. Complete genome sequence of Salmonella enterica serovar Typhimurium LT2. Nature. 413**,** 852-6.

Moura A, Soares M, Pereira C, Leitão N, Henriques I, Correia A (2009) INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. Bioinformatics 25:1096–1098.

Mueller-Doblies, D., Speed, K., Davies, R. H., 2013. A retrospective analysis of Salmonella serovars isolated from pigs in Great Britain between 1994 and 2010. Prev Vet Med. 110**,** 447-55.

Mulvey, M. R., Boyd, D. A., Olson, A. B., Doublet, B., Cloeckaert, A., 2006. The genetics of Salmonella genomic island 1. Microbes Infect. 8**,** 1915-22.

Nekrasov, S. V., Agafonova, O. V., Belogurova, N. G., Delver, E. P., Belogurov, A. A., 2007. Plasmid-encoded antirestriction protein ArdA can discriminate between type I methyltransferase and complete restriction-modification system. J Mol Biol. 365**,** 284-97.

Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D., Wain, J., Churcher, C., Mungall, K. L., Bentley, S. D., Holden, M. T., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Cronin, A., Davis, P., Davies, R. M., Dowd, L., White,

N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T. T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T. S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., Barrell, B. G., 2001. Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18. Nature. 413**,** 848-52.

Romling, U., Galperin, M. Y., Gomelsky, M., 2013. Cyclic di-GMP: the first 25 years of a universal bacterial second messenger. Microbiol Mol Biol Rev. 77**,** 1-52.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M. A., Barrell, B., 2000. Artemis: sequence visualization and annotation. Bioinformatics. 16**,** 944-5.

Ryjenkov, D. A., Tarutina, M., Moskvin, O. V., Gomelsky, M., 2005. Cyclic diguanylate is a ubiquitous signaling molecule in bacteria: insights into biochemistry of the GGDEF protein domain. J Bacteriol. 187**,** 1792-8.

Scholz, P., Haring, V., Wittmann-Liebold, B., Ashman, K., Bagdasarian, M., Scherzinger, E., 1989. Complete nucleotide sequence and gene organization of the broad-host-range plasmid RSF1010. Gene. 75**,** 271-88.

Thomas, A. T., Brammar, W. J., Wilkins, B. M., 2003. Plasmid R16 ArdA protein preferentially targets restriction activity of the type I restriction-modification system EcoKI. J Bacteriol. 185**,** 2022-5.

Thomson, K. S., 2001. Controversies about extended-spectrum and AmpC beta-lactamases. Emerg Infect Dis. 7**,** 333-6.

Thomson, N. R., Clayton, D. J., Windhorst, D., Vernikos, G., Davidson, S., Churcher, C., Quail, M. A., Stevens, M., Jones, M. A., Watson, M., Barron, A., Layton, A., Pickard, D., Kingsley, R. A., Bignell, A., Clark, L., Harris, B., Ormond, D., Abdellah, Z., Brooks, K., Cherevach, I., Chillingworth, T., Woodward, J., Norberczak, H., Lord, A., Arrowsmith, C., Jagels, K., Moule, S., Mungall, K., Sanders, M., Whitehead, S., Chabalgoity, J. A., Maskell, D., Humphrey, T., Roberts, M., Barrow, P. A., Dougan, G., Parkhill, J., 2008. Comparative genome analysis of Salmonella Enteritidis PT4 and Salmonella Gallinarum 287/91 provides insights into evolutionary and host adaptation pathways. Genome Res. 18**,** 1624-37.

VLA., 2012. *Salmonella* in livestock production in GB. Chapter 4 Reports of *Salmonella* in pigs http://www.defra.gov.uk/ahvla-en/files/pub-salm12-chp12.pdf

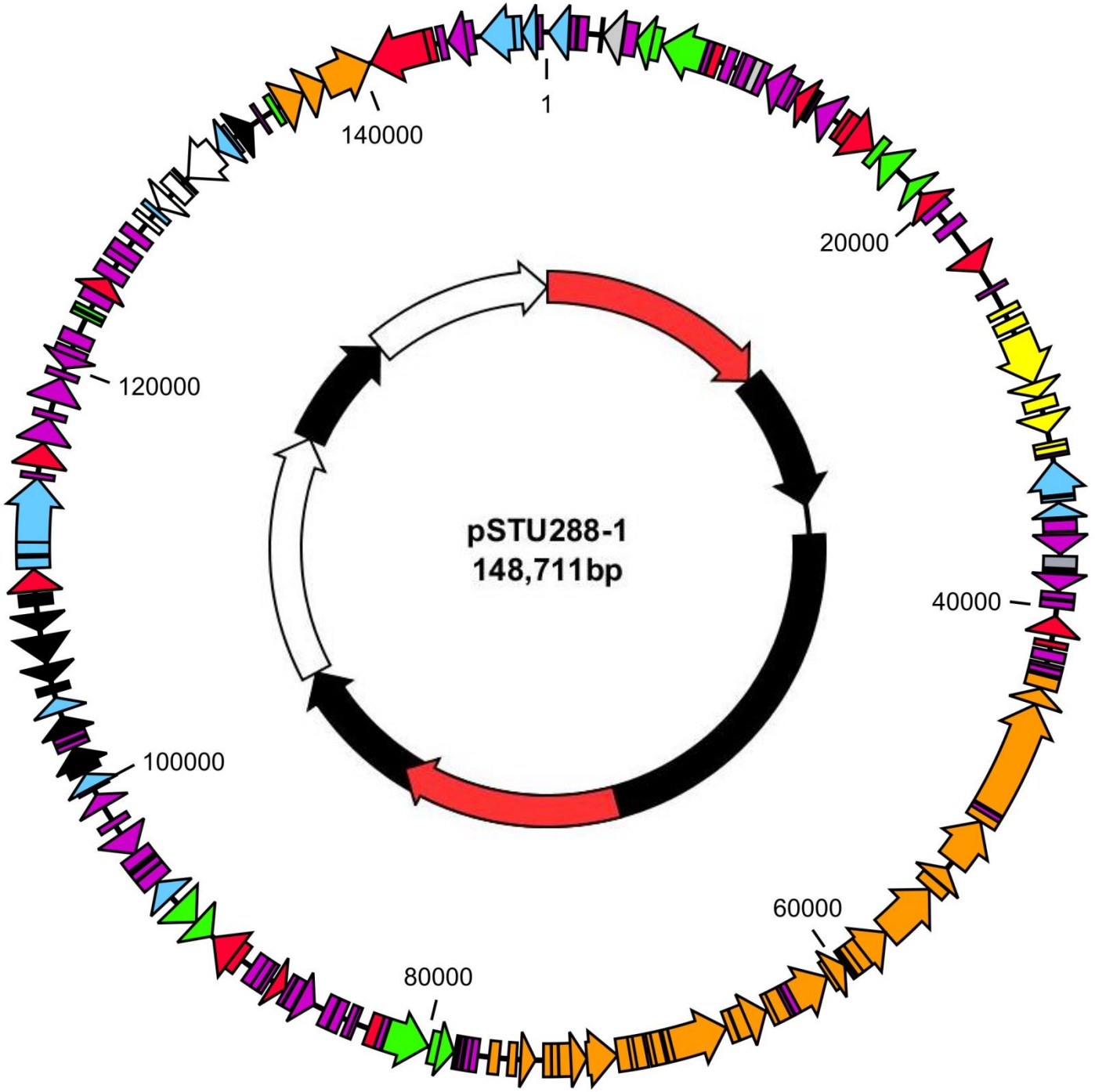
Wu, J. J., Ko, W. C., Chiou, C. S., Chen, H. M., Wang, L. R., Yan, J. J., 2008. Emergence of Qnr determinants in human Salmonella isolates in Taiwan. J Antimicrob Chemother. 62 (6), 1269-72

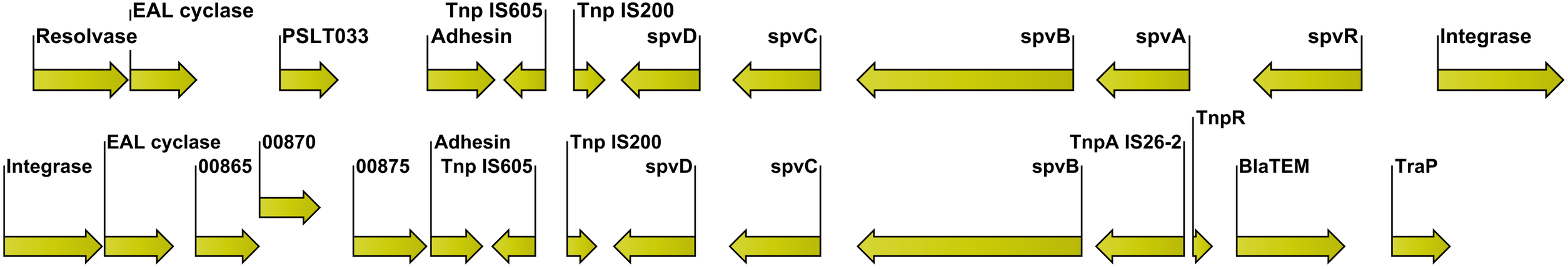
Ye, J., Su, L. H., Chen, C. L., Hu, S., Wang, J., Yu, J., Chiu, C. H., 2011. Analysis of pSC138, the multidrug resistance plasmid of Salmonella enterica serotype Choleraesuis SC-B67. Plasmid. 65**,** 132-40.

Figure legends

Figure 1. A circular representation of the 148,711 bp virulence and antibiotic resistance plasmid pSTU288-1 of *S.* Typhimurium U288. Inner circle - blue arrows (transposons), grey (anti-restriction proteins), red (DNA maintenance/methylation/replication proteins), purple (hypothetical proteins), green (plasmid stabilization/SOS inhibition/maintenance), yellow (plasmid-encoded fimbriae), white (*Salmonella* plasmid virulence proteins). Outer circle - black arrows (pSLT-like), red arrows (duplicated region), and white (non-pSLT).

Figure 2. Comparison of the *spv* genes present on pSLT of *S.* Typhimurium LT2 (upper panel) and pSTU288-1 (lower panel). The *spvR* and *spvA* genes of *S.* Typhimurium U288 have been deleted following a transposition event that introduced a β-lactamase gene (bla_{TEM}) into pSTU288-1. The distal end of the *spv* operon is conserved between the plasmids and features the remnants of previous transposition event(s).


1
20000
40000
60000
80000
100000
120000
140000
pSTU288-1
148,711bp


EAL cyclase
Resolvase
PSLT033
Tnp IS605
Adhesin
Tnp IS200
spvD
spvC
spvB
spvA
spvR
Integrase
Integrase
EAL cyclase
00865
00870
00875
Adhesin
Tnp IS605
Tnp IS200
spvD
spvC
spvB
TnpA IS26-2
TnpR
BlaTEM
TraP

578 **Table 1. Plasmid-encoded antibiotic resistance determinants of *S*. Typhimurium U288**

Gene & locus tag	Function	Predicted MW (kDa.)	Pfam domain
sulIII (pSTU288-1_0137)	Dihydropteroate synthase	29	Pterin binding (PF00809) – TIM barrel superfamily (CL0036)
qacEΔ1 (pSTU288-1_0139)	Small quaternary ammonium compound resistance	11.8	Small drug resistance (PF00893) – Drug/metabolite transporter superfamily (CL0184)
aadA (pSTU288-1_0140)	Streptomycin 3′-adenylyltransferase/aminoglycoside resistance	29.3	Nucleotidyltransferase_2 (PF01909/CL0260)
cmlA (pSTU288-1_0141)	Chloramphenicol resistance	44.3	Major facilitator superfamily (PF07690/CL0015)
aadA2 (pSTU288-1_0142)	Streptomycin 3′-adenylyltransferase/aminoglycoside resistance	29.6	Nucleotidyltransferase_2 (PF01909/CL0260)
dhfR (pSTU288-1_0144)	Dihydrofolate reductase	18.1	DHFR_1 (PF00186) – DHFred (CL0387)
bla_{TEM} (pSTU288-1_0177)	Broad-spectrum β-lactamase	31.5	B-lactamase 2 family (PF13354/CL0013)
sulII (pSTU288-2_006)	Dihydropteroate synthase (Type II)	28.5	Pterin binding (PF00809) – TIM barrel superfamily (CL0036)
strA (pSTU288-2_007)	Streptomycin resistance	29.6	Aph phosphotransferase enzyme family (PF01636)
strB (pSTU288-2_008)	Aminoglycoside/hydroxyurea resistance	30.9	Aminoglycoside/hydroxyurea resistance kinase family (PF04655)
tetA (pSTU288-2_013)	Tetracycline resistance	45.2	Major facilitator superfamily (PF07690/CL0015)
cat (pSTU288-2_014)	Chloramphenicol resistance	13.8	Acetyltransferase GNAT family (PF00583/CL0257)

579

580 **Highlights**

581

582 • We analyse the plasmids of the emerging pig pathogen *Salmonella* Typhimurium U288

583 • Three plasmids encode traits that are of concern for food security and safety

584 • Plasmid-encoded genes confer antibiotic resistance and permit genetic mobility

585 • The plasmids have a mosaic structure with identifiable horizontal gene acquisitions

586