# Microevolution during the emergence of a monophasic *Salmonella* Typhimurium epidemic in the United Kingdom

Running Title: Monophasic *Salmonella* microevolution

Liljana Petrovska[1], Alison E. Mather[2], Manal AbuOun[1], Priscilla Branchu[4], Simon R. Harris[2], Thomas Connor[2], K. L. Hopkins[3], A. Underwood[3], Antonia A Lettini[5], Andrew Page[2], Mary Bagnall[1], John Wain[3], Julian Parkhill[2], Gordon Dougan[2], Robert Davies[1], Robert A. Kingsley[2,4]

[1] Animal and Plant Health Agency, Woodham Lane, New Haw, Addlestone, Surrey, KT15 3NB

[2] The Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, CB10 1SA

[3] Public Health England, Colindale, London, NW9 5EQ.

[4] Institute of Food Research, Norwich Research Park, Colney, Norwich, NR4 7UA

[5] Istituto Zooprofilattico Sperimentale delle Venezie, National Reference Laboratory for Salmonellosis, Legnaro, PD, Italy

Correspondence: rob.kingsley@ifr.ac.uk or Liljana.Petrovska@apha.gsi.gov.uk

## Abstract

Microevolutionary events associated with the emergence and clonal expansion of new epidemic clones of bacterial pathogens hold the key to understanding the drivers of epidemiological success. We describe a comparative whole genome sequence and phylogenomic analysis of monophasic *Salmonella* Typhimurium isolates from the UK and Italy from 2005-2012. Monophasic isolates from this time formed a single clade distinct from recent monophasic epidemic clones described previously from North America and Spain. The current UK monophasic epidemic clones encode a novel genomic island encoding resistance to heavy metals (SGI-3), and composite transposon encoding antibiotic resistance genes not present in other Typhimurium isolates, that may have contributed to the epidemiological success. We also report a remarkable degree of genotypic variation that accumulated during clonal expansion of a UK epidemic including multiple independent acquisitions of a novel prophage carrying the *sopE* gene and multiple deletion events affecting the phase II flagellin locus.

## Introduction

*Salmonella enterica* is one of the most common enteric pathogens of humans and animals. An estimated 94 million cases of non-typhoidal Salmonellosis occur worldwide each year causing considerable morbidity, mortality and an associated economic burden estimated by the US CDC to exceed two billion US dollars per year, in the US alone (1, 2). *S. enterica* consists of more than 2500 serovars, of which *S. enterica* serovar Typhimurium (*S.* Typhimurium) is the most ubiquitous in zoonotic reservoirs for human infection, and the environment (3). The epidemiology of *S.* Typhimurium has been characterized over the past half century by successive waves of dominant multidrug resistant (MDR) clones (4). In Europe, where variants are distinguished by definitive (phage) type (DT), *S.* Typhimurium DT9, DT204, DT104 and DT193 have emerged successively as MDR strains between 1966 and 2010 (5, 6). Epidemic strains dominate for four to 15 years before being replaced by a new dominant phage type. The emergence and spread of *S.* Typhimurium DT104 was global (7) and largely responsible for the increase in *Salmonella* isolates that were MDR in Europe and North America in the 1990s (8). As DT104 incidence has waned in the UK, monophasic variants of *Salmonella* Typhimurium with the antigenic formula 1,4,[5],12:i:- have emerged (9), although it is not clear if this epidemic is related to other epidemics of monophasic variants previously reported in North America (10), Spain (11), and elsewhere in Europe (12). Analysis of the genomic deletions in the phase II flagellum locus responsible for the monophasic phenotype, suggested that there may be multiple independent clones emerging in the US and Europe (10). The first description of a monophasic epidemic in Europe was that of a ‘Spanish clone’ that emerged rapidly during 1997 and was characterized by a deletion in the allantoin-glyoxylate operon and the *fljAB* operon, phage type U302 and a heptaresistance pattern

ACSuGSTSxT (11). Since this time many European countries have reported an increased incidence of this serotype and particularly associated with pig herds (13-16), but later in livestock and poultry (17) (reviewed in 18). However, in contrast to the 'Spanish clone' these have commonly been associated with phage types DT193 or DT120, and a predominant ASSuT resistance pattern suggesting they are distinct epidemics.

The molecular basis for the success of epidemic clones of bacterial pathogens has important implications for the surveillance and management of infectious diseases. Epidemiological success depends on selective advantage of the epidemic clone, due to their unique genotype. We report the whole genome sequence variation of *S.* Typhimurium and *S.* 1, 4,[5],12:i:- isolates from the UK and Italy and the application of these data to phylogenetic reconstruction of the epidemic. We address the questions of whether the monophasic Typhimurium isolates in the UK are part of a single epidemic and how they are related to previously circulating biphasic and monophasic Typhimurium strains. We identify a remarkable level of micro-evolution during clonal expansion of the epidemic that may impact on the antigenicity, pathogenicity and transmission.

## Materials and Methods

Bacterial isolates were from strain collections held by the Animal and Plant Health Agency (APHA), Public Health England (PHE) or Italian Reference Laboratory for *Salmonella* (NRL-IZSVe, Legnaro, Italy). The serotype and phage type were determined as previously described (19). The presence of the *fljB* locus and the occupancy of the *thrW* locus was initially determined by PCR amplification as previously described (12). Strain selection was to represent the diversity of *S.* Typhimurium in the UK and not to be representative of the epidemiology (Supplementary Information and Supplementary Table I). Determination of antimicrobial sensitivity of animal isolates from the UK (APHA) and isolates from Italy (NRL-IZSVe) were tested for susceptibility to the following antimicrobials according to standard procedure (20). Resistance or susceptibility were interpreted on the basis of British Society for Antimicrobial Chemotherapy (BSAC) breakpoints. Intermediate BSAC category was reported here as resistant. Antimicrobial sensitivity of human isolates from the UK (PHE) was determined using a modified breakpoint technique on Isosensitest agar (Oxoid, Basingstoke, UK)(Supplementary Information). MIC for copper sulphate was the concentration at which bacterial growth OD600nm was >0.1 following culture without shaking at 37$^{o}$c for 24 hours in Luria Bertani broth buffered with 25mM HEPES pH7. Determination of whole genome sequence using HiSeq Illumina platform, sequence analysis, de novo assembly, annotation and PCR amplification are described in Supplementary Information.

## Results

**Contemporary *Salmonella* 4,[5],12:i:- strains in the UK are part of a single clonally expanding clade.** The current MDR *Salmonella* 4,[5],12:i:- epidemic in the Europe was first reported around 2005 and is mainly associated with isolates of phage types DT193 and DT120 (21). We investigated the phylogenetic relationship of 206 strains of *S.* Typhimurium (*S.* 1,4,5:i:1,2) and monophasic Typhimurium (*S.* 1,4,[5],12:i:-, *S.* 1,4,12:i:- and *S.* 1,4,5:i:-) isolated from human clinical cases, livestock or contaminated food from the UK or Italy between 1993 and 2010. A total of 97 monophasic *S.* Typhimurium isolates and 142 *S.* Typhimurium isolates were studied (Supplementary Table I). A maximum likelihood (ML) phylogeny of all monophasic and *S.* Typhimurium strains was constructed using 12,793 variable sites in the genome, with reference to the whole genome sequence of strain SL1344, excluding SNPs in prophage, IS-elements and repetitive sequences (Figure 1). The majority of the monophasic strains (77 of 97) were from a single distinct clade that appeared part of the current monophasic epidemic as they were the most abundant and recent isolates. However, older monophasic isolates were also found in at least three other clades within the *S.* Typhimurium tree (* in Figure 1). A clade containing eight isolates including two DT191a (# in Figure 1) was very closely related to a *S.* 1,4,[5],12:i:- isolate from the North American epidemic strain CVM23701 (10). Just six SNPs distinguished this isolate from strain H07 474 0455. In addition, a clade containing six *S.* Typhimurium var Copenhagen (4,12:i:1,2) (e.g., H070160417); and a clade containing four isolates (e.g., H103720606) contained monophasic strains.

**Phylogenetic analysis of the monophasic epidemic in the UK.** In order to study the monophasic Typhimurium epidemic clade in more detail, a maximum likelihood phylogenetic tree was reconstructed using variable sites within the whole genome sequence with reference to the draft genome sequence of a representative strain from within the epidemic (strain SO4698-09). The phylogenetic tree indicated a clonally expanding clade with a maximum root to tip distance of approximately 70 SNPs, indicating that all the strains in the tree shared a common ancestor in the recent past (Figure 2). All isolates form this monophasic clade were of sequence type (ST) 34. The phage type of monophasic epidemic isolates varied depending on phylogeny. The majority of isolates were DT193 (38 of 62 typed) or DT120 (10), and various other phage types including DT7 (3), DT191a (1), DT21 (1), DT21var (1), U311 (3) U302 (2) and RDNC (3). However, while virtually all isolates were of DT193 in subclades A and B, the phage type was highly variable in subclade C. Biphasic DT193 strains (e.g. 4061-1997, Figure 1) isolated before 2005 were not direct ancestors of the current monophasic Typhimurium epidemic as they were present on a distinct lineage. Indeed DT193 isolates were present on four distinct lineages highlighting the polyphyletic nature of this phage type (Figure 1). There was a relative paucity of UK isolates from animals in cluster C; one of 21 isolates in this cluster was from a UK animal. Instead, this cluster was dominated by human isolates from the UK, and isolates from Italy. In contrast cluster A was dominated by isolates of livestock origin (18 of 32), with just five human isolates. Clade B contained approximately equal number of human and livestock isolates. Furthermore, while UK pig isolates were present in all three clusters, UK cattle isolates were only present in cluster A, consistent with epidemiological reports that the epidemic originated in pig herds and were later transmitted to cattle herds (18). Despite only 6 isolates from avian species being included in the analysis

these were distributed throughout the tree suggesting multiple transmission events into these animal populations. There was also a strikingly uneven distribution of human and livestock (pigs, cattle and sheep) isolates within subclades of the phylogenetic tree of UK monophasic isolates. Most isolates (64 of 77) were tetra-resistant (ASSuT) and the corresponding resistance genes were detected in *de novo* assembled sequences (Supplementary Figure 1), suggesting that the MRCA of the clade had this complement of resistance genes. However, during clonal expansion seven strains had lost their resistance genes entirely and a further seven had an altered complement of resistance genes.

**The monophasic epidemic clade contain a novel genetic island encoding heavy metal resistance.** A large novel genomic island (designated SGI-3) specific to the monophasic Typhimurium epidemic clade is inserted at the *yjdC* locus (Supplementary Figure 2) in strain SO4698-09. The island contained approximately 90 genes some of which had sequence similarity to those associated with plasmid transfer and conjugation, and an integrase gene, suggested the island may have originated by integration of a plasmid. Determination of the accessory genome indicated the island was present in 74 of 77 isolates within the monophasic clade (Figure 2), but absent from all strains from outside the clade. Ancestral state reconstruction using ACCTRAN (Supplementary Figure 3A) suggested that introduction of this island likely occurred shortly before clonal expansion of the monophasic clade. Three clusters of genes with similarity to genes involved in resistance to heavy metals are present on the island. Consistent with the island contributing to enhanced resistance to copper sulfate, a common additive to animal feed, isolates within the monophasic Typhimurium clade exhibited a significantly greater MIC ($p=0.015$) to copper sulphate (24.2 +/- 1.9 mM)

than Typhimurium isolates from outside of this clade (21.2 +/- 1.1 mM) that did not encode the island (Supplementary Figure 4).

**Genotypic variation in the *fljBA and thrW loci* and loss of the virulence plasmid.** Monophasic phenotype is due to the absence of phase-2 flagellin monomer FljB. Determination of the presence of the *fljBA* genes and neighbouring genome sequence of *S*. Typhimurium and monophasic variants by mapping raw sequence read data to *fljB* locus region of the SL1344 whole genome sequence (Supplementary Figure 5A) indicated that the UK epidemic strains are monophasic due to multiple independent deletion events that occurred during clonal expansion. Four *S*. Typhimurium isolates (two DT7 isolates, SO5416-06 and H09164 0090, a DT135 isolate, SO6221-07, and a DT177 isolate, H08390 0191) that were closely related and shared a common ancestor with the monophasic epidemic strains (Figure 1), encoded the entire *fljBA* locus indicating that the MRCA with these strains and the epidemic strains was biphasic. In contrast, 67 of 77 monophasic Typhimurium strains from the epidemic clade lacked at least part of the *fljBA* locus, due to deletions ranging in size and with a distribution that was consistent with the phylogenetic relationship of the strains (Supplementary Figure 5A). The eight epidemic strains that did not have a deletion in the *fljB* locus were deeply rooted in the tree, consistent with multiple deletion events (1kb-36kb) occurring since clonal expansion of the clade. Most deletions shared a common junction in the intergenic region of *fljB* and *iroB*. As it was not possible to assemble short read sequence data across the *fljB* locus deletion region, in order to investigate the nature of the deletion and we generated long read sequence data for a representative isolate SO4698-09 using the PACBIO sequencing platform. A single contig assembly of these data revealed a deletion of 15,726bp of the genome relative to SL1344 and an insertion

of 27,473bp of novel sequence (Supplementary Figure 5B). The inserted sequence included sequence with similarity to a number of genes from transposon Tn21, mercury resistance genes (*merTABCDE* and *merR*) and antibiotic resistance genes, consistent with the resistance profile of this strain (*strA*, *strB*, *sul2*, *tet(B)* and *blaTEM-1*). The composite transposon insertion was not present in closely related isolates eg SO5416-06 (Figure 1) that were outside of the monophasic clade, suggesting that it was acquired by the MRCA of the monophasic clade and not prior to clonal expansion. The deletions in the *fljB* locus of monophasic strains from outside the main UK clade were distinct from that in the UK monophasic clade, but identical to those described previously for strains from North American epidemic (eg CVM23701)(10), and Spain (eg 1115/25) (11)(Supplementary Figure 5A).

In addition to hyper-variability at the *fljB* locus, isolates from the epidemic group exhibited sporadic loss of the virulence plasmid pSLT. The pattern of plasmid loss within the clade could be most parsimoniously explained by loss during clonal expansion. Intriguingly, the loss of pSLT was not uniform across the monophasic tree. While just 13% and 20% of isolates tested contained pSLT in sub-clades A and C respectively, in contrast over 70% of isolates in sub-clade B contained the plasmid (Figure 2).

**The *sopE* virulence gene was acquired on a novel prophage mTmV by multiple independent events during clonal expansion of the epidemic clade.** The *thrW* locus of contemporary monophasic Typhimurium isolates has previously been reported to harbor either a prophage, a novel genetic island, or neither (12). In strain SO4698-09, the *thrW* locus contains the novel genetic island described previously, but also an

additional prophage element encoding the *sopE* gene that together total 55 kb. Determination of the accessory genome using the Roary pan genome pipeline (22) indicated that 23 of 77 monophasic isolates from the epidemic clade contained the *sopE* gene (Figure 2). SopE is a guanine exchange factor (GEF) involved in subversion of the host enterocyte cytoskeleton, a key component of the infection process (12, 23, 24). The *sopE* gene was present in six distinct clusters of the monophasic clade, and ancestral state reconstruction indicated that multiple independent acquisitions followed by clonal expansion of the *sopE* -positive variant was the most likely explanation of their distribution (Supplementary Figure 3B). The *sopE* gene of strain SO4698-09 is present on a 55 kb region (designated mTmV phage, for monophasic Typhimurium V) that was absent from strain SL1344 and shared the greatest similarity with the SfV prophage from *Shigella flexneri* (Supplementary Figure 6)(25). SfV_SO4698-09 was not related to the FELS-2 prophage of *S.* Typhimurium strain SL1344 that also encodes the *sopE* gene, except in a 2443bp region that encoded the *sopE* gene and flanking sequence. Examination of partial assemblies of other monophasic strains encoding *sopE* revealed that the gene was associated with the same prophage, and inserted between the genome region corresponding to the *thrW* locus. These data indicated that a novel *sopE*-phage entered the genome on at least six occasions during the clonal expansion of the epidemic clade. Since the *sopE* gene was present in phylogenetic clusters toward the terminal branches of the monophasic clade tree and subsequently exhibited clonal expansion, we addressed the question of whether the proportion of strains isolated from each year in our strain collection that encoded the *sopE* gene changed during the period 2005 to 2010. The frequency distribution for each year was determined from collated data from strains for which date of isolation and sequence data were available (59 strains) and an additional 41 randomly selected monophasic strains from the UK for

255 which the presence of the *sopE* gene was determined by PCR (Figure 3, Supplementary
256 Table II). An increase in frequency ranging from 0% in 2005 and 2006, to 33% in 2010
257 was observed suggesting that acquisition of this gene may have conferred a competitive
258 advantage.

## Discussion

The phylogenetic relationship of *S.* 1,4,[5],12:i:- isolated from the US and Europe since the late 1990's is unclear from reports to date. Our analyses suggest that there are at least three distinct epidemics associated with *S.* 1,4,[5],12:i:- and that the vast majority of the monophasic isolates in livestock and humans in the UK since 2006 are not directly related to either the epidemic described in Spain around 1997 (11), or the epidemic described in the US around 2004 and 2007 (10). Instead, the UK epidemic is related to that reported in Germany and elsewhere since around 2005 (12). The US clone is characterized by a large deletion in the *fljB* locus and acquisition of a prophage, neither of which were present in the UK monophasic clone. Furthermore, whole genome sequence for a single isolate from the US epidemic (CVM23701), placed this isolate in a small clade of monophasic isolates from the UK isolated around 1995, distinct from the current UK clade. The Spanish clone is characterized by variable size deletions in the *fljB* locus, all distinct from deletions observed in the UK isolates, and a deletion in the allantoin metabolism locus, also absent from the main UK clade. The MRCA of the UK *S.* 1,4,[5],12:i:- epidemic in our strain collection was shared with a biphasic *S.* Typhimurium isolate with DT7 (strain H09164 0090), a relatively rare phage type that has not been associated with epidemics in the epidemiological record. The common ancestor with strain H091640090 likely existed in the recent past (~20 years), as only about 10 SNPs have accumulated in the genome since the lineages diverged, based on short term substitution rate (1-2 SNPs / per genome / per year) previously reported for *Salmonella* epidemics (26, 27).

Since virtually all monophasic strains from the current epidemic clade encoded SGI-3 but isolates from outside the clade did not, it is likely that initiation of clonal expansion was accompanied by the acquisition of this genomic island. SGI-3 encodes resistance to heavy metals including copper and zinc is potentially significant since these are common supplements in pig feed, as a micronutrient and a general antimicrobial (28). Indeed, in the EU heavy metals have been used increasingly in response to the ban on non-specific use of antibiotics in animal feed as a growth promoter (29). Heavy metals are concentrated in the pig intestine and this may represent a significant selective pressure contributing to the success of this clone. Indeed, a recent study reported that enhanced MIC (20-24 mM) compared with the baseline MIC (16mM) to copper sulphate was significantly more likely in isolates from pig feces (30).

A remarkable feature of the monophasic epidemic in the UK is the considerable number of polymorphisms that impact coding capacity that occurred during the short period (~10-15 years) of clonal expansion of the epidemic clade. They include a complex pattern of deletions in the *fljB* locus and surrounding genome sequence, insertions in the *thrW* locus, and acquisition of a novel phage carrying the *sopE* gene. These polymorphisms appear to be stable and not deleterious as they all appear in parts of the tree that have subsequently undergone further clonal expansion. Deletions in the *fljB* locus that occurred subsequent to the initial clonal expansion of the epidemic clade, accounted for the monophasic phenotype exhibited by most of these isolates. The high frequency of deletions in this locus may be the result of a composite Tn21-like transposable element that is inserted in the *hin* - *iroB* intergenic region, a well-known characteristic of such insertions (31).

The acquisition of the *sopE* gene on a novel prophage element that occurred through multiple recent independent events may be highly significant to the pathogenesis and epidemiology of the current epidemic. Lysogeny by phage carrying the *sopE* gene has been associated previously with epidemic strains of *S.* Typhimurium and of other *Salmonella* serotypes (32). The expression of SopE may increase the fitness of the pathogen a possibility consistent with the observation that recent acquisition of the *sopE* gene by monophasic epidemic isolates has been followed by an increase in the frequency with which *sopE* positive isolates have been isolated. The ability to induce inflammatory diarrhea is an important strategy for the transmission of *Salmonella* Typhimurium. SopE is a guanine exchange factor that activates both cdc42 and rac1 while *sopE2* only activates cdc42 (33). All *S.* Typhimurium strains sequenced to date encode the *SopE2* gene that exhibits 59% identity with SopE. The additional activity of *SopE* has a marked impact on the outcome of the interaction of *S.* Typhimurium with the intestinal mucosa, resulting in increased burden of *Salmonella* in the intestinal lumen and shedding in the faeces. *SopE* expression results in increased production of host nitrate, a valuable electron acceptor utilized by *S.* Typhimurium for respiration (34).

The current monophasic Typhimurium clone associated with many animal species and human clinical infections in the UK arose in the recent past and subsequent microevolution in a short period of time has resulted in considerable genotypic variation impacting important antigens, virulence factors and resistance loci. Some genomic features, such as resistance to heavy metals may have resulted in initial selection for the current clone, while more recent horizontal gene transfer or deletions and plasmid loss may be generating variation selected during the epidemic.

## Figure Legends

**Figure 1. Phylogeny of *S*. Typhimurium and *S.* 1,4,[5],12:i:- isolates.** Maximum likelihood tree of 212 *S*. Typhimurium and monophasic isolates constructed using 12793 SNPs outside of prophage elements, IS elements and sequence repeats identified by reference to the whole genome sequence of *S*. Typhimurium strain SL1344. The tree is rooted with *S*. Enteritidis whole genome sequence as an outgroup (note shown). The lineage containing the *Salmonella* 1,4,[5],12:i:- current UK epidemic group is conflated for simplicity (filled triangle). * Monophasic isolates outside of the main epidemic clade, # monophasic clade closely related to the North American monophasic clone CVM23701 (10). The designation of the isolates (left column) and phage type are shown (right column), ND, not determined. The bar indicates the approximate number of SNPs determined by genetic distance and the number of SNPs used to construct the tree.

**Figure 2. Phylogeny of *S.* 1,4,[5],12:i:- epidemic clade isolates.** Maximum likelihood tree of 77 *Salmonella* 1,4,[5],12:i:- isolates rooted with *S*. Typhimurium strain SL1344, constructed using 1058 SNPs outside of prophage elements, IS elements and sequence repeats identified with reference to whole genome sequence of *S*. Typhimurium strain SO4698-09. Bootstrap values are indicated at nodes where less than 70. Subclades A (blue lineages), B (red lineages) and C (green lineages) indicated. Strain designations are colour coded for human isolates (red type) and animal isolates (blue type). Epidemiological data for the source of isolate, phage type, Country of origin, presence of the virulence plasmid (pSLT), presence of the *sopE* gene, occupancy of the *thrW* locus and the presence of *Salmonella* genetic island-3 are indicated (right). The scale

bar indicates the approximate number of SNPs determined by genetic distance and the number of SNPs used to construct the tree.

**Figure 3. Frequency of carriage of the *sopE* gene in *S.* 1,4,[5],12:i:- epidemic isolates for each year 2005-2010.** The presence of the *sopE* gene was detected in draft genome assemblies by sequence comparison or by PCR amplification of genomic DNA using primers specific for the *sopE* gene of randomly selected monophasic isolates from each year. The number of isolates investigated for each year is indicated above the bar.

**Supplementary Figure 1. Presence of antibiotic resistance genes in the monophasic Typhimurium epidemic strains from the UK.** The presence (red) or absence (blue) of antibiotic resistance genes are shown in the context of the maximum likelihood tree described in Figure 2. Data unavailable due to poor quality sequence assembly (black).

**Supplementary Figure 2. Gene arrangement of the novel genomic island of *S.* 1,4,[5],12:i:- strain SO4698-09.** Arrows indicate predicted genes within the island. The position of genes with predicted functions by sequence comparison are indicated for arsenic resistance (red), cadmium, zinc and copper resistance (green). The nucleotide sequence flanking the insertion in the whole genome sequence of SO4698-09 (PRJEB10340) is indicated.

**Supplementary Figure 3. Ancestral state reconstruction of SGI-3 and sopE gene within the monophasic epidemic clade.** Maximum likelihood trees for 77 UK and

Italy monophasic isolates as previously described in Figure 2. Ancestral state for presence (red edges) or absence (blue edges) of SGI-3 (A) or *sopE* (B) were reconstructed based on maximum parsimony using ACCTRAN. * indicate the inferred acquisition of the genetic element.

**Supplementary Figure 4. Minimum inhibitory concentration of monophasic Typhimurium and Typhimurium isolates to copper sulphate in rich broth culture.** The ability of monophasic Typhimurium (filled circles) or Typhimurium (filled squares) isolates to grow in Luria Bertani broth in the presence of copper sulfate (pH7) were monitored by the optical density of culture. The MIC was defined as the concentration at which cultures attained at least $OD_{600nm}$ of 0.1. The mean for each phylogenetic group (grey bar) +/- standard deviation are indicated. Student's t test was used to test significance.

**Supplementary Figure 5. Heat map showing deletions around the *fljB* locus of the *S.* 1,4,[5],12:i:- epidemic clade isolates.** The heat map (A) indicating mapped sequence read coverage for *S.* 1,4,[5],12:i:- epidemic clade isolates to the *fljB* locus and flanking sequence of the whole genome sequence of *S.* Typhimurium strain SL1344. Color indicates 0 mapped reads (blue) to ≥20 bases (red). Filled arrows indicate genes in the SL1344 genome sequence as described previously (35). A maximum likelihood tree of phenotypically monophasic isolates from the strain collection is shown.

**Supplementary Figure 6. Prophage element mTmV from strain SO4698-09 and BLAST results with SfV and FELS-2 prophage.** Predicted open reading frames in

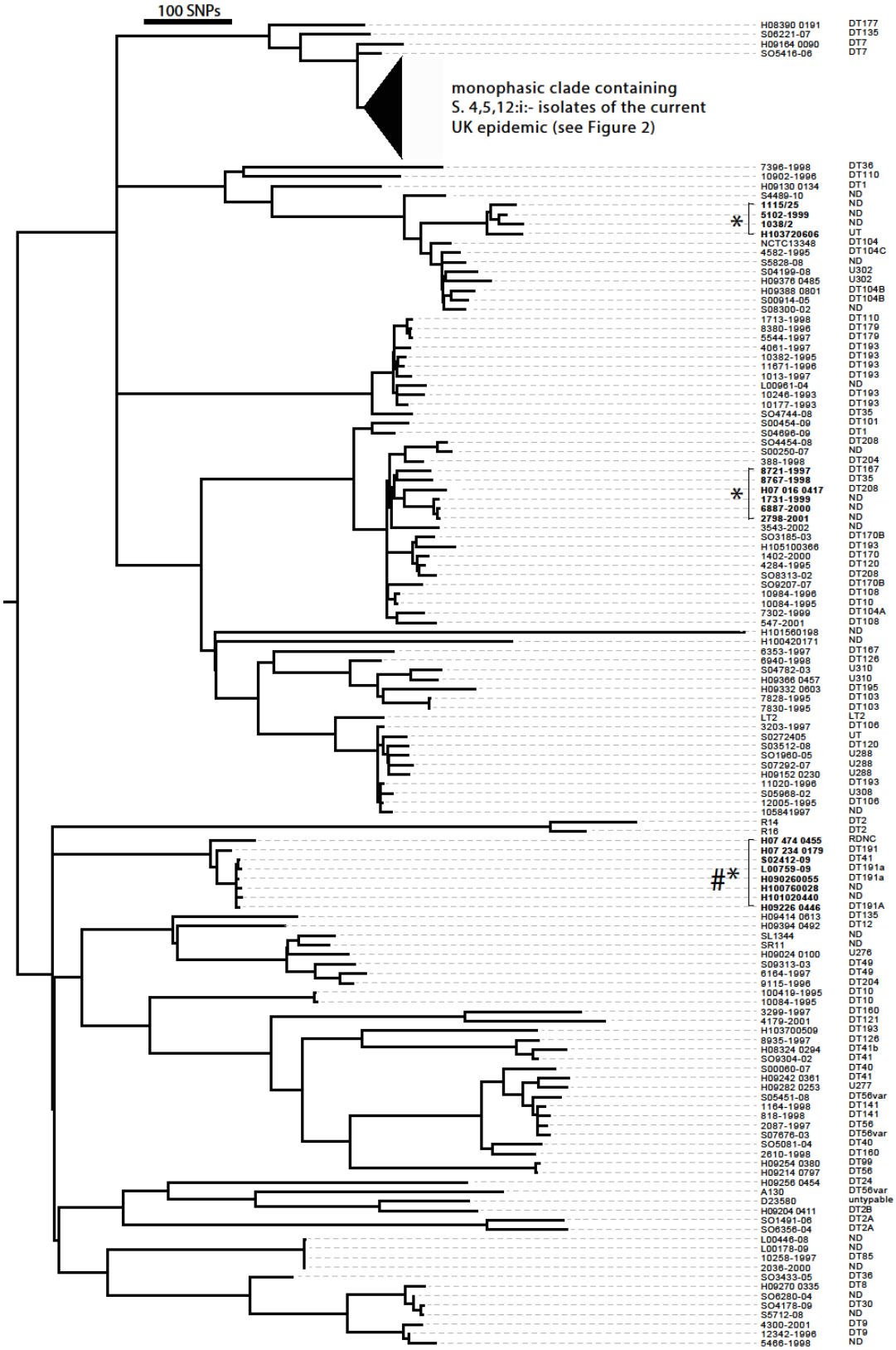
408 the 55 kb mTmV prophage of strain SO4698 are shown with flanking nucleotide
409 sequence for orientation. Regions with significant BLAST results (red bar) in the
410 related prophage SfV prophage and FELS-2 prophages are indicated below.
411
412

## References


1. Herikstad H, Motarjemi Y, Tauxe RV. *Salmonella* surveillance: a global survey of public health serotyping. Epidemiology and infection. 2002;129(1):1-8.
2. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clin Infect Dis. 2010 Mar 15;50(6):882-9.
3. Wales A, Davies RH. Environmental aspects of *Salmonella*. In: Barrow PA, Methner U, editors. *Salmonella* in domestic animals. 2 ed: CABI; 2013. p. 399-425.
4. Rabsch W, Truepschuch S, Windhorst D, Gerlach RG. Typing phages and prophages of Salmonella. In: Porwollik S, editor. *Salmonella*, from genome to Function. Norfolk, UK: Caister Academic Press; 2011. p. 25-48.
5. Rabsch W, Truepschuch S, Windhorst D, Gerlach RG. Typing phages and prophages of *Salmonella*. Norfolk, UK: Caister Academic Press; 2011.
6. Rabsch W, Tschape H, Baumler AJ. Non-typhoidal salmonellosis: emerging problems. Microbes Infect. 2001 Mar;3(3):237-47.
7. Threlfall EJ. Epidemic *Salmonella typhimurium* DT 104 - a truly international multiresistant clone. J Antimicrob Chemoth. 2000 Jul;46(1):7-10.
8. Threlfall EJ, Frost JA, Ward LR, Rowe B. Epidemic in cattle and humans of *Salmonella typhimurium* DT104 with chromosomally integrated multiple drug resistance. Veterinary Record. 1994;134:577.
9. Agency VL. *Salmonella* in livestock production in GB—2011 report 2014 [cited; Available from: http://vla.defra.gov.uk/reports/rep_salm_rep11.htm
10. Soyer Y, Switt AM, Davis MA, Maurer J, McDonough PL, Schoonmaker-Bopp DJ, et al. *Salmonella enterica* Serotype 4,5,12:i:-, an Emerging *Salmonella* Serotype That Represents Multiple Distinct Clones. Journal of Clinical Microbiology. 2009 Nov;47(11):3546-56.
11. Laorden L, Herrera-Leon S, Martinez I, Sanchez A, Kromidas L, Bikandi J, et al. Genetic evolution of the Spanish multidrug-resistant *Salmonella enterica* 4,5,12:i:- monophasic variant. J Clin Microbiol. 2010 Dec;48(12):4563-6.
12. Trupschuch S, Laverde Gomez JA, Ediberidze I, Flieger A, Rabsch W. Characterisation of multidrug-resistant *Salmonella* Typhimurium 4,[5],12:i:- DT193 strains carrying a novel genomic island adjacent to the thrW tRNA locus. International journal of medical microbiology : IJMM. 2010 Jun;300(5):279-88.
13. de la Torre E, Zapata D, Tello M, Mejia W, Frias N, Garcia Pena FJ, et al. Several Salmonella enterica subsp. enterica serotype 4,5,12:i:- phage types isolated from swine samples originate from serotype *typhimurium* DT U302. J Clin Microbiol. 2003 Jun;41(6):2395-400.
14. Mossong J, Marques P, Ragimbeau C, Huberty-Krau P, Losch S, Meyer G, et al. Outbreaks of monophasic Salmonella enterica serovar 4,[5],12:i:- in Luxembourg, 2006. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2007 Jun;12(6):E11-2.
15. Hauser E, Tietze E, Helmuth R, Junker E, Blank K, Prager R, et al. Pork contaminated with *Salmonella enterica* serovar 4,[5],12:i:-, an emerging health risk for humans. Appl Environ Microbiol. 2010 Jul;76(14):4601-10.
16. Barone L, Dal VA, Pellissier N, Vigano A, Romani C, Pontello M. [Emergence of *Salmonella* Typhimurium monophasic serovar: determinants of

antimicrobial resistance in porcine and human strains]. Annali di igiene : medicina preventiva e di comunita. 2008 May-Jun;20(3):199-209.
17. DEFRA. *Salmonella* in Livestock Production in GB. DEFRA; 2013.
18. Switt AIM, Soyer Y, Warnick LD, Wiedmann M. Emergence, Distribution, and Molecular and Phenotypic Characteristics of Salmonella enterica Serotype 4,5,12:i:-. Foodborne Pathog Dis. 2009 May;6(4):407-15.
19. Anderson ES, Ward LR, Saxe MJ, de Sa JD. Bacteriophage-typing designations of Salmonela typhimurium. The Journal of hygiene. 1977;78(2):297-300.
20. Anonymous. http://bsac.org.uk/. 2010 [cited 2014 September]
21. Anonymous. Scientific Opinion on monitoring and assessment of the public health risk of "*Salmonella* Typhimurium-like" strains. EFSA Journal. 2010;8(10):1826.
22. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015 Jul 20.
23. Wood MW, Rosqvist R, Mullan PB, Edwards MH, Galyov EE. SopE, a secreted protein of *Salmonella dublin*, is translocated into the target eukaryotic cell via a *sip*-dependent mechanism and promotes bacterial entry. Mol Microbiol. 1996;22:327-38.
24. Hardt WD, Chen LM, Schuebel KE, Bustelo XR, Galan JE. *S. typhimurium* encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. Cell. 1998;93(5):815-26.
25. Allison GE, Angeles D, Tran-Dinh N, Verma NK. Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of *Shigella flexneri*. J Bacteriol. 2002 Apr;184(7):1974-87.
26. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, et al. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. Nature genetics. 2012 Nov;44(11):1215-21.
27. Mather AE, Reid SWJ, Maskell DJ, Parkhill J, Fookes MC, Harris SR, et al. Distinguishable Epidemics of Multidrug-Resistant *Salmonella* Typhimurium DT104 in Different Hosts. Science. 2013 Sep 27;341(6153):1514-7.
28. Nicholson FA, Chambers BJ, Williams JR, Unwin RJ. Heavy metal contents of livestock feeds and animal manures in England and Wales. Bioresource Technol. 1999 Oct;70(1):23-31.
29. Slade RD, Kyriazakis I, Carroll SM, Reynolds FH, Wellock IJ, Broom LJ, et al. Effect of rearing environment and dietary zinc oxide on the response of group-housed weaned pigs to enterotoxigenic Escherichia coli O149 challenge. Animal : an international journal of animal bioscience. 2011 Jun;5(8):1170-8.
30. Medardus JJ, Molla BZ, Nicol M, Morrow WM, Rajala-Schultz PJ, Kazwala R, et al. In-feed use of heavy metal micronutrients in U.S. swine production systems and its role in persistence of multidrug-resistant salmonellae. Appl Environ Microbiol. 2014 Apr;80(7):2317-25.
31. Hughes KT, Roth JR. Directed formation of deletions and duplications using Mud(Ap, lac). Genetics. 1985 Feb;109(2):263-82.
32. Hopkins KL, Threlfall EJ. Frequency and polymorphism of sopE in isolates of Salmonella enterica belonging to the ten most prevalent serotypes in England and Wales. J Med Microbiol. 2004 Jun;53(Pt 6):539-43.
33. Friebel A, Ilchmann H, Aepfelbacher M, Ehrbar K, Machleidt W, Hardt WD. SopE and SopE2 from *Salmonella* typhimurium activate different sets of RhoGTPases of the host cell. J Biol Chem. 2001 Sep 7;276(36):34035-40.

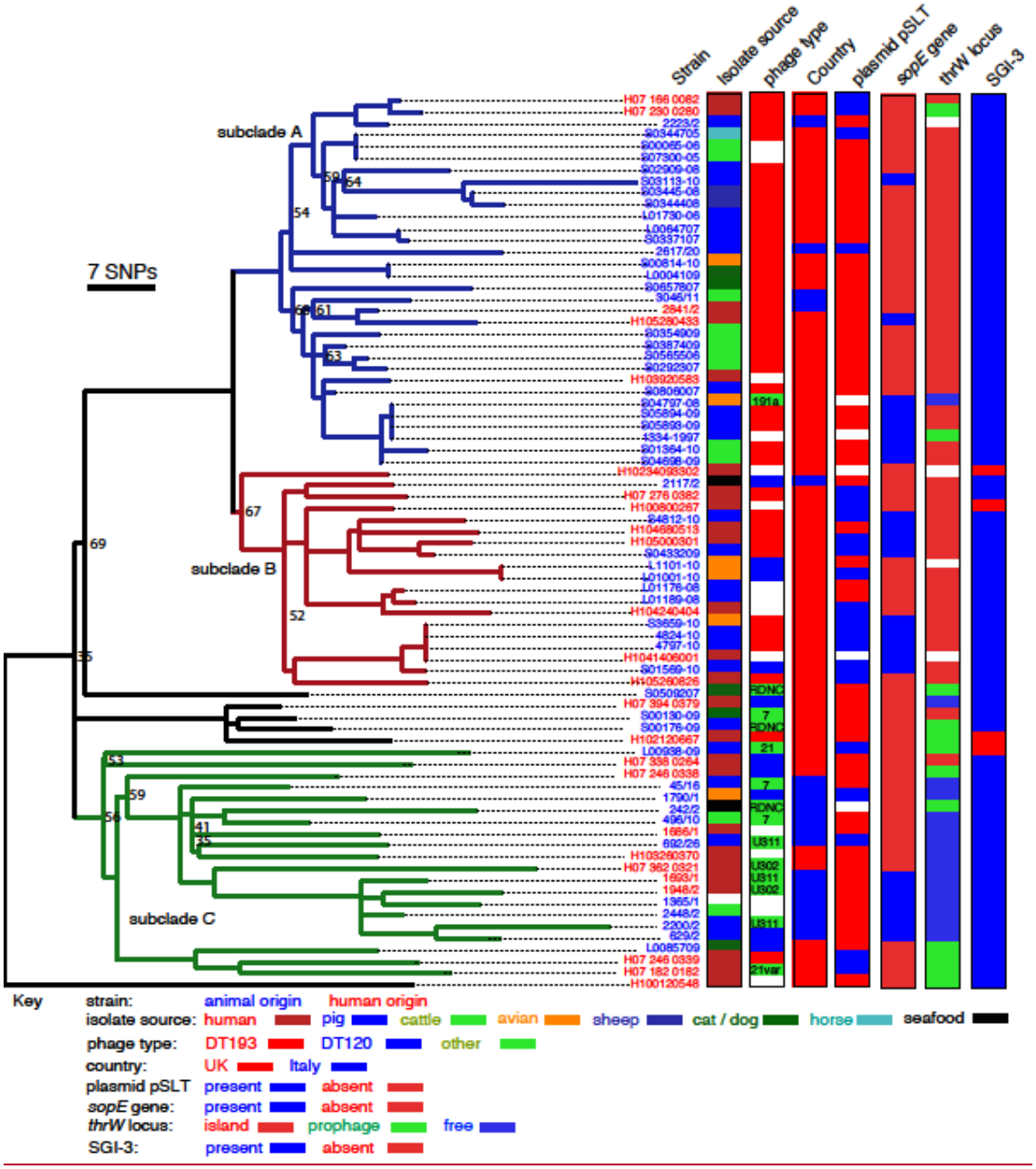
510 34. Lopez CA, Winter SE, Rivera-Chavez F, Xavier MN, Poon V, Nuccio SP, et
511 al. Phage-mediated acquisition of a type III secreted effector protein boosts growth of
512 *Salmonella* by nitrate respiration. mBio. 2012;3(3).
513 35. Kroger C, Dillon SC, Cameron AD, Papenfort K, Sivasankaran SK, Hokamp
514 K, et al. The transcriptional landscape and small RNAs of *Salmonella* enterica serovar
515 Typhimurium. Proc Natl Acad Sci U S A. 2012 May 15;109(20):E1277-86.
516
517

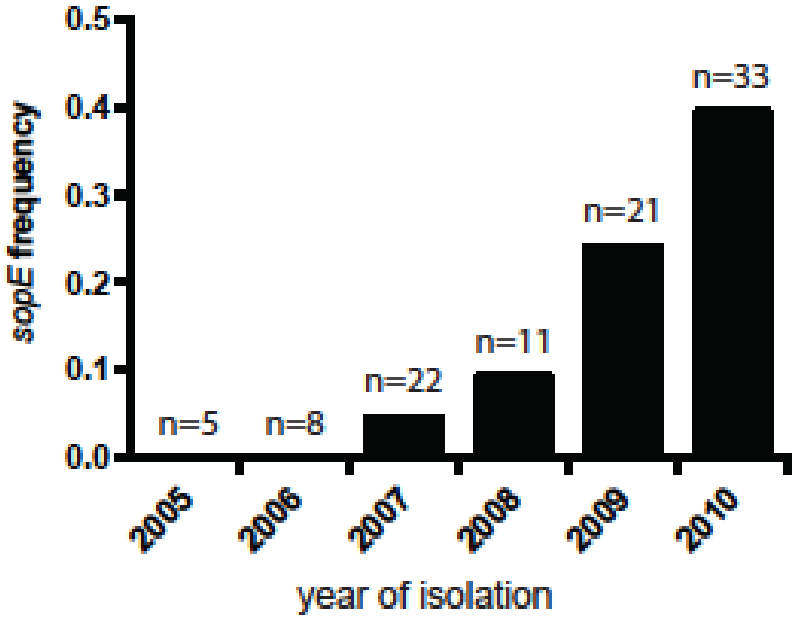

100 SNPs
monophasic clade containing
S. 4,5,12:i:- isolates of the current
UK epidemic (see Figure 2)
*
*
#*

518

519

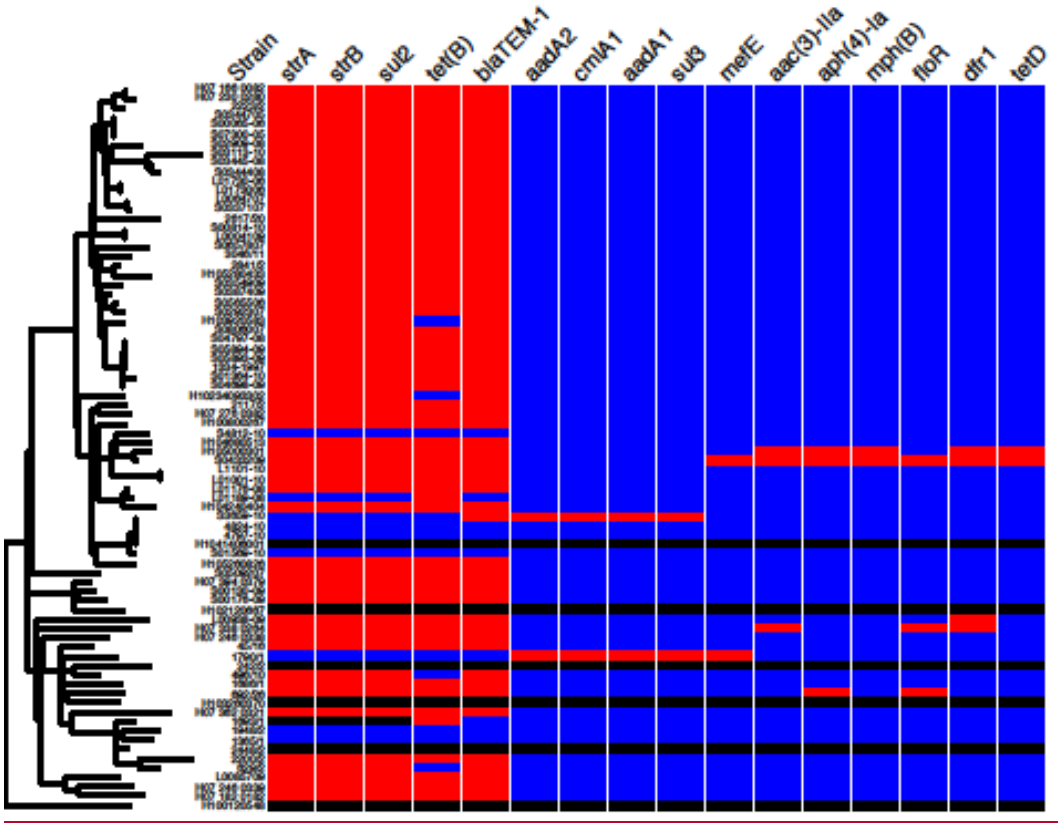
520


Strain
isolate source
phage type
Country
plasmid pSLT
sopE gene
thrW locus
SGI-3
7 SNPs
subclade A
subclade B
subclade C
Key
strain: animal origin human origin
isolate source: human pig cattle avian sheep cat / dog horse seafood
phage type: DT193 DT120 other
country: UK Italy
plasmid pSLT present absent
sopE gene: present absent
thrW locus: island prophage free
SGI-3: present absent


0.5
0.4
0.3
0.2
0.1
0.0
sopE freq uency
n=5
n=8
n=22
n=11
n=21
n=33
2005
2006
2007
2008
2009
2010
year of isolation

524

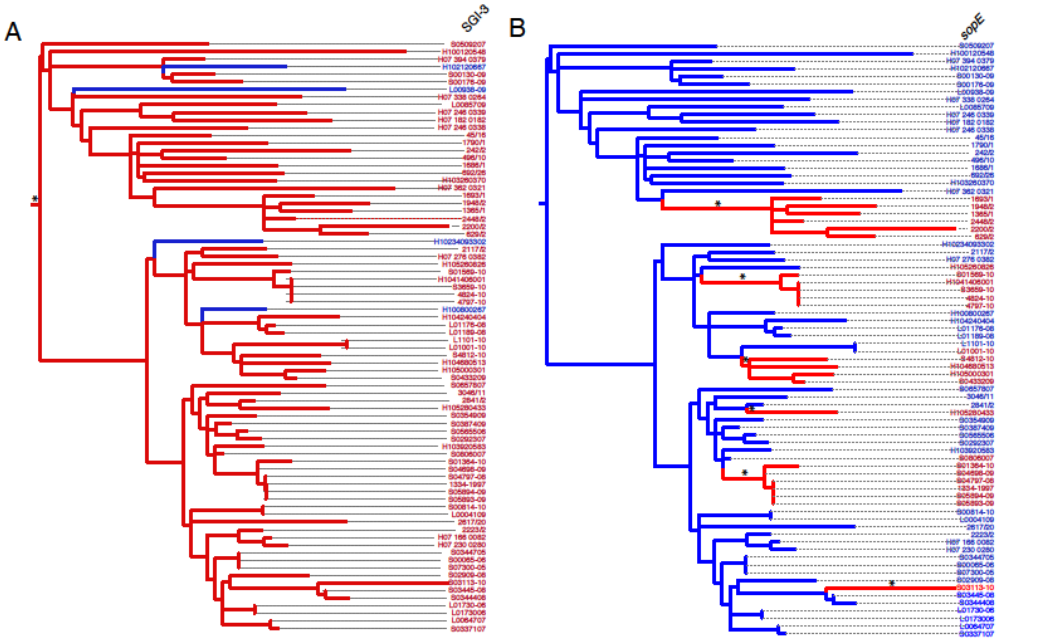
525


Strain
strA
strB
sul2
tet(B)
blaTEM-1
aadA2
cmlA1
aadA1
sul3
mefE
aac(3)-IIa
aph(4)-Ia
mph(B)
floR
dfr1
tetD

526

527

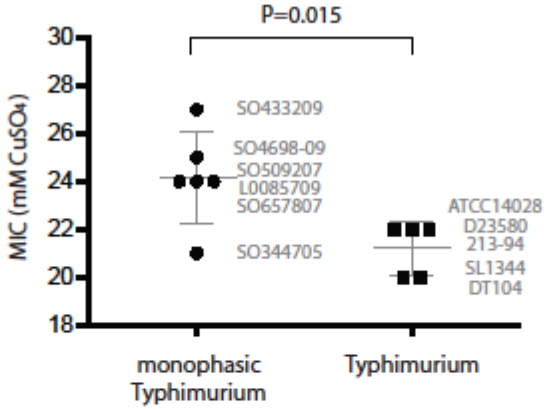
528




529

530

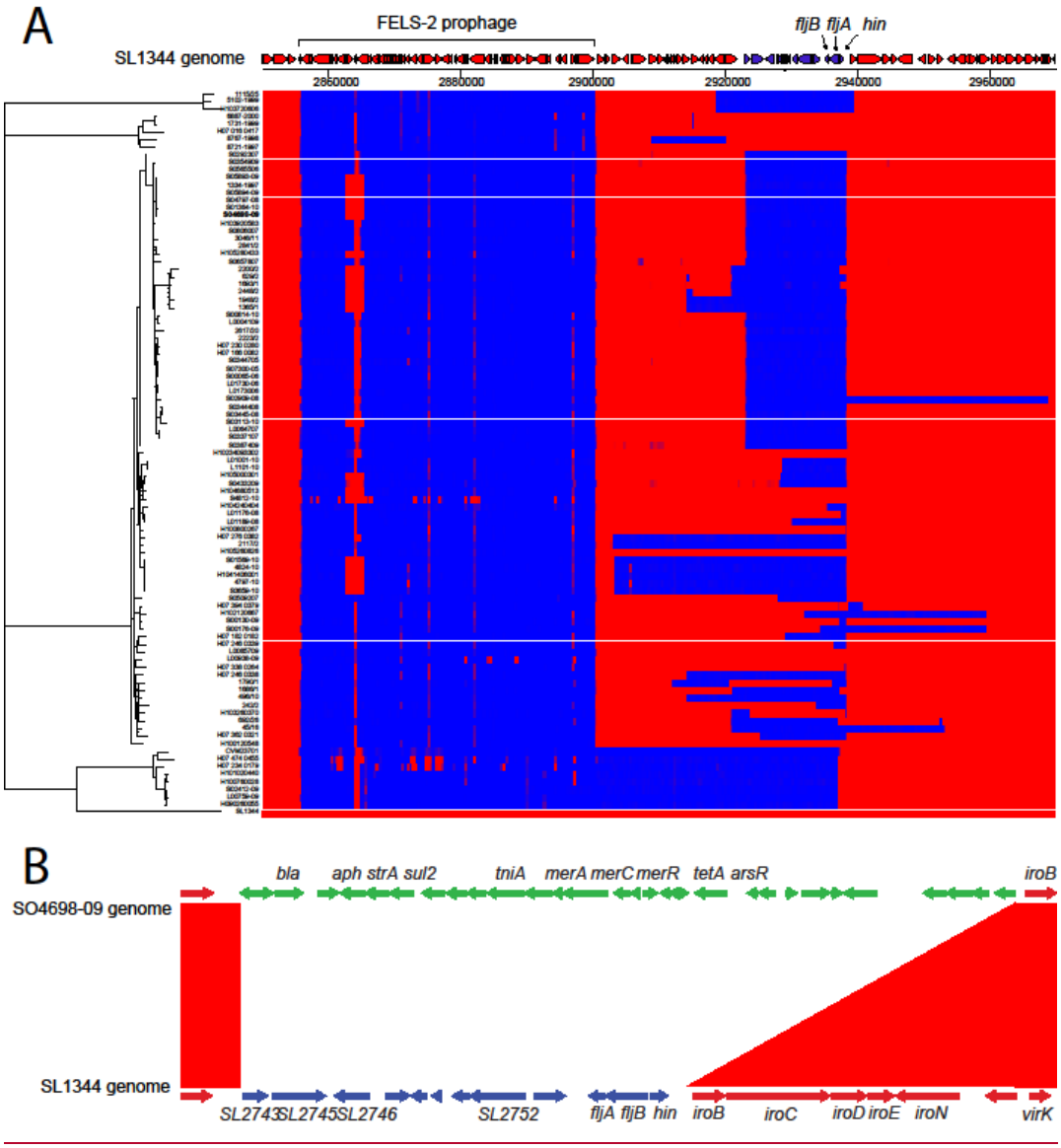
531


P=0.015
MIC (mM CuSO4)
30
28
26
24
22
20
18
SO433209
SO4698-09
SO509207
L0085709
SO657807
SO344705
ATCC14028
D23580
213-94
SL1344
DT104
monophasic Typhimurium
Typhimurium

532

533

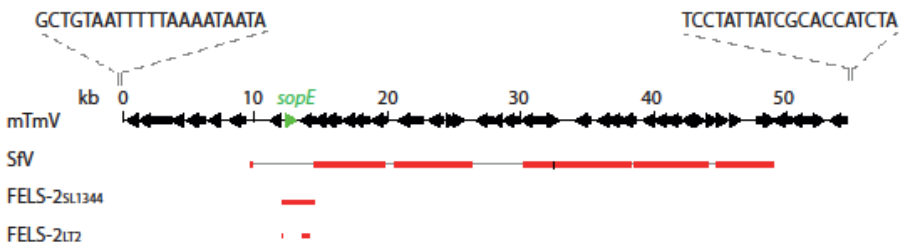
534


A
FELS-2 prophage
fljB fljA hin
SL1344 genome
2860000
2880000
2900000
2920000
2940000
2960000
B
bla
aph strA sul2
tniA
merA merC merR
tetA arsR
iroB
SO4698-09 genome
SL1344 genome
SL2743 SL2745 SL2746
SL2752
fljA fljB hin
iroB
iroC
iroD iroE
iroN
virK

535

536

537


GCTGTAATTTTTAAAATAATA
TCCTATTATCGCACCATCTA
kb
0
10
sopE
20
30
40
50
mTmV
SfV
FELS-2SL1344
FELS-2LT2

538

539