Microevolution of monophasic variant of epidemic *Salmonella* Typhimurium (*S*. 4,[5],12:i:-)

Eleonora Tassinari

Thesis for the degree of Doctor of Philosophy

University of East Anglia

Quadram Institute Bioscience

September 2019

Abstract

Salmonella Typhimurium is a major cause of food-borne gastroenteritis. In 2005 a monophasic variant of *S*. Typhimurium (*S*. 4,[5],12;i:-) sequence type ST34 and ASSuT antimicrobial resistance profile (AMR) emerged as a worldwide epidemic. Pigs are the main reservoir of this pathogen, pork and pork products represent the predominant vehicles of *S*. 4,[5],12;i:-.

In this study, whole-genome sequencing was used to investigate the population structure and the persistence of *S*. Typhimurium and *S*. 4,[5],12;i:- ST34 on Irish pig farms. Additionally, the ability of the isolates to produce biofilms was assessed as this represents an important strategy for environmental survival and persistence. Farm-specific clones persisting between production cycles were identified. No unequivocal association between biofilm formation and persistence was observed; however, epidemic clones formed more biofilms than non-epidemic genotypes suggesting the advantage conferred by this trait. During persistence on farm, microevolution enhancing biofilm formation, but also AMR and heavy metal resistance was observed following the acquisition of IncHI2 plasmids.

The phage-encoded virulence factor SopE was sporadically described in *S*. Typhimurium but was associated with epidemic types. *sopE* was acquired by *S*. 4,[5],12;i:- ST34 following lysogenic conversion by the novel prophage mTmV. The findings that the acquisition of *sopE* occurred multiple, independent times in *S*. 4,[5],12;i:- ST34 and was concomitant with the emergence of this pathogen in pigs suggested that the acquisition of *sopE* was positively selected and may have contributed to the success of *S*. 4,[5],12:i:- ST34. mTmV was virtually absent in *S*. Typhimurium serovar types despite its ability to infect a variety of *S*. Typhimurium genotypes, but disseminated to diverse *S*. *enterica* serovars, potentially contributing to the emergence of novel pathovariants. Variants of mTmV and novel bacteriophages carrying *sopE*, which were identified in association with *sopE* copy number variation, likely promote the dissemination of *sopE*.

Table of Contents

List of Figures

List of Tables

List of Appendices

Accompanying Material

Publication:

Tassinari E., Duffy G., Bawn M., Burgess C.M., McCabe E.M., Lawlor P.G., Gardiner G., Kingsley R.A., Microevolution of antimicrobial resistance and biofilm formation of *Salmonella* Typhimurium during persistence on pig farms,
Scientific Reports (2019), (9),1.
DOI: 10.1038/s41598-019-45216-w
Based on Chapter 3

The study was conceived by G. Duffy and R.A. Kingsley, the experiments were performed by E. Tassinari, data analysis and expertise were provided by E. Tassinari, G. Duffy, M. Bawn, C.M. Burgess, E.M. McCabe and R.A. Kinglsley, material and expertise of samples was provided by P.G. Lawlor and G. Gardiner.

Acknowledgements

First, I would like to thank Rob for giving me the opportunity, together with Geraldine, to carry out this PhD, and I would like to express him my sincere gratitude for his constant guidance during these years, for his teachings and for being a supportive supervisor. I am also feeling grateful to him for the assistance and encouragement he provided me throughout the production of this thesis.

I would like then to thank Geraldine and Kaye for their supervision and support, for their contribution the research project and for taking the time to travel to Norwich.

A thanks to Evonne who has been an important guidance during the very beginning of my journey, always available with her support inside and outside the lab. It has been a pleasure working with you and getting to know you as a person.

A huge thank you then goes to Matt, who had the hard task of teaching me the bases of bioinformatics, I like to think that, despite the countless times that I "broke the cluster", eventually he was partially successful. I really appreciated your help and encouragement as well as your considerable contribution to the project, it has been stimulating working with someone as creative as you.

I am also grateful to all the colleagues in the Kingsley group, past and present, with a special mention to Priscilla, Oli, Mark, Jen and Gaetan, who created a supportive work environment and a nice atmosphere in the lab and in the office.

I would like then to say thank you to my companions in adventure, the other postgraduate students who I met in Teagasc (especially to Elena and Juan) and at Quadram (Marina, Dan, Dim, Ann-Marie) and with who I shared the good and the bad in the past 4 years (as well as a considerable amount of pints). A special mention for Erika who has been a valuable friend and a source of laughs and hugs here in Norwich.

I would like to express my infinite gratitude to my family, in particular to my parents who gave me their unconditional support and love across these years, but also to my siblings Fabrizio and Elena, my adorable nephews Francio and Richi, and my brother-in-law Massimo for their encouragement. How I could not thank my aunt Emanuela, who showed me all her love by cooking terrific dinners every time I was coming back home, and Anna for her precious friendship. And a final thanks to a special person for bravely entering my life in the past few months bringing

joy and a good dose of fun, making this period so enjoyable despite the struggles of the writing.

Abbreviations

AMR: Antimicrobial resistance

bp: base pair

DT: definitive type

ECM: extra-cellular matrix

iNTS: invasive non-typhoidal *Salmonella*

NTS: non-typhoidal *Salmonella*

MDR: multidrug resistance

ML: maximum-likelihood

SGI-4: *Salmonella* genomic island 4

SNP: single-nucleotide polymorphism

ST: sequence type

WT: wild type

WGS: whole-genome sequencing

1 Literature review

1.1 Introduction

1.1.1 *Salmonella* genus

Salmonella genus is constituted by Gram-negative, rod-shaped and non-spore-forming bacteria.

The genus contains only two species: *Salmonella enterica* and *Salmonella bongori*. *S. enterica* comprises six subspecies, which can be designated either with their taxonomic name or with Roman numerals: *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtenae* (IV) and *indica* (VI).

S. *enterica* subsp. *enterica* accounts for almost the totality of human infections (Aleksic et al., 1996).

Within the subspecies, serovars are distinguished based on antigenic and biochemical characteristics; subspecies *enterica* currently comprises more than 1,500 serovars (Grimont and Weill, 2007).

Salmonella adapted to live in the intestinal tract of animals. Based on the host range, *Salmonella* bacteria can be classified as generalists, meaning they are able to circulate in multiple animal species, or host-adapted, namely strongly associate with a single animal species. Infections with generalist and host-adapted serovars often have different outcomes of infection, ranging from gastroenteritis to systemic infections (Baumler and Fang, 2013).

Salmonella classification of serovars is performed worldwide accordingly to the White-Kauffman scheme, which is constantly updated by the WHO Collaborating Centre for Reference and Research on *Salmonella* (Grimont and Weill, 2007).The White-Kauffman and Le Minor classification is based on the analysis of three antigens: the somatic antigen (O antigen), the flagellar antigen (H antigen) and the virulence antigen (Vi antigen), which enable to type *Salmonella* up to a serovar level (Grimont and Weill, 2007).

The somatic antigen is determined based on the glucidic portion of the lipopolysaccharide (LPS). The flagellar antigen includes both phase 1 and phase 2 flagellin.

While all *Salmonella* subspecies harbour phase 1 flagellin encoded by *fliC*, *S*. enterica subsp. I, II, IIIb and VI acquired a second phase flagellin locus (*FljABhin*)

through horizontal gene transfer becoming biphasic (Figure 1.1) (Porwollik et al., 2002).

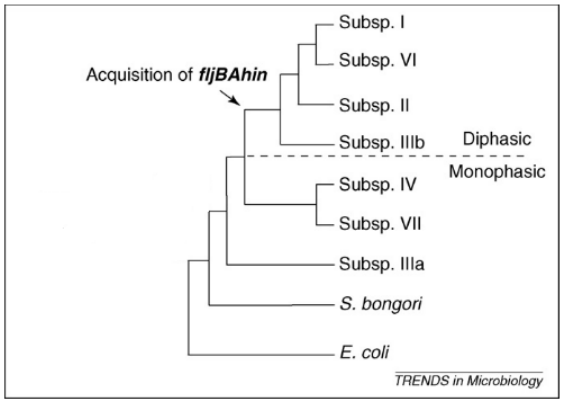

Figure 1.1 Acquisition of phase 2 flagellin during evolution of *Salmonella enterica* subspecies. The acquisition of the *fljABhin* operon is shown within the phylogenetic context of *S*. *enterica* subspecies. Figure adapted from (McQuiston et al., 2008).

Phase 1 and 2 flagellins, which induce different humoral responses, are alternatively expressed through a reversible process termed phase variation mediated by the Hin recombinase. The activation of the *fljAB* locus depends on the orientation of the promoter which is periodically inverted following recombination of the flanking inverted repeated sequences *hixL* and *hixR* by the Hin recombinase. The *fljAB* operon encodes a repressor of *fliC* (FljA) and the phase 2 flagellin, FljB. When the promoter is in the "off" orientation, FliC is expressed. When the promoter is turned "on" FljA repress *fliC* and FljB is expressed (Figure 1.2). Alternation between the two phases occurs with a frequency of 10^{-3} - 10^{-5} per cell generation in *S. enterica* (Stocker, 1949, Gillen and Hughes, 1991).

Nevertheless, some biphasic serovars have lost one or both the flagellin antigens becoming monophasic, such as most of *S*. Typhi, or non-flagellated, such as *S*. Gallinarum and *S*. Pullorum. Although phase 1 and 2 flagellins are chromosomally encoded, rare examples of plasmid mediated acquisition of flagellin genes were reported. For instance, while the majority of *S*. Typhi only harbour *fliC* encoding the H:d antigen or a *fliC* variant deriving from an in-frame deletion (H:j antigen) reported in some *S*. Typhi isolates from Indonesia, some *S*. Typhi from this country have acquired a phase 2 flagellin (H:z66 antigen) encoded on the linear plasmid pBSSB1 (Baker et al., 2007a, Baker et al., 2007b).

The three distinct forms of flagellin of *S*. Typhi influence the motility of the microorganisms and their interactions with the host in relationship, such as the immune response and invasiveness of the host cells (Schreiber et al., 2015). The virulence antigen is a capsular polysaccharide expressed by only some strains belonging to *S*. Typhi, *S*. Paratyphi C and *S.* Dublin serovars (Grimont and Weill, 2007).

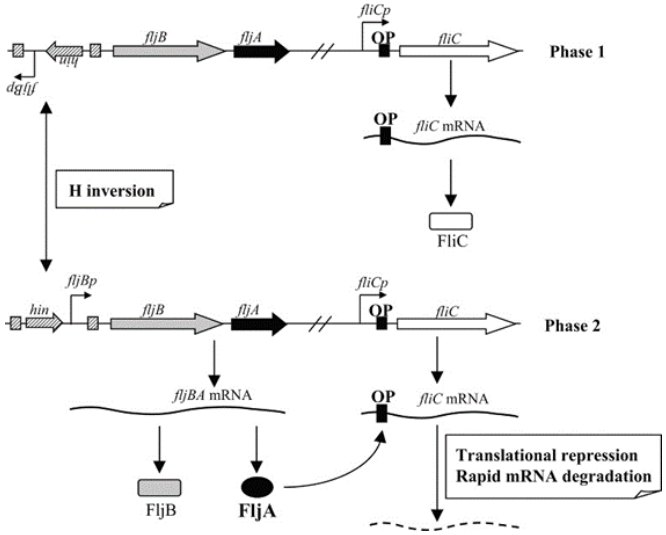

Figure 1.2 Mechanism of phase variation of the flagellar antigens FliC and FljB in *Salmonella* (Yamamoto and Kutsukake, 2006). When the promoter of the *fljAB* operon is in "off" orientation, fliC is expressed. Following the inversion of the orientation of the *fljAB* promoter by the Hin DNA invertase, *fljAB* is expressed resulting in the production of the phase 2 flagellin, FljB, and in FljA, which represses FliC expression.

1.1.1.1 Methods for detection and typing/subtyping of *Salmonella*

Typing consists in the characterisation of microbial isolates beyond species/subspecies level with the aim to discriminate related isolates. Typing methods may rely on the analysis of phenotypic characteristics, such as serotyping, phage typing, antimicrobial resistance and biochemical tests, or on molecular features.

Historically, phenotypic methods were the first to be developed for outbreak investigation and surveillance purposes. Despite their limited resolution, phenotypic typing methods are still widely used for the initial characterisation of *Salmonella* isolates (EFSA, 2013).

A greater discriminatory power can be reached with molecular typing methods which discriminate the bacteria based on their genotypic characteristics (Foley et al., 2007). The mechanisms of discrimination underlying molecular methods can

consist of restriction analysis of the DNA, amplification of target DNA or in the analysis of specific DNA polymorphisms (Foley, Lynne, & Nayak, 2009). Due to their ability to fingerprint isolates in a more sensitive way, subtyping methods are suitable for a broad range of applications, such as evolution and phylogenetic analysis, early outbreaks detection, identification of contamination source and dissemination route, bacteria tracking and risk assessment (EFSA, 2014).

1.1.1.1.1 Methods for *Salmonella* isolation

Methods for the isolation of *Salmonella* from food, animal feed or primary production developed by the International Organization for Standardization and accepted as European standards (EN ISO) are available and described in the EU legislation for the official monitoring in these matrices (EFSA, 2010b). Despite some variation, all the EN ISO protocols for *Salmonella* isolation comprise a pre-enrichment step in non-selective broth, followed by selective enrichment in one/two selective media, isolation in one/two agar media and confirmation (EFSA, 2010b). Biochemical tests were traditionally performed to check for *Salmonella* isolation (EFSA, 2010b). Once the genus was confirmed, further subtyping analyses are performed (EFSA, 2010b).

1.1.1.1.2 Phenotypic typing methods

1.1.1.1.2.1 Phage typing

The susceptibility to a set of bacteriophages, defined by the ability of bacteriophages to infect and induce cell lysis, can be used as a phenotypic feature to further classify *Salmonella* (EFSA, 2010b). Phage typing schemes have been developed for the characterisation of *S. enterica* serovars Typhi, Enteritidis and Typhimurium (Rabsch, 2007). Among the several phage typing schemes developed for *S*. Typhimurium, the one described by Callow and Anderson represents the gold standard and is performed globally (Callow, 1959, Anderson et al., 1977, Rabsch, 2007). Phage type is determined by the pattern of lysis exhibited by a pure culture to a defined pool of bacteriophages. Several mechanisms affect the ability of bacteriophage to invade the host cells and induce the lysis, such as presence/absence of specific surface proteins used as receptor, superinfection mechanisms by integrated prophages and bacterial several immunity systems (i.e. CRISPR-Cas) (Rabsch, 2007). Currently, more than 300 definitive *S*. Typhimurium phage types are defined; they are designated by DT (definitive type) or PT (provisional type) followed by a number (Anderson et al.,

1977). Historically, phage typing was useful for the identification of outbreak sources and epidemiological surveillance (Branchu et al., 2018). However, the technique is labour-intensive and requires trained personnel. Additionally, it does not enable an assessment of the phylogenetic relationships of the isolates.

1.1.1.1.3 Molecular typing methods

1.1.1.1.3.1 PCR

PCR is a powerful tool that can be used for diagnostic purposes, to detect *Salmonella* in contaminated food and environmental matrices, allowing a more timely detection compared to the conventional methods (Whyte et al., 2002). In particular, real time PCR is particularly suitable thanks to its rapidity, simplicity, high sensitivity and lower risk of contamination compared to conventional PCR (Adzitey et al., 2013).

Several PCR assays for *Salmonella* based on the analysis of different target genes have been developed (Cohen et al., 1996, Malorny et al., 2004, Pathmanathan et al., 2003, Ellingson et al., 2004). Depending on the gene target, *Salmonella* identification can be performed at a genus or species/subspecies level.

PCR is also routinely used for the confirmation of monophasic variants of *S*. Typhimurium, as recommended by the European Food Safety Authority (EFSA, 2010b). The PCR protocol simultaneously targets the *fljB* gene and the *S*. Typhimurium-specific intergenic region *fliB-fliA* which enables the discrimination of *S*. Typhimurium and the monophasic variants from other serovars with the same O and H antigens (EFSA, 2010b).

1.1.1.1.3.2 Pulse-field gel electrophoresis (PFGE)

PFGE involves the fragmentation of the genomic DNA with specific restriction enzymes followed by the separation of large DNA fragments (20 - 800 kb) achieved by switching the polarity of the current at regular intervals (Foley et al., 2007). The resulting banding pattern enables the fingerprinting of the isolates (Foley et al., 2007). The technique was considered the “gold standard” for molecular characterisation of *Salmonella* due to its high discriminatory power and high degree of reproducibility. However, it can fail to differentiate highly-related isolates and it is labour-intensive and time consuming (Foley et al., 2007).

1.1.1.1.3.3 Multiple locus variable-number tandem repeat analysis (MLVA)

MLVA typing exploits the variation in number of tandem repeats (variable-number tandem repeats or VNTR) in multiple loci of the genomic DNA (Chiou, 2010). The MLVA scheme for *S*. Typhimurium commonly involves the analysis of four genomic regions (STTR3, STTR5, STTR6 and STTR9) and a plasmid region situated on the virulence plasmid pSLT (STTR10) (Lindstedt et al., 2013). Based on the number of the repeats, a digit is assigned to each locus so that a sample can be easily referred to using an array on numbers (Lindstedt et al., 2003). Thanks to its high discriminatory power MLVA is suitable for for outbreak investigations (Chiou, 2010).

1.1.1.1.3.4 Multi-locus sequence typing (MLST)

Multi locus sequence typing (MLST) discriminates isolates based on DNA sequence variation in a set of housekeeping genes. The MLST scheme for *Salmonella* assesses variation in seven housekeeping genes (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA* and *thrA*) (Kidgell et al., 2002). Sequence types (ST) are assigned based on the allelic profile. MLST is suitable for phylogenetic studies but not for outbreaks investigation due to limited discriminatory power (EFSA, 2013)

1.1.1.1.3.5 Whole-genome sequencing (WGS)

Whole genome sequencing (WGS) is the method with the highest discriminatory power enabling the discrimination of isolates at the nucleotide level. With the advent of next-generation sequencing (NGS) in the early 2000s the determination of nucleotide sequences became progressively quicker and cheaper compared to the traditional Sanger sequencing method (von Bubnoff, 2008).

Depending on the length of the sequence generated, sequencing platforms can be distinguished in short-read (50-400 bp) or long-read (>10kb reads) platforms (WHO, 2018).

The most common short-read technologies are the Illumina and the Ion Torrent suites, which are based on a sequencing-by-synthesis approach (WHO, 2018). In Illumina sequencing the nucleotide sequence is determined by fluorescently-labelled nucleotides incorporated into the elongated strand, whereas Ion Torrent detects the nucleotide bases by measuring the hydrogen ions released by the elongation (WHO, 2018).

Short-read sequencing generates fragmented genome assemblies and cannot resolve complex, repetitive sequences such as phage genomes (Besser et al., 2018). Long-read sequencing technologies developed by Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT) overcome these limitations. While in the PacBio platforms the sequencing occurs by synthesis, ONT technology translates small perturbations of current generated by a single DNA molecule passing across a biological pore into a nucleotide sequence (WHO, 2018).

In comparison with traditional typing methods, surveillance using WGS has the highest resolution power, enabling the detection of potential outbreaks which would have been undetected with the previous methods (WHO, 2018). Because of the high amount of information generated, WGS can also help with the prediction of emerging threats (i.e. following acquisition of virulence genes within a microbial population) and enables the monitoring of antimicrobial resistance (WHO, 2018). The cost of sequencing per isolate is currently comparable to that of traditional typing systems (ECDC, 2015). WGS is also able to assess the phylogenetic relationships of the isolates and investigate their evolutionary history at the highest possible resolution (WHO, 2018).
For all these advantages a number of Reference Laboratories in EU/EAA already routinely perform WGS on *Salmonella* isolates. In order to enhance the use of WGS in food safety and public health protection the ENGAGE project (Establishing Next Generation sequencing Ability for Genomic analysis in Europe) was established between eight public health, food and veterinary institutions across the European Union (EU) (Hendriksen et al., 2018).

1.1.2 Salmonellosis in humans

Infection by *Salmonella* can result in either a limited gastrointestinal disease, or in a systemic infection: enteric fever, including typhoid and paratyphoid disease, or bacteraemia associated with invasive non-typhoidal *Salmonella* (iNTS disease) (Andrews-Polymenis et al., 2010, Feasey et al., 2012).
The transmission of *Salmonella* occurs through the oral-faecal route, mainly following the ingestion of contaminated food or water, and more rarely by direct contact with infected animals and humans (Hoelzer et al., 2011). Animals

constitute the main reservoir for non-typhoidal *Salmonella*, whereas humans are the sole reservoir for typhoidal and iNTS serovars (Gal-Mor et al., 2014).

Non-typhoidal salmonellosis consists in a self-limited gastroenteritis, characterised by diarrhoea, abdominal pain, vomit and fever (McGovern and Slavutin, 1979). The infective dose for non-typhoidal salmonellosis has been estimated to be 10^6 (Mintz et al., 1994); however a lower dose (10^2-10^3 cells) may be sufficient, such as in patients with low gastric pH (Giannella et al., 1971). The symptoms of *Salmonella* appear after 12-72 hours following the ingestion of contaminated food. In many cases, recovery occurs spontaneously in 4-7 days (Coburn et al., 2007). Treatment with antimicrobials is not recommended, except in case of complications, which include bacteraemia, focal infections, vascular, neurological and muscular issues, and chronic arthritis (Sirinavin and Garner, 2000).

Some categories of individuals, such as children, elderly people and immunocompromised, are more at-risk of developing complications (Crum-Cianflone, 2008).

The term enteric fever comprises typhoidal salmonellosis, caused by S. Typhi, and a related disease termed paratyphoid fever, caused by *S*. Paratyphi A, B or C (Fabrega and Vila, 2013).. Symptoms such as headache, abdominal pain, diarrhoea or constipation, precede the onset of the fever (Bhan et al., 2005). Besides the fever, the infected individuals can also develop hepatomegaly and splenomegaly, enlarged liver and spleen, respectively (Eng et al., 2015). In the reticuloendothelial system typhoid *Salmonella* can persist and cause relapses in the 10% of the patients (Parry et al., 2002). Colonisation of the gallbladder is associated with the establishment of the carrier state in which the individuals asymptomatically carry and shed the pathogen through the faeces (Andrews-Polymenis et al., 2010).

Invasive non-typhoidal salmonellosis is a severe systemic infection caused by certain *S*. Typhimurium strains adapted to humans, such as *S*. Typhimurium ST313. iNTS is common in sub-Saharan Africa where the disease is associated with malnourished children and immunocompromised individuals, particularly those infected by HIV, and has a case fatality of 20-25% (Okoro et al., 2015). The symptoms include fever, hepatosplenomegaly and pneumonia, diarrhoea is often absent (Feasey et al., 2012).

The distinct pathogenesis between enteric fever and non-typhoidal salmonellosis is caused by the ability of typhoidal serovars to elude the immune system and prevent the acute inflammatory response in the intestinal lumen, which enables the bacteria to invade the intestinal barrier and disseminate (Raffatellu et al., 2008). By contrast, iNTS exploit immunological defects of the host to spread to systemic sites (Feasey et al., 2012).

1.1.3 Epidemiology of food-borne salmonellosis

Salmonella is the second most common cause of foodborne gastroenteritis in humans, after *Campylobacter* spp. (ECDC, 2018). It is estimated that 93.8 million cases of non-typhoidal salmonellosis (NTS) occur worldwide every year causing 155,000 deaths (Majowicz et al., 2010).

NTS is particularly common in developing countries, where the poor hygienic and sanitary conditions facilitate the dissemination of the pathogen. Nonetheless, NTS represents one of the main public health concerns in developed countries where it represents a major cause of foodborne gastroenteritis and it associated with a considerable economic burden (Majowicz et al., 2010).

In the USA, NTS disease was the leading cause of hospitalisation and death by foodborne diseases. The economic burden associated with NTS disease was estimated as high as $3.7 billion per year in the USA (Hoffman et al., 2015). .

In Europe, *Salmonella* caused 91,662 cases, accounting for over 25% of all the foodborne infections reported in 2017 (ECDC, 2018). It also represented the most frequent cause of foodborne outbreaks which were mainly caused by contaminated animal products, including eggs and related products, meat and pork, but also bakery products and mixed food (ECDC, 2018).

Nevertheless, the notification rate of *Salmonella* in the EU/EAA significantly decreased since 2007 falling from 35.2 to 19.7 cases per 100,000 people in 2017 (ECDC, 2018) (Figure 1.3).

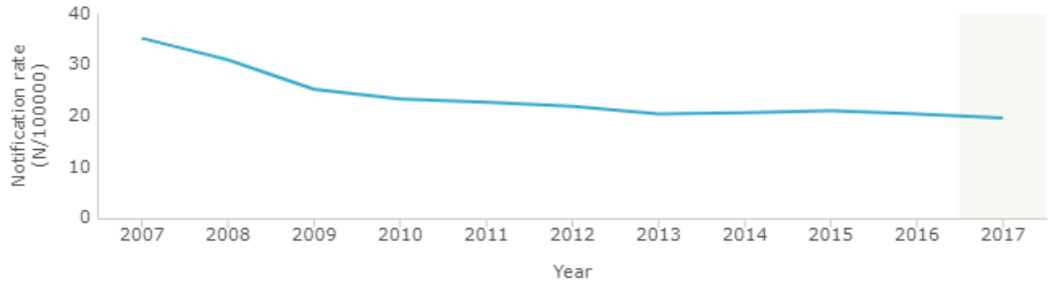

Figure 1.3 Notification rate of non-typhoidal salmonellosis in the EU/EEA between 2007 and 2017. The plot shows the notification rate, reported as number of cases per 100,000 inhabitants, per year. The graph was downloaded from the Surveillance Atlas of Infectious Disease (ECDC).

The most common serovars responsible for the human infections were *S*. Enteritidis, which accounted for almost half of the *Salmonella* infections in 2017 (49.1%), followed by *S*. Typhimurium (including the monophasic variants of *S*. Typhimurium) which accounted for the 21.4% of human cases (ECDC, 2018). While *S*. Enteritidis was predominantly associated with poultry (broilers and layers), *S*. Typhimurium was ubiquitous (mainly broilers and pigs) and monophasic variants of *S*. Typhimurium (*S*. 4,[5],12:i:-) were mainly isolated from pigs and also broilers sources (ECDC, 2018) (Figure 1.4).

Other common serovars were Infantis (2.3%), associated with broiler sources, Newport (1.2%), associated with broilers and turkey, Agona (0.8%), Kentucky (0.8%), Derby (0.8%),Stanley (0.7%) and Virchow (0.6%) (ECDC, 2018).

In Europe, over the year, salmonellosis is more common during the summer months when warmer weather and unrefrigerated foods creates more favourable conditions for *Salmonella* growth (ECDC, 2018).

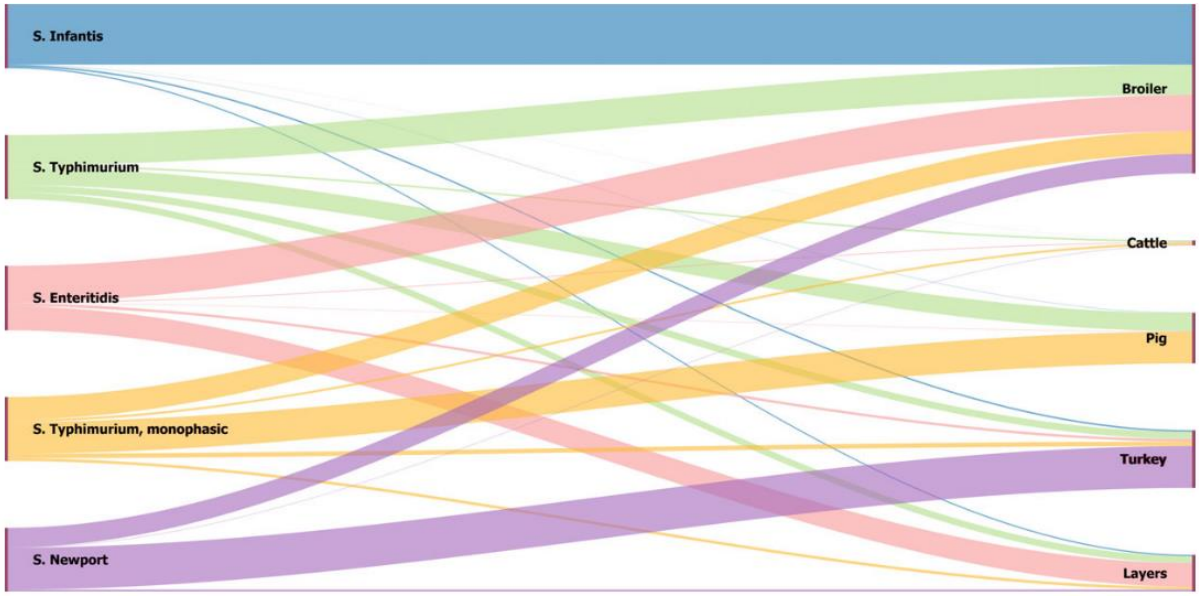

Figure 1.4 Sankey diagram showing the distribution of the five most common *Salmonella* serovars isolated in different livestock species in the EU (ECDC, 2018). The diagram shows the proportion of the five most common *S. enterica* serovars in the EU (Enteritidis, Typhimurium, monophasic Typhimurium, Infantis and Newport) isolated from different livestock species in the EU in 2017.

1.1.4 Molecular basis of *S. enterica* subsp. *enterica* pathogenesis

Virulence can be defined as the ability of a microorganism to damage the host (Casadevall and Pirofski, 2003). The infection by *Salmonell*a is a complex, multi-step process which requires the coordinated expression of several virulence factors. The outcome of the disease depends on multiple factors not only related to the microorganisms, such as pathogen load and virulence potential, but also associated with the host species and status of the immune system (Eng et al., 2015).

The infection of *Salmonella* occurs in the small and large intestines where the microorganisms contact enterocytes and M cells (Schulte and Hensel, 2016, Zhang et al., 2003).

The adhesion of the microorganisms to the epithelial cells, mediated by fimbrial and non-fimbrial adhesins, such as BapA, SiiE, is followed by the invasion of the cells triggered by a type III secretion system (T3SS), a complex syringe-like machinery injecting effector proteins into the cytosol of the host cells (Fabrega and Vila, 2013, Zhang et al., 2000).

The genes encoding the T3SS, as well as the secreted effectors and the chaperones required for the delivery, cluster in a 40-kb locus termed *Salmonella* pathogenicity island-1 (SPI-1), which plays a key role in the adhesion, invasion and early-phase post invasion of the host cell. The effector proteins injected by the SPI-1 T3SS (SipA, SipC, SopB/SigD, SopD, SopE/SopE2) induce rearrangements in the actin cytoskeleton of the host cell resulting in membrane ruffling and uptake of the microbial cells. Additionally, they contribute to the generation of proinflammatory response (Fabrega and Vila, 2013).

Following internalisation, *Salmonella* replicates in the host cell within modified phagosomes called *Salmonella*-containing vacuoles (SCVs) (Knodler and Steele-Mortimer, 2003).

The expression of genes present on a second pathogenic island, termed *Salmonella* pathogenicity island 2 (SPI-2), are crucial for the survival of the microorganisms within the intracellular environment (Ochman et al., 1996). SPI-2 encodes for a type III secretion system and the relative effectors, such as SseG, SseF, SifA, SipA, PipB2, SseJ, involved in the maturation of the SCVs and in the formation of tubulovesicular structures, termed *Salmonella*-induced filaments (SIF) (Fabrega and Vila, 2013). Additionally, SPI-2 harbours the *ttrRSBCA* operon involved in the reduction of tetrathionate. The use of tetrathionate as a respiratory electron acceptor confers a competitive advantage to *Salmonella* over the resident microbiota in the inflamed gut (Winter et al., 2010b).

Many clinically-relevant *Salmonella* serovars, including *S*. Typhimurium, harbour further virulence genes encoded on the pSLT plasmid, a 90-95 kb IncF plasmid (Fabrega and Vila, 2013). The plasmid-encoded genes contribute to intracellular multiplication, systemic infection and persistence of *Salmonella* in the host and confer protection to serum killing (Fabrega and Vila, 2013). Invasion of the intestinal epithelium by *S*. Typhimurium induces the production of pro-inflammatory cytokines and massive recruitment of neutrophils resulting in diarrhoea. The microbial factors responsible for the immune response include surface antigens (flagellin, curli fimbriae, lipopolysaccharides, etc.) recognised by Toll-like receptors and resulting in the activation of the NF-kB transcription factor (Tukel et al. 2005; Zeng et al. 2006). The secretion of the T3SS-SPI-1 effectors further stimulate the pro-inflammatory response. For example the guanine

nucleotide exchange factor activity of SopE is detected by the inflammasome resulting in the activation of the NF-kB-dependent inflammatory process (Keestra et al. 2013).

The pro-inflammatory response constitutes a host defence mechanism aimed at limiting the dissemination of *Salmonella* to systemic sites. However, *Salmonella* is adapted to the colonisation of the inflamed gut where it outcompetes the resident microbiota and proliferates by using alternative electron acceptors (Lopez et al., 2012). Diarrhoea also contributes in perpetuating the life cycle of *Salmonella* by promoting the faecal excretion of the microorganisms in the external environment from where a novel cycle of infection can initiate (Winfield and Groisman, 2003b). In this respect, the acquisition of pro-inflammatory virulence genes, such as *sopE*, is beneficial for *Salmonella*. The SopE-induced inflammation greatly contributes to *Salmonella* proliferation in the inflamed gut by inducing nitrate production, the preferential anaerobic electron acceptor of *Salmonella*, derived from the nitrogen reactive species of the activated immune system (Lopez et al., 2012).

1.1.4.1 Acquisition of virulence factors mediated by bacteriophage in *Salmonella*

The lysogenic conversion of temperate bacteriophage harbouring virulence determinants as cargo genes can further expand the repertoire of virulence factors in *Salmonella* (Figueroa-Bossi et al., 2001). A number of phage-encoded virulence factors in *sensu lato*, namely toxins/invasins which damage the host and fitness factors which enhance the survival in the host without damaging it, have been described for *S*. Typhimurium (Figueroa-Bossi et al., 2001).

The acquisition of the guanine nucleotide exchange factor (GEF) SopE, a SPI-1 T3SS effector, was mediated by a P2-like bacteriophage, termed SopEΦ. Upon translocation, SopE activates the Rho-GTPases Rac-1 and Cdc42 responsible for actin rearrangements resulting in membrane ruffling and microbial invasion of the host cells (Hardt et al., 1998a). SopE also contributes to the induction of the pro-inflammatory response and promotes intracellular replication of the bacteria in the early SCV (Muller et al., 2009, Vonaesch et al., 2014).

The lambdoid Gifsy-1 phage encodes three virulence genes involved in the intra-cellular survival of *Salmonella*: *gogB*, *sarA* and *pagK2*. GogB prevents the inflammatory response through inhibition of NF-κB (Pilar et al., 2012). The SPI-2

T3SS effector SarA activates anti-inflammatory pathways in the eukaryotic host cells through the activation of the transcription factor STAT3 (Jaslow et al., 2018). PagK2 is secreted in outer membrane vesicles and contributes to survival in macrophages through an unknown mechanism (Yoon et al., 2011). The inhibition of the inflammatory response is suggested to be important for intracellular growth and long term infection of the host in systemic sites (Wahl et al., 2019).

The lambdoid Gifsy-2 phage encodes a variety of virulence genes, such as the periplasmic CuZn-superoxide dismutase SodCI, which enhances the survival of *Salmonella* in macrophages, and the putative protein GtgE of unknown function which contributes to the virulence of *S*. Typhimurium in a mouse model (Ho et al., 2002). Additional putative virulence factors were identified in Gifsy-2 based on sequence homology, including *gtgA*, whose product is 75% identical to PipA and encoded on *Salmonella* pathogenicity island 5, the *ailT* gene homologous to the Lom/Ail family of proteins conferring serum resistance, and *gtgF* homologous to *msgA*, involved in survival in macrophages. Gifsy-2 further harbours the SseI/Srf, an effector of the SPI-2 T3SS (Uzzau et al., 2002).

The P2-like Fels-1 phage harbours the superoxide dismutase sodCIII, the neuraminidase/sialidase and the NanH which contributes to the degradation of the host glycocalyx required for invasion of the host (Arabyan et al., 2016, Brussow et al., 2004).

1.2 *Salmonella* in the pork chain

1.2.1 *Salmonella* in pigs

Pigs are one of the most common animal reservoirs of *Salmonella* (Campos et al., 2019). Contaminated pork and pork products are well-established vehicles of salmonellosis and have been responsible for a number of outbreaks in humans in Europe (EFSA, 2010b). Pig colonisation by *Salmonella* can occur either by vertical (i.e. from sow to piglet) or horizontal (environmental source, feed, trucks) transmission, a combination of which may results in a permanent cycle of contamination on farm (Campos et al., 2019)

A variety of *Salmonella* serovars can infect pigs. The host-adapted *S*. Choleraesuis and *S*. Typhisuis are responsible for severe systemic disease in pigs, but only occasionally cause salmonellosis in the human population (Chiu et al., 2004). The symptoms, which appear 36-48 hours after the infection, comprise

enterocolitis, septicaemia, and also pneumonia and/or hepatitis (Bonardi, 2017). Other common serovars circulating in pig populations are *S*. Typhimurium and *S*. 4,[5],12:i:-, *S*. Derby and *S*. Infantis, which are among the most common causes of NTS in humans (ECDC, 2018, Bonardi, 2017). In contrast to Choleraesuis and Typhisuis, these serovars cause diarrhoea and enterocolitis in pigs and are host generalists as they can circulate in different animal species (EFSA, 2010a). Nevertheless, asymptomatic carriage of *Salmonella* is common in pigs. Asymptomatic pigs can carry *Salmonella* in tonsils, gut and gut-associated lymphoid tissue (Bonardi, 2017). The shedding of the bacteria by asymptomatic carriers through faeces is considered a risk factor for the spread of the pathogen within the herd (San Roman et al., 2018). A lack of symptoms favours the spread of the pathogen within the herds as the immediate identification of the infected animal is not possible.

1.2.2 *Salmonella* in pork production chain

The contamination by *Salmonella* can occur at every step along the pork chain (Lo Fo Wong et al., 2002). Following a farm-to-fork model, animal feed is at the beginning of the food chain and it represents a well-established source of *Salmonella* on pig farms, although the prevalence of *Salmonella* on animal feed is low (Burns et al., 2015). As environmental contamination plays a major role in herd colonisation, the presence of wild animals on farms carrying and excreting the bacteria constitutes a risk factor for the infection of pigs (Lynch et al., 2018, Andres-Barranco et al., 2014). As reservoir of the bacteria, wild animal may also contribute to the persistence of *Salmonella* by continuous re-introduction of the pathogen (Andres and Davies, 2015).

Multiple risk factors are associated with *Salmonella* contamination and persistence in pig farms, including herd size and the density of the animals, the typology of the floor, herd health and the control of other pathogens, the lack of biosecurity measures regarding the movement of the animals, and the access of the personal and cleaning measures (Rostagno and Callaway, 2012, Bonardi, 2017, Andres and Davies, 2105).

During the pre-slaughter process, the proportion of animals infected by *Salmonella* increases (Rostagno and Callaway, 2012). *Salmonella* contamination can potentially occur during handling, transport or lairage stage. The stresses to

which the pigs are exposed before the slaughter can increase faecal shedding of the *Salmonella*, therefore increasing dissemination (Rostagno and Callaway, 2012). The duration of the lairage also positively correlates with enhanced detection of *Salmonella* in pigs (Rostagno and Callaway, 2012, Bonardi, 2017).

At the harvest stage (slaughter) the carcasses may be contaminated by the contact with the faecal matter or with contaminated equipment at the slaughterhouse, nevertheless, as the slaughtering process proceeds, contamination of the carcasses tend to decrease as a result of decontamination steps (Bonardi, 2017).

1.3 Phylogenetic reconstruction

Phylogenetics is the study of the evolutionary relationships between biological entities. Molecular markers generated using a variety of techniques such as MLVA and MLST can be used to determine the phylogenetic relationships of bacteria. However, with the advent of WGS, SNP-based phylogenies greatly improved the resolution of microbial phylogenies especially of highly clonal microorganisms such as *S*. Typhi (Baker et al., 2010).

In the public health sector, microbial molecular phylogenetics is used for outbreak detection, discrimination between epidemic versus sporadic isolates, for the identification of vehicles of infection and understanding the transmission routes of pathogens (WHO, 2018). Molecular phylogenetics is a useful tool to study population dynamics of pathogens as it enables us to distinguish the distinct clones circulating within a population, identify predominant clones and detect clonal replacement. For example, phylogenetics used in combination with genomics can provide valuable information about the efficacy of public health measures such as immunisation programmes and their impact on the microbial populations (Dyson et al., 2017).

The combination of phylogenetics and genomics (phylogenomics) aids us understanding how gene flux and genetic mutations contribute to the evolution and the adaptation of bacteria to a new niche. For example, phylogenomics enables us to understand how epidemics have emerged following the acquisition of genetic elements absent in related, non-epidemic microorganisms and it can help us to elucidate the dynamics which lead to acquisition and spread of antimicrobial resistance (Petrovska et al., 2016, Mather et al., 2018).

Phylogenomics can also be used to understand the process of pathogen adaptation to a novel ecological niche (Kingsley et al., 2009, Mather et al., 2018). Methods for phylogenetic reconstruction are either distance-based (e.g. Neighbor-Joining method) or character-based (e.g. maximum parsimony, maximum-likelihood and Bayesian inference methods). While distance-based methods compute the distance between each pair of sequences and use the resulting matrix to build a phylogeny, character-based methods simultaneously compare all the sequences, one character (single site in the alignment) at a time.

1.3.1 Maximum Parsimony reconstruction (MPR) methods

Maximum parsimony reconstruction (MPR) methods reconstruct the ancestral state character favouring the tree topology which assumes the least number of sequence mutations. The main drawbacks of the parsimony method consists of the impossibility of incorporating models of sequence evolution in the construction of the tree and long-branch attraction, in which long branches of the tree are grouped together preventing the inference of the correct tree topology (Yang and Rannala, 2012).

1.3.2 Maximum likelihood (ML) methods

Maximum-likelihood methods generate the phylogenetic tree with the highest probability of producing the observed data given the specified model of evolution. Multiple DNA substitution models are available (Yang and Rannala, 2012). The Jukes-Cantor model assumes equal base frequencies (P(A) = P(C) = P(G) = P(T) = (1/4) and assigns the same probability to each mutation (Jukes and Cantor, 1969). Similarly, the Kimura model assumes equal base frequencies but assigns a different probability score to transitions and transversions (Kimura, 1980). The Felsenstein model, by contrast, does not distinguish between transitions and transversions, but the base frequency can vary (Felsenstein, 1981). The Hasegawa, Kishino and Yano model allows unequal base frequencies and makes a distinction between transitions and transversions (Hasegawa et al., 1985). The generalised time-reversible (GTR) model is the general time-reversible model which incorporates different rates for every change and different nucleotide frequencies (Tavaré and Miura, 1986). As the mutation rate of the nucleotide sites may not be the same, the concept of rate variation was introduced in the models assuming a gamma (Γ) distribution (Yang, 1993), so that some sites were allowed to mutate at higher rates than others. The Randomized Axelerated Maximum

Likelihood (RAxML) program (Stamatakis, 2014, Stamatakis, 2006) is one of the most popular software for estimation of ML phylogenetic trees due to its speed and accuracy (Stamatakis, 2006, Zhou et al., 2017). RAxML first generates a starting tree by performing a fast MPR search, followed by the so-called “lazy subtree rearrangement” search which identifies the tree with the highest likelihood (Stamatakis, 2006).

1.3.3 Bayesian inference

Bayesian inference methods are based on Bayes’ theorem and combine a prior probability of a phylogenetic tree with the likelihood of the data to generate a posterior probability distribution on phylogenetic trees (Yang and Rannala, 2012). Bayesian inference methods treat trees and tree parameters as random variables. Markov chain Monte Carlo (MCMC) algorithms are used to generate a sample from the posterior distribution. One of the most common MCMC algorithms is the Metropolis-Hastings algorithm. This method generates a random tree (Ti) then selects a neighbour tree (Tj) and compares the probabilities of the two trees. If the probability of Tj is higher than Ti then this is accepted, otherwise Ti is kept and a new tree is considered and compared with it. The process is repeated N times (Yang and Rannala, 2012). The phylogenetic tree with the highest posterior probability is then returned.

Several phylogenetic tools performing Bayesian inference have been developed. The early methods assumed a molecular clock, namely the uniform rate at which all the sequences evolve. However, software which eliminate this assumption have been developed, such as MrBayes (Ronquist et al., 2012). More recently, software which use relaxed-clock models to infer roots were developed, such as BEAST (Bouckaert et al., 2019). In this way different substitution models can be applied in different parts of the tree. A major limitation of Bayesian methods comprise the fact that they are computationally intensive (Yang and Rannala, 2012). Contrary to ML methods which return a single “true” phylogenetic tree, Bayesian methods return a distribution of phylogenetic trees.

1.4 Antimicrobial resistance (AMR) in *S.* Typhimurium

The adoption of intensive farming practices combined with the introduction of antimicrobials in veterinary medicine contributed to the emergence of antimicrobial resistance (AMR) in *S*. Typhimurium (Rabsch et al., 2001). The first

drug-resistant *S*. Typhimurium, which belonged to phage type DT29, was isolated in England and Wales in 1963 from human and bovine samples, suggesting its emergence in cattle. Shortly after its emergence, the antimicrobial-resistant DT29 caused a large outbreak which lasted from 1964 to 1966 and became the predominant type responsible for human disease (Rabsch et al., 2013).

Epidemiological records from England, Wales and Germany showed that the successive epidemics of *S*. Typhimurium in livestock and humans were characterised by dominant multi-drug (MDR) resistant clonal populations which could be discriminated by phage typing (Rabsch et al., 2001, Magiorakos et al., 2012). Notably, a clonal replacement has occurred approximately every 15 years, when a novel epidemic emerged and the previous clone almost disappeared (Rabsch et al., 2001) (Figure 1.5). All the epidemic clones first emerged in cattle, then spread to other animal species. The drug-susceptible *S*. Typhimurium DT9 was predominant in Germany until 1972 when it was replaced by the *S*. Typhimurium DT204 resistant to sulphonamides (Su) and tetracyclines (T). Following the acquisition of an IncH plasmid, *S*. Typhimurium DT204 acquired further resistance to ampicillin (A), chloramphenicol (C), kanamycin (K) and streptomycin (S) (Rabsch et al., 2001). *S*. Typhimurium DT204 resistant up to nine antimicrobials was also reported (Rabsch et al., 2013).

In the 1980s DT204 was replaced by *S*. Typhimurium DT104 characterised by ACSSuT resistance profile which dominated until 2010. However, it should be noted that while less than 40% of DT204 were MDR, over 95% of DT104 were MDR (Rabsch et al., 2001). In DT104 the drug resistance was encoded on composite class I integron integrated in the genome, the *Salmonella* genomic island 1 (SGI-1) (Boyd et al., 2002). However, variations of the AMR profile were reported as a result of rearrangements within the SGI-1 (Boyd et al., 2002). The further acquisition of resistance to nalidixic acid by *S*. Typhimurium DT104 resulted in a sudden increase in the number of infections caused by the clone around 2010 (Rabsch et al., 2013). Nonetheless, since 2006 a monophasic variant of *S*. Typhimurium (S. 4,[5],12:i:-) with ASSuT resistance profile has been emerging, replacing *S*. Typhimurium DT104.

The factors responsible for the emergence and succession of dominant MDR *S*. Typhimurium clones are not known. Calf-dealing trades were suggested as a

cause for the rapid spread of drug-resistant clones as both *S*. Typhimurium DT29 and DT204 were associated with extensive trading (Rabsch et al., 2001). Additionally, the use of antimicrobials in agriculture may have contributed to the emergence of epidemic clones. However, AMR alone cannot explain the selective advantage conferred by one clone over the other as resistance genes in the successive clone are often present in the previous clone. Nonetheless, in case of DT104 the resistance was chromosomally encoded not plasmid encoded as for DT204, that may result in more stably maintained resistance. Consistent with this idea, a greater proportion of DT104 were MDR compared with DT204 suggesting the selective advantage conferred by the AMR. Finally, the acquisition of mobile genetic elements has been suggested as a key driver for the emergence of dominant epidemic clones. For instance, DT204 acquired a phage-mediated virulence gene (*sopE*) which was absent in the previous epidemic clone and in non-epidemic strains (Mirold et al., 1999).

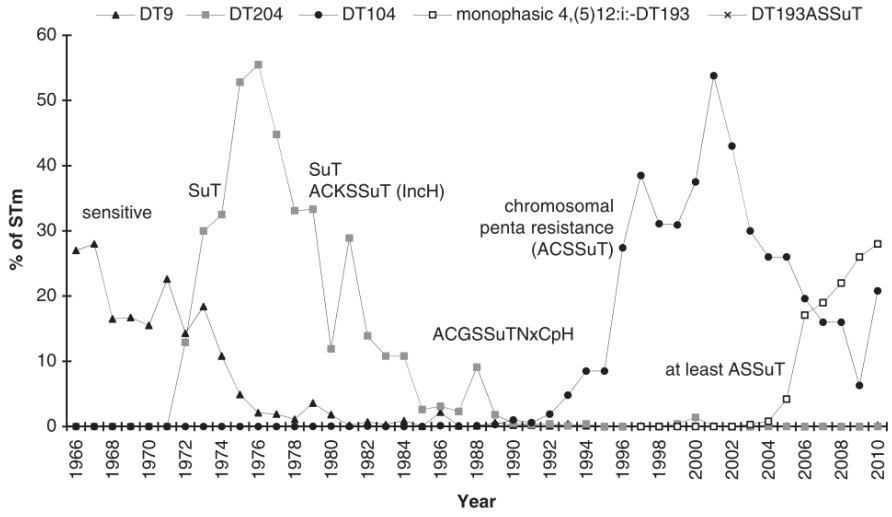

Figure 1.5 Succession of epidemic drug-resistant *S*. Typhimurium lineages in Germany (Rabsch et al., 2013). The plot shows the percentage of human salmonellosis cases caused by successive lineages of *S*. Typhimurium in Germany between 1986 and 2010. The distinct lineages are distinguished by phage type; the AMR profile associated with each lineage is reported.

1.5 Emergence and dissemination of monophasic variants of *S*. Typhimurium (*S*. 4,[5],12:i:-)

1.5.1 The evolutionary history of *S*. 4,[5],12:i:-

Since the late 1990s monophasic variants of *S*. Typhimurium lacking of second-phase flagellar antigen (FljB), defined by the serotype formula *S*. 4,[5],12:i:-, emerged and replaced the previous dominant epidemic clone DT104 (Rabsch et al., 2013). Contrary to the previous MDR *S*. Typhimurium epidemics which first emerged in cattle, then spread to poultry and swine, *S*. 4,[5],12:i:- emerged in pigs, later spreading to poultry and cattle (Hauser et al., 2010). Pork and pork products represent the main vehicle of the pathogen (Hauser et al., 2010, Xie et al., 2020).

Before the spread of the current dominant pandemic *S*. 4,[5],12:i:- clone, multiple lineages of *S*. 4,[5],12:i:- independently emerged but were to confined geographical areas.

In Europe, *S*. 4,[5],12:i:- first emerged in Spain where its first isolation dates back to 1997. However, only three years later it was the fourth most common *Salmonella* serovar isolated in the country (Echeita et al., 1999).

The so-called Spanish clone was sequence type ST19 and DT U302 phage type. Additionally, the majority of the isolates were resistant to ampicillin, chloramphenicol, streptomycin, sulphonamide, tetracycline, gentamicin and trimethoprim-sulfamethoxazole (ACSSuT-GSxT) (Echeita et al., 1999, Guerra et al., 2001, Garcia et al., 2013). The determinants of MDR were encoded on large non-transferable plasmids (~120-200 kb) belonging either to incompatibility group IncA/C or IncA/C and IncN (Garcia et al., 2011).

This novel serotype was suggested to be a monophasic variant of *S*. Typhimurium because of the presence of an IS200 sequence in the *fliA-fliB* intergenic region previously described as *S*. Typhimurium-specific, and of a 162-bp fragment specific to *S*. Typhimurium phage types DT104 and U302 (Echeita et al., 2001). Moreover, the high degree of sequence identity to *S*. Typhimurium strain LT2 provided further evidence about the genetic origin of the monophasic strains. Only five deletions situated within the allantoin-glyoxylate operon, the prophage Fels-1-, Gifsy-1-, and Fels-2- related genes, as well as in the *fljAB* operon, differentiated the genomes of *S*. 4,[5],12:i:- and *S*. Typhimurium LT2

(Garaizar et al., 2002). The disruption of the flagellar operon resulted from the insertion of an IS26 element in this region (Garaizar et al. 2002). Various pattern of deletions were identified: while the start of the mutation was conserved (same nucleotide in the STM2758 gene), the terminal portion was subjected to more variability, reaching the *iroB*, *iroC* or *emrA* gene (Laorden et al., 2010). Sub-typing analyses (PFGE) suggested that the Spanish *S*. 4,[5],12:i:- evolved from or shared a common ancestor with *S*. Typhimurium DT U302 with which it shared the same phage type and plasmids (de la Torre et al., 2003, Echeita et al., 2001).

A distinct *S*. 4,[5],12:i:- lineage, sequence type ST19, was reported in the USA in the 1990s (EFSA, 2010b, Hoelzer et al., 2010). The majority of isolates belonging to the American clonal complex were pan-susceptible or resistant to few antimicrobials (i.e. namely streptomycin, sulfamethoxazole, tetracycline) (Switt et al. 2009). The deletion of the phase 2 flagellar locus encompassed 76 genes in strain CVM23701, in which a 7-kb genetic element encoding prophage related sequences integrated (Soyer et al., 2009).

An additional *S*. 4,[5],12:i:- clonal complex ST19 designated as Southern European clone was reported as the third most common monophasic *S*. Typhimurium variant in Portugal (Mourao et al., 2014). The Southern Europe clone was unrelated to the Spanish clone but shared a common ancestor with the U.S. clone. The most common resistance pattern was CSSuTTm (chloramphenicol, streptomycin, sulphonamide, tetracycline and trimethoprim) conferred by an atypical class 1 integron in a large non-transferable IncR plasmid (Mourao et al., 2014).

In 2006, a novel monophasic variant of *S*. Typhimurium was first reported in Luxembourg and in Germany, then in other European countries, including Austria, Ireland, France, Italy, and the Netherlands (Trupschuch et al., 2010). Because of the numerous differences with the Spanish clone, these isolates were suggested to be part of a distinct “European” lineage, arisen independently from the Spanish clone (Garcia et al., 2013). The European *S*. 4,[5],12:i:- was ST34 and phage types DT193 and DT120, although other phage types were reported with lesser frequency (Petrovska et al., 2016). In addition, a complex pattern of

deletions was described within the *fljAB* flagellar operon, that were distinct from the deletion in the Spanish clone (Petrovska et al. 2016). For instance, one isolate (S04698-09) representative of *S*. 4,[5],12:i:- ST34 clonal complex lacked a region of 15,726 bp encompassing the *fljB* gene. However, a composite transposon of 24.5 kb integrated in the *fliAB* operon harbouring genes encoding for mercury resistance (*merTABCDE* and *merR*) and antimicrobial resistance to ampicillin, streptomycin, sulphonamide and tetracycline (ASSuT) conferred by bla_{TEM-1}, *straA/strB*, *sul2* and *tet(B*) genes, respectively (Petrovska et al. 2016).
Contrary to the Spanish monophasic *S*. Typhimurium, the European clone had an intact allantoin-glyoxylate operon and harboured Gifsy-1 and Gifsy-2 prophages (Trupschuch et al., 2010).
Notably, most of *S*. 4,[5],12:i:- ST34 isolates lacked the virulence plasmid pSLT (Hopkins et al., 2010, Petrovska et al., 2016)

The reconstruction of the phylogenetic relationships of *S*. Typhimurium and *S*. 4,[5],12:i:- ST34 from UK and Italy based on whole genome sequences showed that the isolates from the European clone emerged from a common ancestor approximately 20 years ago (Petrovska et al., 2016). The emergence of the clone coincided with the acquisition of a novel, 80 kb genomic island termed *Salmonella* genomic island 4 (SGI-4) that was integrated in the tRNA-Phe locus and encoded the resistance to multiple heavy metals (Petrovska et al., 2016). Notably, the SGI-4 conferred enhanced resistance to copper, particularly accentuated under anaerobic, that is supplemented in pigs feed as a growth promoter and antimicrobial (Nicholson et al., 1999, Branchu et al., 2019). This together with the observation that the SGI-4 was exclusively associated with *S*. 4,[5],12:i:- ST34 which was stably maintained with the exception of few losses suggested the relevant role played by the SGI-4 in the success of the clone (Branchu et al., 2019). Additionally, the acquisition of the virulence gene *sopE* was reported in *S*. 4,[5],12:i:- ST34 following lysogenic conversion of a novel prophage termed mTmV (Petrovska et al., 2016).

Currently, *S*. 4,[5],12:i:- ST34 is pandemic. In China, *S*. 4,[5],12:i:- ST34 is the most common *S*. Typhimurium type isolated from humans and animals (Wong et

al., 2013, Sun et al., 2014). Similarly, *S*. 4,[5],12:i:- ST34 is the predominant clone in swine and cattle in Japan (Arai et al., 2018).

In Australia an increasing proportion of salmonellosis was recorded in concomitance with the emergence of *S*. 4,[5],12:i:- phage type DT193 with ASSuT resistance profile likely related to *S*. 4,[5],12:i:- ST34 clone (Hamilton et al., 2015). In Latin American, *S*. 4,[5],12:i:- ST34 was reported in Colombia, although it was more susceptible to antimicrobials than the isolates circulating in Europe (Li et al., 2019).

S. 4,[5],12:i:- ST34 has been reported in swine samples from the United States Midwest (Elnekave et al., 2018). However, along with the tetra-resistance pattern ASSuT, a small proportion of the isolates acquired resistance to quinolones either following nucleotide mutations in the *gyrA-gyrB* and *parC-parE* genes, plasmid-mediated acquisition of *qnrB2* resistance gene or altered expression of efflux pumps/porin channels (Elnekave et al., 2018).

Additionally*, S*. 4,[5],12:i:- ST34 was described in Vietnam (Mather et al., 2018). Nonetheless, a cluster of Vietnamese isolates presented a number of genetic differences compared with European *S*. 4,[5],12:i:- ST34. In particular, the isolates had re-acquired the phase 2 flagellin gene and encoded further antimicrobial resistance to bleomycin, fluoroquinolones, sulphonamides, trimethoprim, kanamycin, streptomycin, chloramphenicol, spectinomycin, florfenicol, hygromycin, apramycin, β-lactams and rifampicin harboured by a large IncHI2 plasmid (Mather et al., 2018). Notably, the Vietnamese ST34 variant strongly associated with invasive salmonellosis in HIV-infected individuals, which is uncommon outside sub-Saharan Africa where it is caused by *S*. Typhimurium ST313 (Mather et al., 2018, Kingsley et al., 2009). Unlike the latter, no evidences of genome degradation were identified in the Vietnamese iNTS isolates.

1.5.2 Epidemiology and outbreaks caused by *S*. 4,[5],12:i:-

The monitoring of *Salmonella* in humans is mandatory in the European Union (Commission Decision 2000/96/EC). However, the European Surveillance System (TESSy) to which the countries in the EU and European Economic Area (EEA) report their surveillance data was established only in 2007. For this reason, the data about the prevalence of monophasic variants of *S*. Typhimurium in the EU/EEA is available only from 2008 (EFSA, 2010b).

Between 2007 and 2009, monophasic *S*. Typhimurium was reported by seven out of the ten countries providing the serotyping analysis of the *Salmonella* strains (Denmark, Germany, Ireland, Italy, Luxembourg, The Netherlands and Spain). During this period, the number of cases increased from 360 to 1,416, although also the number of countries reporting the serotyping formula increased over the period considered (EFSA, 2010). The *S*. 4,[5],12:i:- reported likely belonged to the European clone as the isolates belonged to phage types DT193 (71%) and DT120 (19%) (EFSA, 2010b).

Since 2007 the number of reported cases went up so sharply that *S*. 4,[5],12:i:- became the fourth most common *Salmonella* serotype in 2010, and the third since 2011. Along with the rise of the human infections, the number of the countries reporting the emerging serotype increased, indicating its spread across the Europe (EFSA, 2010b) (Table 1.1 Figure 1.6).

Table 1.1 Human infections caused by *S*. 4,[5],12:i:- in EU/EAA between 2010 and 2017. The table reports the number of the confirmed infections, the percentage accounted by *S*. 4,[5],12:i:- over the total number of cases of salmonellosis, and the number of European member states (MS) which reported 4,[5],12;i serotype. The data was extracted from the annual report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks by EFSA and ECDC.

Year	Number of human cases	%	MS
2010	1,426	1.7	-
2011	3,739	4.6	10
2012	5,932	7.2	12
2013	6,313	8.6	14
2014	5,773	7.8	13
2015	5,786	7.9	15
2016	5,697	8.4	15
2017	6,324	8.0	16

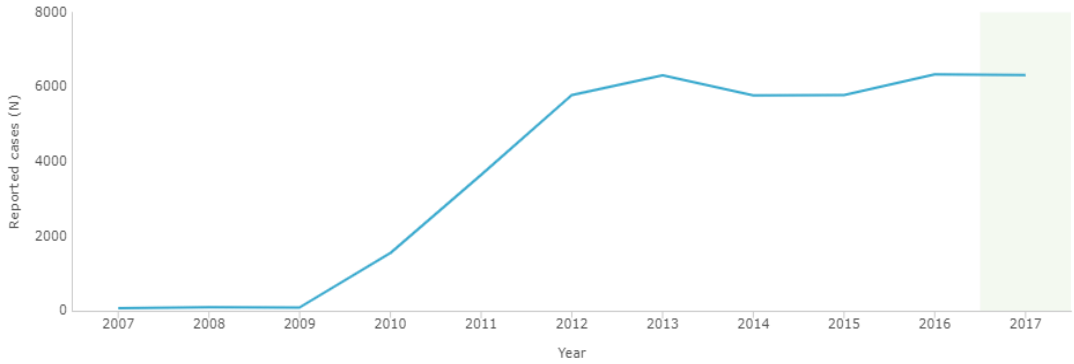

Figure 1.6 Number of reported cases of human salmonellosis caused by *S.* 4,[5],12:i;- in the EU/EAA between 2007 and 2017 (plot from the Surveillance Atlas of Infectious Disease, https://atlas.ecdc.europa.eu/public/index.aspx).

In Ireland, *S*. 4,[5],12:i:- was the second most common serovar responsible for human infections in 2013, accounting for the 16% of reported cases (n=53), according to the report of the Irish Health Protection Surveillance Centre (HPSC). The proportion of infections by *S*. Typhimurium remained constant over the following two years, accounting for the 15.7% and the 14.1% of cases in 2014 and 2015, respectively (HPSC).

In England and Wales the number of the infections caused by *S*. 4,[5],12:i:- increased from 47 to 151 cases between 2005 and 2010, although the number of cases was likely underestimated as full serotyping was not performed by all the laboratories (EFSA, 2010). Currently, *S*. 4,[5],12:i:- and S. Typhimurium are the second most common Salmonella serovars responsible for human infection in the UK (PHE, 2018).

Several foodborne outbreaks caused by *S*. 4,[5],12:i:- occurred in Europe.

In 2006, 133 confirmed cases were reported in Luxembourg, of which 24 individuals were hospitalised and one died. *S*. 4,[5],12:i:-, characterised by phage type DT193 and ASSuT R-type, was acquired following ingestion of locally-produced pork meat (Mossong et al. 2007). In the same year, Germany experienced some diffuse outbreaks caused by the same DT193 monophasic variant (Trüpschuch et al. 2010).

In France, the consumption of contaminated dried pork sausages caused two nationwide outbreaks. In May 2010 an outbreak affected fifty-four people of which 20 were hospitalised (Bone et al. 2010). A large outbreak with 337 patients

occurred in November and December 2011 (Gossner et al. 2012). A more circumscribed outbreak, involving 50 students attending four schools in Poitiers, France, occurred in October 2010 (Gossner et al. 2012).

In June and July 2014, a multi-state outbreak consisting of 38 human cases occurred in six Member States (Denmark, Finland, Germany, The Netherlands, Norway and Sweden). All the isolates belonged to R-type ASSuT, but exhibited two usual MLVA profiles (Food & Authority 2014).

Twenty-one cases in 10 geographically dispersed countries occurred following the consumption of food or drinks contaminated by *S*. 4,[5],12:i:- ST34 (Siira et al., 2019).

Outsides Europe, outbreaks caused by *S*. 4,[5],12:i:- have been reported in the USA, Canada, but also Brazil, Chile and Costa Rica (EFSA, 2010).

1.6 Biofilms

1.6.1 Biofilms as a strategy for environmental survival

The ability to persistent outside the host is of particular relevance for *Salmonella* (Winfield and Groisman, 2003a). Environmental reservoirs are an important intermediate step in the transmission of the pathogen as suggested by multiple outbreaks in developed countries (MacKenzie et al., 2017).

The formation of biofilms is considered an important strategy for environmental persistence (MacKenzie et al., 2017). Biofilms are sessile communities of bacteria irreversibly attached to a substratum or interface, embedded in an extracellular matrix (ECM) (Donlan and Costerton, 2002). Compared to planktonic cells, microbial cells in biofilms exhibit enhanced tolerance to antimicrobials (Hoiby et al., 2010), disinfectants (Corcoran et al., 2014, Condell et al., 2012), highly reactive chemicals (Scher et al., 2005) and desiccation (MacKenzie et al., 2017). For these reasons, the ability of microorganisms to form biofilm has important implications in medical, environmental and food safety sectors.

The role of biofilms in the survival of *Salmonella* outside the host is suggested by their formation at temperature below 37°C. Several studies have shown the ability of food, animal and human *Salmonella* isolates to produce biofilms in order to explain their persistence (MacKenzie et al., 2017). For instance, the ability of different *S*. *enterica* serovars to persist on industrial plants correlated with their

biofilm-forming ability (Vestby et al., 2009). Additionally, *Salmonella* persisted in scaly biofilms in the toilet of domestic homes up to 4 weeks after the infection ceased despite the use of cleaning products (Barker and Bloomfield, 2000).
However, a study of the colonisation of egg conveyer belts by a *S*. Enteritidis which repetitively contaminated a flock over a 3-year period showed that the material of the belt rather than biofilm formation ability displayed *in vitro* was the factor decisive for the persistence of the strain (Stocki et al., 2007). This suggested the difficulty of recreating in laboratory the conditions encountered in the real world by *Salmonella* (MacKenzie et al., 2017).

1.6.2 Composition of *Salmonella* biofilms

As most of the bacteria, *Salmonella* can live as either as free-living, planktonic cells or in sessile communities embedded in an exopolysaccharide matrix, called biofilms. *Salmonella* can produce biofilms on multiple types of surfaces, either on biotic or abiotic, and in surface-liquid or liquid-air interface (Fux et al., 2005, Steenackers et al., 2012). It should be noted that while generalists *Salmonella* are generally able to form biofilms, host-adapted Salmonella such as *S*. Typhimurium ST313 lost this ability (MacKenzie et al., 2017).
The composition of the ECM is complex and can vary with the type of the surface or interface on which the biofilms are attached to. The most abundant components of *Salmonella* biofilms are curli fimbriae, amyloid proteins, and the polysaccharide cellulose, polysaccharide constituted by β 1-4 linked glucose molecules (Steenackers et al., 2012). These physically interact with each other forming a complex network which confers architectural support to the biofilms, but also a very specific microenvironment, modulating the diffusion of nutrients.
Other types of fimbriae are involved in biofilm formation. Type 1 fimbriae (*fim*), long polar fimbriae (l*pf*) and plasmid-encoded fimbriae (*pef*) are necessary for biofilm formation on epithelial cells (Steenackers et al., 2012).
Additionally, type 1 fimbriae contribute to biofilm formation on gallstones (Steenackers et al., 2012). Another protein identified in biofilms was is BapA (biofilm-associated protein), a large multidomain protein, secreted by a type I secretion system (Latasa et al., 2005). However, BapA role in biofilms formation and architecture is not clear (Jonas et al., 2007). Similarly, the role of flagella in biofilm formation is controversial due to contradictory results published

(Steenackers et al., 2012). However, the contradictory results may be due to the use of different *Salmonella* serovars and experimental systems used in the studies.

Additionally, colanic acid, a capsular lipo-polysaccharide, contributes to the 3D structure of biofilms on epithelial cells, but is not required for biofilm formation on abiotic surfaces nor alfalfa seeds (Steenackers et al., 2012). Nevertheless, it contributed to the formation of a tight pellicle on the air-liquid interface in LB conditions, but not ATM medium, further suggesting importance of environmental conditions on the ECM composition.

Also, LPS is involved in the multicellular behaviour and in the formation of biofilms on certain types of surfaces, such as glass (Steenackers et al., 2012).

1.6.3 The process of biofilm formation

The formation of biofilms is a multi-step process. The initial adhesion of the planktonic cells to the surface is mediated by weak electrostatic forces. Changes in gene expression resulting in the production of extracellular molecules, such as cellulose, enable a more stable attachment to the substratum (Brooks and Flint, 2008). The properties of the surface influence the rate of microbial attachment. For instance, biofilms grow quicker on rough and hydrophobic materials (Donlan, 2001).

Attached cells then replicate forming microcolonies and produce extracellular polymeric substances, mainly polysaccharides, which provide the scaffold for the biofilm (Donlan, 2001). With the maturation of the biofilm, complex and heterogeneous 3D structures are then formed. They are characterised by the presence of “water channels” which enable the transport of oxygen and nutrients to the growing cells (Lewandowski and Evans, 2000). The final stage comprises the detachment of microbial cells (Donlan, 2001).

The nature and the characteristics of the substrate, as well as environmental conditions (pH, temperature), heavily influence the process of biofilm formation.

1.6.4 Investigation of biofilms *in vitro*

Both qualitative and quantitative assays have been developed to study the ability of *Salmonella* to produce biofilms *in vitro*.

The ability of *Salmonella* to produce biofilms can be phenotypically assessed by growing the bacteria at environmental temperatures (<37°C) on LB agar supplemented with the azo dye Congo Red, which binds cellulose and curli fimbriae, the two major components of the ECM. On the media, biofilms appear as red, dry and rough (RDAR) macro-colonies (Romling, 2005). The multicellular structures produced on Congo Red agar plates have a heterogeneous structure and contain cells in each of the different stages of biofilm formation (Simm et al., 2014). The centre of the colonies represents the oldest part of the biofilm, characterised by a complex 3D structure resulting from cellulose and curli fimbriae interaction (Simm et al., 2014). The rim of the colonies, by contrast, are smooth with little ECM and comprise flagellated bacteria (Simm et al., 2014). Other morphotypes comprise smooth and white (SAW) colonies when neither fimbriae nor cellulose are produced, pink, dry and rough (PDAR), indicative of the production of cellulose, but not curli fimbriae, brown, dry and rough (BDAR) expressing curli fimbriae but not cellulose, smooth, brown and mucoid (SBAM), which do not produce cellulose or fimbriae, but is able to form biofilm relying on the overproduction of capsular polysaccharide (Romling et al., 1998, Malcova et al., 2008).

The formation of pellicle represents another model for biofilm formation. *Salmonella* are able to produce biofilm on the air-liquid interface, namely pellicles, when inoculated into a nutrient rich medium and incubated at 22°C, whereas most microorganisms attach to the walls or fall to the bottom in liquid broth (Romling and Rohde, 1999).

The microtiter plate assay is a common, high-throughput method which enables the quantification of the biomass of the biofilm produced on the plastic walls and of the bottom of the wells of a microtiter plate. This process is achieved by staining the adherent cells adhering with the Crystal violet dye, which binds to the peptidoglycans layer, followed by measurement of the optical density (Merritt et al., 2005). Similarly, the Calgary Biofilm Device (CBD) allows to study biofilm formation on plastics: biofilms are formed on the polystyrene pegs on the coverlid of a microtiter plate (Ceri et al., 1999). Sonication is then used to disrupt biofilms before plating and CFU enumeration.

The methods described above are performed in static condition with no renewal of the medium nor aeration which may impair the formation of mature biofilms

(Merritt et al., 2005). Continuous flow systems overcome these limitations and enable to study the effect of materials and shear stresses on biofilms formation. However, these methods are generally low throughput require expensive equipment.

1.6.5 Environmental conditions promoting biofilm formation

Several environmental conditions influencing the formation of biofilms by *Salmonella* were identified.

In vitro the formation of biofilms is promoted during the stationary phase when the availability of nutrients is limited. The depletion of nitrogen and phosphate but not carbon source leads to the transcription of *csgD* (Gerstel and Romling, 2001). In addition, the presence of gluconeogenesis substrates (i.e. amino acids) favours biofilm formation, whereas the presence of glucose inhibits their production (Romling et al., 1998, White et al., 2010). Iron limitation, low osmolarity and microaerophilic levels of oxygen positively influence biofilm production (Romling et al., 1998, Gerstel and Romling, 2001). Also, temperatures below 30°C promote the formation of biofilms (Romling et al., 2003).

1.7 Aims of the study

a) Investigate *S*. Typhimurium and *S*. 4,[5],12:i:- circulating in Irish pig farms with the aim to place them in a wider European context, infer possible transmission events and investigate population structure of *Salmonella* on pig farms using whole genomic sequencing and phylogenomics.
b) Investigate biofilm-forming ability of *S*. Typhimurium and *S*. 4,[5],12:i:- from Irish pig farms using a combination of methods (Congo red agar plating, pellicle formation assay and microtiter plates assay) to test any possible correlation with between persistence on farms and biofilm-forming ability of the isolates.
c) Investigate the association of *sopE* with *S*. 4,[5],12:i:- ST34 in a collection representative of the diversity of S. Typhimurium and understand the impact of *sopE* acquisition on the epidemiological success of *S*. 4,[5],12:i:- ST34 using phylogenomics and ancestral state reconstruction methods.

d) Investigate the ability of the mTmV prophage to replicate and disseminate to diverse *S*. Typhimurium genotypes and *S*. *enterica* serovars through *in vitro* experiments and *in silico* analysis.
e) Investigate the *sopE* copy number variation in *S*. 4,[5],12:i:- ST34 associated with phage recombination and the transfer of the *sop*E cassette between unrelated phage using WGS and comparative genomics.

2 Materials and Methods

2.1 Molecular microbiology and molecular biology methods

2.1.1 Strains and plasmids

Table 2.1 List of microbial isolates used in the study. The name of the isolates, their genotype and source are reported together with the chapter and the purpose they were used for.

Strain	Genotype	Source	Chapter	Aim
IR715	*S.* Typhimurium ATCC 14028 Na^R	(Stojiljkovic et al., 1995)	Chapter 3	IncHI2 plasmid conjugation
3593A	*S.* Typhimurium	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
1468A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3500B	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
1465A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
1466E	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
1469C	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
1495C	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3476A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3496A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3508A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3519A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3520A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3836A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation

3844A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
3845A	*S.* 4,[5],12:i:-	(Burns, 2015)	Chapter 3	IncHI2 plasmid conjugation
E. coli BW25141 (pKD3)	Cm^R		Chapter 5	Mutagenesis
E. coli BW25141 (pKD4)	Kan^R		Chapter 5	Mutagenesis
S. Typhimurium SL1344 (pSIM18)	Hyg^R	(Kingsley et al., 2013)	Chapter 5	Mutagenesis
2882-06 (WITS1)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
SL1344 (WITS2)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
01960-05 (WITS4) *	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
03-715 (WITS6)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
L01157-10 (WITS7)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
4179-2001 (WITS10)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
S07676-03 (WITS11)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
SO7292-07 (WITS15)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
11020-1996 (WITS16)*	*S*. Typhimurium Kan^R	Dr. Gaëtan Thilliez	Chapter 5	Phage transfer
NCTC 13348	*S*. Typhimurium wt (ACSSuT)	National Collection of Type Cultures	Chapter 5	Phage transfer
D23580	*S*. Typhimurium wt	(Kingsley et al., 2009)	Chapter 5	Phage transfer
S01569-10	*S.* 4,[5],12:i:- wt	(Petrovska et al., 2016)	Chapter 5	Phage transfer
0343A	*S.* 4,[5],12:i:- wt	(Burns, 2015)	Chapter 5	Phage transfer

0343A mTmV_Cm	*S.* 4,[5],12:i:- Cm^R	This study	Chapter 5	Phage transfer
0343A mTmV_Kan	*S.* 4,[5],12:i:- Kan^R	This study	Chapter 5	Phage transfer

*WITS: wild-type isogenic tagged strain harbouring *aph* gene (Kan^R) inserted in *iciA-yggE* intergenic region.

Table 2.2 List of plasmids used in the study with their genotypes and sources.

Plasmid	Genotype	Source	Chapter	Aim
pKD3	Amp^R. pANTSγ plasmid derivative with FRT-flanked Cm^R cassette.	(Datsenko and Wanner, 2000)	Chapter 5	Mutagenesis
pKD4	Amp^R. pANTSγ plasmid derivative with FRT-flanked Kan^R cassette.	(Datsenko and Wanner, 2000)	Chapter 5	Mutagenesis
pSim18	Hyg^R. Temperature-sensitive pSC101 plasmid derivative encoding lambda recombination system (γ, β, *exo* genes) controlled by temperature-dependent repressor cI857 (inactivated at 42°C).	(Chan et al., 2007)	Chapter 5	Mutagenesis

2.1.2 Media and buffers

2.1.2.1 Lysogeny Broth (LB) Miller

LB broth Miller is a rich culture medium which was commonly used to grow the isolates, 36.0 g of the powder were dissolved in 1.0 litre of deionised water and autoclaved at 121°C for 15 minutes. The composition of the broth is the following:

Reagent	g/l
Tryptone	10.0
Yeast extract	5.0
NaCl	10.0

2.1.2.2 LB agar plates

A suspension containing 40 g of LB agar Miller in 1 litre of deionised water was autoclaved at 121°C for 15 minutes. Plates of 20 ml were poured at 50°C.

2.1.2.3 SOB (Super Optimal Broth)

SOB is a nutrient-rich broth used for bacterial growth. SOC is a variant of SOB which was used in the final step of bacterial cell transformation to enhance the efficiency (Table 2.3). The media was sterilised by autoclave at 121°C for 15 minutes.

Table 2.3 SOB and SOC composition (Hanahan, 1983).

Reagent	g/l
Peptone	20.0
Yeast extract	5.0
NaCl	0.58
KCl	0.19
$MgCl_2$+$MgSO_4$*	20 mM
Glucose#	20 mM

*From a stock solution of Mg^{2+} (1 M $MgCl_2$ + 1 $MgSO_4$)

only in SOC

2.1.2.4 Congo red agar plates

Tryptone, yeast extract and agar were dissolved in 1 litre of deionised water and autoclaved at 121°C for 15 minutes (Table 2.4). Filter-sterilised (0.22 µm diameter) Congo red and Coomassie blue dyes were added when the temperature dropped to 55°C. Plates of 20 ml were poured at 50°C.

Table 2.4 Composition of Congo red agar.

Reagent	Final concentration
Tryptone	10.0 g/l
Yeast extract	5.0 g/l
Agar	15.0 g/l
Congo red dye	40.0 µg/ml
Coomassie blue dye	20.0 µg/ml

2.1.2.5 Solutions for preparation of large plasmids using the Kado-Liu method.

The E buffer (Table 2.5) and lysis buffer (Table 2.6) were used to perform the plasmid extraction according to the Kado-Liu method (Kado and Liu, 1981).

Table 2.5 E buffer (pH 8.0). The solution is sterilised by autoclave.

Reagent	Amount
1 M Tris, pH 8.0	5 ml
0.5 M EDTA, pH 8.0	0.2 ml
dH_2O	Up to 100 ml

Table 2.6 Lysis buffer (sterilised by filtration).

Reagent	Amount
Tris powder	1.211 g
SDS	6.0 g
NaOH*	8.2 ml,
dH_2O	Up to 200 ml

*From fresh 2 M NaOH solution

2.1.2.6 TAE buffer

TAE (Tris-acetate-EDTA) buffer was used for gel electrophoresis. A 50X stock was prepared then used at a final concentration of 1X.

Reagent	Concentration (50X)
Tris base	242.2 g/l
EDTA sodium salt dihydrate	18.612 g/l
Acid acetic	60.5 ml/l
dH_2O	Up to 1 litre

2.1.2.7 Antimicrobial solutions

Stock of sterile solutions of antibiotics were prepared and used at the following final concentrations: chloramphenicol (25.0 µg/ml), kanamycin (50.00 µg/ml), nalidixic acid (50.00 µg/ml)

2.1.3 Growth conditions

For routine culture, strains were grown in 5.0 or 10.0 ml of LB Broth at 37 °C with shaking (200 rpm) for 16-17 hours.

2.1.4 Allelic exchange by recombineering methodology

Site directed mutagenesis by allelic exchange with the *cat* or *aph* antimicrobial resistance genes was achieved using a recombination-based system which employs lambda-red recombinase (Datsenko and Wanner, 2000). The λ-red recombineering methodology was used to make targeted insertion in a nucleic acid whose sequence was known. In particular, the methods was used to generate mutants strains of *S*. 4,[5],12:i:- isolate 0343A harbouring the *cat* or *aph* cassette within the mTmV prophage (Chapter 5).

Briefly, the microbial cells were transformed by electroporation (2.5 kV) with a DNA fragment composed of either the *cat* or *aph* gene flanked by sequence identical to the region of the chromosome that was to be replaced with the resistance gene, into the target strain containing pSim18, that encoded λ-Red recombinase (Chan et al., 2007). Recombinant strains in which the target gene was replaced with the resistance gene were selected on LB agar containing either chloramphenicol or kanamycin (Figure 2.1).

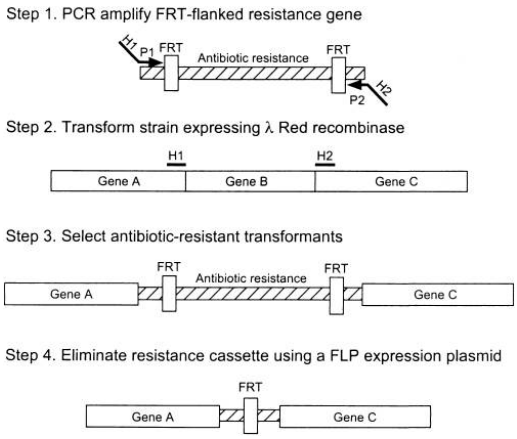

Figure 2.1 Schematic representation of the mutagenesis method based on the λ-red recombination system (Datsenko & Wanner 2000). The fourth step was not performed in the current study.

2.1.4.1 Preparation of the PCR amplicon

The PCR product containing the resistance gene flanked by homologous sequence was amplified using pKD3 (*cat* gene conferring resistance to chloramphenicol) or pKD4 (*aph* gene conferring resistance to kanamycin) as

template and oligonucleotides of 70 bases, of which the 50 bases at the 5' end were complementary to sequence flanking the desired insertion site, while the remaining are complimentary to the plasmid pKD3/pKD4 (Figure 2.2).

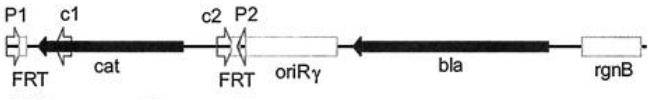

Figure 2.2 Linear representation of the pKD3 plasmid (Datsenko & Wanner 2000). P1 and P2 are the priming sites used for the amplification of the chloramphenicol cassette (*cat*).

The PCR was performed using the mTmV_mut oligonucleotides (Table 2.7)and the Q5® high-fidelity DNA polymerase (New England BioLabs), following the manufacturer's instructions:

Reagent	Amount
5X Q5 Reaction Buffer	5 µl
10 mM dNTPs	0.5 µl
10 µM Forward Primer	1.25 µl
10 µM Reverse Primer	1.25 µl
Template DNA	Variable
Q5 DNA Polymerase	0.25 µl
Nuclease-Free Water	To 25 µl

The PCR cycle program comprised an initial denaturation of the nucleic acid at 95°C for 300 seconds, followed by 35 cycles of amplification consisting on DNA denaturation at 94°C for 45 seconds, primers annealing at 45°C for 30 seconds and extension of the growing DNA at 72°C for 120 seconds. A final elongation step at 72°C for 300 seconds was included to ensure the extension of any remaining single stranded molecules. The PCR product was then separated by agarose gel electrophoresis, the band of the predicted size excised and extracted using the QIAquick Gel Extraction kit (QIAGEN).

2.1.4.2 Transformation with the pSim18 plasmid

The microbial strain was then transformed with the pSim18 by electroporation. Briefly, fresh electrocompetent cells were generated by growing the microorganisms in a flask with 50 ml of LB broth (37°C, 200 rpm shaking) to an $OD600_{nm}$ of 0.6, followed by ice-cold washes in 25 ml and 10 ml of cold sterile

UP water. After the last wash the cells were resuspended in 1.5 ml of cold sterile 10% glycerol and kept on ice; 1-5 µl of plasmid DNA were added to 100 µl of electrocompetent cells, which were then transferred into a cooled electroporation cuvette and electroporated (Bio-Rad MicroPulser™ at preset protocols, 2.5 Kv voltage, 200 Ω resistance, 25 uF capacitance, with a path length 2 mm). The cells were recovered by adding 0.95 mL of SOC and incubated at 30°C for 1 hour before the plating on hygromycin agar plates (Fast-Media® Hygro, InvivoGen).

2.1.4.3 Transformation with the linear PCR amplicon

Bacteria were transformed with the PCR product by electroporation. An overnight culture of the strain, grown in 10 ml of TBH broth (Fast-Media® Hygro TB, InvivoGen) at 30°C with shaking (200rpm) was used to inoculate 150 ml of the same medium. The fresh culture was grown at 30°C with shaking to an approximate OD_{600nm} of 0.3. When ready, the cells were heat shocked at 42°C for 15 minutes, then put immediately on ice for 10 minutes and centrifuged at 4000 rpm for 10 min at 4°C. They were resuspended in 75 ml of sterile ice-cold PBS a first time, washed in 1.5 ml five times and finally resuspended in 150 µl of sterile ice-cold PBS. A volume of 40 µl of bacteria was mixed with the PCR product (up to 10 µl, a function of the concentration) and subsequently electroporated (Bio-Rad MicroPulser™ Electroporator). Finally, 400 µl of SOC was added straight away to the cells, which were allowed to recover at 37°C for 2 hours and 30 minutes, before being plated in the selective medium and incubated at 37°C overnight.

2.1.4.4 Screening and selection of the mutants

The colonies that grew on selective media were screened by PCR amplification of the target gene sequence using combinations of primers (Check_ext, cat cassette and kan cassette oligonucleotides reported in Table 2.7) that annealed within the antibiotic resistance gene and flanking sequence. The reaction was performed with the GoTaq® DNA polymerase (Promega) in accordance with the manufacturer's instructions:

Component	Final Volume	Final Concentration
5X Green GoTaq® Reaction Buffer	10 µl	1X (1.5 mM $MgCl_2$)
dNTPs Mix, 10 mM each	1.0 µl	0.2 mM each dNTP
Primer Forward	X µl	0.1 – 1.0 µM
Primer Reverse	Y µl	0.1 – 1.0 µM
GoTaq® DNA polymerase (5u/ µl)	0.25 µl	1.25 u
Template DNA	Z µl	< 0.5 µg/50µl
Nuclease-free Water	Up to 50.00 µl	

After an initial denaturation of the DNA at 95°C for 300 seconds, 35 cycles of amplification were performed, consisting of denaturation at 94°C for 45 seconds, annealing at 56°C for 30 seconds and elongation at 72°C for 120 seconds. A final elongation step at 72°C for 300 seconds was included.

2.1.4.5 Curing pSim18 plasmid and verification

The colonies positive for the PCR screening were grown overnight in 5ml of LB broth supplemented with either chloramphenicol (25µg/ml) or kanamycin (50µg/ml) at 37°C with shaking in order to enrich for bacteria that had lost plasmid pSim18.

To check the effective plasmid loss, single colonies, obtained after the plating of serial dilutions, were streaked both on LB agar plates with either chloramphenicol (25 µg/ml) or kanamycin (50 µg/ml) and in hygromycin-supplemented agar plates and incubated overnight at 37°C, the first, and at 30°C, the latter. Colonies that grew on chloramphenicol or kanamycin plates, but not in presence of hygromycin, were selected. Microbial stocks were then prepared and conserved at -80°C in 20% glycerol.

2.1.5 PCR and oligonucleotides used

PCR was performed to screen mutant strains, check for the presence of plasmids and investigate the insertion site of a prophage in Chapters 3, 5 and 6. The oligonucleotides used in the study are listed in Table 2.7.

Table 2.7 List of oligonucleotides used in this study.

Name	Forward	Reverse
mTmV_mut	AATATAGATATGTAGTAGAGA CTGCTGCTATATGTTATATAG CAGCAATGGTGTAGGCTGGA GCTGCTTCG	CCGTCATTTACCAGTGCTAAAATT ATACACTCAACCATCAAAAAAATA GC CATATGAATATCCTCCTTAGT
Check_ext	CCAGACTGGCCGGAGAAA	GGAACTAACTCCGGAGCTAAAC
cat cassette	ACAAACGGCATGATGAACCT	GCACAAGTTTTATCCGGCCT
kan cassette	GAATGAACTGCAGGACGAGG	AGCAATATCACGGGTAGCCA
HI2	TTTCTCCTGAGTCACCTGTTA ACAC	GGCTCACTACCGTTGTCATCCT
attL	GCGTGAACACACCCTTCTCA	CGAAAGCGCCAACAGTAAGG
attR	CAAAACTGCCAACTTCCGCT	TCTCGTTTCGCTTCCATCGG
ddrA		GCCTAAATCGAGGATCGCCA
pS01569 _out_1	TACCCTTTCTTCCGCCTGTT	GTGGCATGTATCGGCTTGTT
pS01569 _out_2	TTCTTCCGCCTGTTCTGTGA	GGAGAACGAGGATATTGCGC
pS01569_int_right	GCTGTTTCGCCCATTCTCAT	CACCTTTGCTGACACGTTGA

2.1.6 Agarose gel electrophoresis

Gel electrophoresis was performed to check the products of the PCR reactions and to visualise the plasmid and genomic DNA extracted in Chapters 3, 5 and 6. Gels containing a concentration of agar ranging from 0.8% to 2.0% (w/v) depending on the size of the molecules to separate and containing 5% (v/v) Ethidium Bromide were made and run in TAE buffer. For size comparison, a 100 bp or 1 kb ladder (NEB) was incorporated into the run. The imaging was performed under exposure to UV light.

2.1.7 Genomic and plasmid DNA extraction

Genomic DNA extraction (Chapters 3 and 6) was performed from overnight liquid microbial culture with the Wizard® Genomic DNA Purification kit according to the manufacturer's instructions (Promega).

The extraction of plasmid DNA (Chapter 3) based on a method suitable for the extraction of large plasmids consisting of cell lysis under alkaline conditions followed by phenol chloroform extraction in order to remove cell debris and proteins (Kado and Liu, 1981). Pelleted cells from 1-1.5 ml of an overnight culture were resuspended in 20 µl of E buffer (Table 2.5) and vortexed briefly before the addiction of 150 µl of lysis buffer (Table 2.6). The mixture was then incubated at

55°C for 30 minutes, after which 150 µl of phenol:chlorophorm:isoamyl alcohol was added and the tube inverted repeatedly in order to create an emulsion. A centrifugation step (5 minutes, 13000 rpm, room temperature) was performed to break the emulsion. After incubation in the fridge for 30 minutes, the upper aqueous phase containing the plasmid was collected and stored in the freezer at -20°C. To check the extraction, a few microliters of plasmid were run in agarose gel. DNA purification was performed to remove residues of phenol chloroform from plasmid DNA which may interfere with downstream applications, such as PCR.

Small plasmids (up to 50 kb) were extracted from overnight cultures of donor strains cultured in liquid medium (supplemented with antibiotics, as appropriate), with shaking at 37°C (or 30°C for plasmids with temperature sensitive replicons such as pSIM18), using the QIAprep Spin Miniprep plasmid kit (QIAGEN). The plasmids prepared comprised pSIM18 (6.8 kb), pKD3 (2.8 kb) and pKD4 (3.2 kb) which were extracted from *S*. Typhimurium strain SL1344 pSim18, *E. coli* BW25141 pKD3 and E. coli BW25141 pKD4, respectively. DNA species were separated by horizontal agarose gel electrophoresis and the concentration of nucleic acids measured using the NanoDrop 2000 (Thermo Scientific).

2.1.8 DNA purification

2.1.8.1 Chloroform extraction

The extraction with chloroform was performed to remove the residues of phenol present in the solution after the extraction of the nucleic acids (Chapter 3). To this end, an equal volume of chloroform:isoamyl alcohol was added to the DNA solution, then it was mixed by vortex and centrifuged at full speed for 5 minutes. The aqueous upper part was transferred in a new microtube.

2.1.8.2 Ethanol precipitation

The ethanol precipitation enabled to concentrate and desalt the extracted nucleic acids. Sodium acetate 3.0 M (1/10 of the recovered volume) and ice-cold absolute ethanol (2.5 volumes) were added to the DNA solution, which was then incubated at -20°C overnight. The following day, the solution was centrifuged for 30 minutes at 4°C, and the supernatant aspired and washed with cold 70% ethanol. After an additional centrifugation, the ethanol was taken out and the pellet containing the DNA was air-dried and resuspended in the desired volume of water or TRIS

buffer. The DNA was kept in the fridge at 4°C or frozen at -20°C for long-term storage. The method was used in Chapter 3.

2.1.9 Plasmid conjugation

Conjugation experiments were performed to test the ability of plasmids belonging to IncHI2 incompatibility group to be self-transmissible (Chapter 3). To this end, donor and recipient (IR715) strains were cultured for 18 hours in 5 ml of LB broth at 37°C without shaking. Cells were harvested by centrifugation, re-suspended in phosphate buffered saline pH7 after which they were mixed in 5 ml LB broth to OD_{600nm} of 0.1 in a 1:1 ratio and incubated without shaking for 24 hours at 26°C. The conjugation of IncHI2 plasmids is, in fact, thermosensitive: it is inhibited at 37°C but promoted at 22°C -30°C (Phan and Wain, 2008). Ex-conjugants were enumerated on LB agar supplemented with nalidixic acid (50 µg/ml) or nalidixic acid (50 µg/ml)/chloramphenicol (25 µg/ml) and incubated at 30°C for 18 hours, and CFU enumerated. The acquisition of the plasmids was confirmed by PCR with the IncHI2 oligonucleotides previously published (Carattoli et al., 2005).

2.1.10 Determination of the minimum inhibitory concentration (MIC)

The determination of the MIC of the recipient and transconjugant cells was performed to phenotypically assess the acquisition of antimicrobial resistance encoded on the IncHI2 plasmids (Chapter 3). To this end, a microdilution assay was performed as previously described (Petrovska et al., 2016). An overnight culture grown at 37°C with shaking (200 rpm) was harvested and resuspended in sterile phosphate-buffered saline (PBS) and adjusted to an OD_{600nm} of 1.0. Then, 2-fold dilutions of antibiotic were prepared in LB broth, of which 200 µl were dispensed in a single well of a microtiter plate and inoculated with 15 µl of the adjusted culture. The plate was then incubated at 37°C for 24 hours, after which the optical density was measured. The MIC was the concentration which reduced the growth by 80% or more.

2.1.11 Congo red agar plating

The Congo red phenotypic assay was performed to test the ability of the isolates to produce biofilms in Chapter 3. The microbial strains were streaked and a few µl of an overnight microbial culture were spotted onto agar plates supplemented with Congo red dye (Table 2.4). The plates were then incubated at 28°C for 72 hours before visual inspection (O'Leary et al., 2013).

2.1.12 Pellicle formation assay

The ability of the isolates to produce pellicles in the air-liquid interface was tested by inoculating each strain in glass tubes containing 5 ml of LB broth without salts. The tubes were incubated at 22°C for 72 hours without shaking after which the formation of a pellicle was visually evaluated. The type of pellicle produced was classified either as strong and rigid or fragile basing on its ability to resist tube shaking (O'Leary et al., 2013). The method was used in Chapter 3.

2.1.13 Microtiter plate assay

The ability of the isolates to produce biofilms on polystyrene was assessed performing the microtiter plate assay (Chapter 3). The protocol used was previously described (Walia et al., 2017). A pre-culture from a single colony was grown in 5ml LB without salt at 37°C with shaking (150 rpm) for 6/7 hours, after which it was diluted 1:100 in 5 ml LB broth without salt and incubated overnight at 37 °C under static conditions. Subsequently, the OD_{600nm} of the culture was adjusted to 0.02 using LB broth without salt, and 200 µl aliquots were dispensed in eight wells (8 technical replicates per biological replicate) of a microtiter plate. The 96-well plates were then incubated at 22°C and 37°C without shaking for 24 or 48 hours, after which the supernatant was removed, and the cells washed with 200 µl of tryptone salt (Oxoid Limited), before fixing with 300 µl of pure ethanol for 20 minutes. Biomass was stained with crystal violet and the OD was read at 595 nm. At least two biological replicates were performed.

2.1.14 Assessment of prophage excision and frequency of lysogeny in susceptible microbial isolate

The assay was performed to test the ability of the mTmV prophage to excise and form active viral particles by infecting and integrating into a susceptible host (Chapter 5). To this end, the supernatant of an overnight culture of an isolate carrying the tagged prophage and containing the free phage particles was mixed with the recipient isolate and the lysogeny monitored by growth on selective media. In detail, strains 0343A mTmV_Cm/0343A mTmV_Kan were grown in 5.0 ml LB broth at 37°C for 16-17 hours with shaking (200 rpm), after which the cells were pelleted by centrifugation (10 minutes at 4,000 rpm) and the supernatant collected and filter sterilised (0.22 µm pore size). The sterility of the filtered supernatant was checked by assessing the lack of microbial growth on LB agar plates spread with 100 µl of the filtered supernatant and incubated overnight at

37°C. The recipient isolate, previously grown at the same conditions as the donor strains, was then adjusted to an OD_{600nm} of 0.01 in falcon tubes with 4 ml LB broth and 1 ml of sterile supernatant and incubated at 37°C for 24 hours with shaking (100 rpm). Serial dilutions and plating in LB agar plates with and without antimicrobials were carried out, followed by overnight incubation at 37°C. The rate of lysogeny was calculated as the ratio between concentration (CFU/ml) of lysogen cells grown on selective media and concentration of total recipient cells grown on non-selective media. Finally, Lysogens were checked for the integration of mTmV prophage by PCR using the attL, attR and ddrA oligonucleotides.

2.2 Bioinformatics and genomics methods

2.2.1 Whole-genome sequencing (WGS)

In Chapter 3, whole-genome sequencing (WGS) was performed at the Earlham Institute (EI). Sequencing libraries were prepared using the Low Input Transposase-Enabled (LITE) methodology and 125 bp paired-end reads were generated using the Illumina HiSeq 2500 with version 4 chemistry and in High Output mode according to manufacturer's instructions (Illumina).

In Chapter 5, WGS was performed in-house at the Quadram Institute Bioscience (QIB); libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina, California, USA) and 150 bp paired-end reads were generative with the Illumina NextSeq 500 system (Illumina, California, USA) according to manufacturer's instructions.

Long-read WGS was performed using the PacBio technology at EI and the sequences were assembled as previously described (Bawn et al., 2019).

2.2.2 Construction of maximum-likelihood (ML) phylogenetic tree

Maximum-likelihood phylogenetic trees described in Chapters 3, 4 and 6 were constructed using single nucleotide polymorphisms (SNPs) in the core genome, namely in sequences present in all the isolates considered, of either whole genome or plasmid sequences. The SNPs were identified using the snippy v3.0 (Seemann, 2015) software which aligned the raw reads to a reference sequence using BWA-MEM (Li and Durbin, 2009) and subsequently performed variant calling and SNP filtering with FreeBayes (Garrison and Marth, 2012) and vcflib/vcftools (Danecek et al., 2011). The removal of SNPs in recombinant regions was performed using Gubbins v2.3.4 which scans for the presence of regions with a high density of point mutations (Croucher et al., 2015). Finally,

RAxML v8.2.12 was used to generate the ML phylogenetic tree (Stamatakis, 2014). Rapid bootstrapping was performed with 450 replicates. The phylogenetic tree was visualised, manipulated and annotated using the R package ggtree (Yu et al., 2017) and the online tool iTOL (Letunic and Bork, 2019).

2.2.3 Phylogeny of *sopE* gene

The phylogeny of the *sopE* gene (Chapter 4) based on nucleotide variation at synonymous sites in the coding sequence was generated using the SNAP software (http://www.hiv.lanl.gov; Korber, 2000) as previously described (Mirold et al., 2001b). The sequences of *sopE* came from the following isolates: *S*. *enterica* subspecies IV SARC9 (accession number AF378111), *S*. *enterica* subspecies VII SARC16 (accession number AF378112), *S*. *enterica* subspecies I serovar Hadar strain X3230 (accession number AY034828), Enteritidis strain ATCC BAA-708 (accession number NZ_CP025554.1), Dublin strains 3149B and 3108A (this study), Typhi strains SARC2 (accession number AF378115) and X3744 (accession number AF153829.1), Typhimurium strain SL1344 (accession number AF043239.1), strain H142780372 (SRA accession number SRR1965151), strain H142940536 (SRA accession number SRR3049928), strain H143660450 (SRA accession number SRR1958481), strain H143740716 (SRA accession number SRR1970113) and strain H142520441 (SRA accession number SRR3049530); *S*. 4,[5],12:i:- strain S04698-09 (accession number NZ_LN999997.1).

2.2.4 Phylogeny of bacteriophages and prophages using VICTOR

VICTOR (Meier-Kolthoff and Goker, 2017) was used to generate the whole genome-based phylogeny of phages/prophages described in Chapter 6. The software used the Genome BLAST Distance Phylogeny (GDBP) approach in which BLAST was used to identify high-scoring segment pairs (HSPs), namely subsegments of pairs of sequences which share a high level of similarity, between each genome. Genome distance matrices were then computed and used by FastME 2.1.4 (Lefort et al., 2015) to infer the phylogenetic tree. In addition, the software used OPTSIL (Goker et al., 2009) to perform a taxonomic classification aimed at clustering of the bacteriophages at species, genus and family levels. OPTSIL optimised the distance thresholds gained from an in-depth statistical analysis performed on a large ICTV (International Committee on Taxonomy of Viruses) reference dataset.

2.2.5 Time-dependent ancestral state reconstruction (BactDating)

BactDating was used to generate a time-dependent phylogeny (Didelot et al., 2018) (Chapter 4). The software uses a Bayesian inference approach combined with a Markov chain Monte Carlo (MCMC) methodology to estimate the molecular clock and the coalescence rate in order to infer a time-dependent phylogeny from a provided rooted phylogeny, previously generated by the maximum likelihood method described above. Before constructing the dated phylogeny, the presence of a molecular clock signal within the collection of isolates included in the phylogenetic analysis was tested by evaluating the correlation of the root-to-tip distance with the isolation dates by regression analysis. Inference was performed under a mixedgamma model with 10^6 MCMC iterations to ensure convergence. The effective sample size of the α (coalescent time unit), μ (mean substitution rate) and σ (standard deviation of the per-branch substitution rates) parameters was > 200.

2.2.6 Ancestral state reconstruction (PastML)

PastML was used in Chapter 4 to infer ancestral states on a rooted maximum likelihood phylogenetic tree using the MPPA (marginal posterior probabilities approximation) method, a maximum-likelihood method which chose the state for each node based on the Brier score (Ishikawa et al., 2019).

2.2.7 Comparative genomics

Comparative genomic analyses described in Chapters 5 and 6 were carried out by aligning the nucleotide sequences of interest using the BLAST algorithm (Altschul et al., 1990) then plotting the comparison using the R package genoPlotR (Guy et al., 2010).

2.2.8 Detection of sequences of interest in whole genome raw reads

In Chapters 3, 4, 5 and 6 the presence of sequences of interest in whole genome sequences was investigated using either SRST2 v0.1.7 (Short Read Sequence Typing) (Inouye et al., 2014) or ARIBA (Antimicrobial Resistance Identification By Assembly) (Hunt et al., 2017b). SRST2 mapped next generation sequence short reads to the sequence provided as reference using Bowtie2 (Langmead and Salzberg, 2012), the alignment was scored and a report was then generated. Matches with >90% coverage and <10% sequence divergence were reported as present. ARIBA combined the mapping of the paired-end reads to an input reference with the local assembly. The reference sequences were first clustered

by similarity using CD-HIT (Fu et al., 2012). The reads were then mapped and assembled using minimap and fermi-lite, respectively (Li, 2016, Li, 2015). The closest reference was then identified and the completeness of the assembly evaluated with nucmer (MUMmer package) (Kurtz et al., 2004). Finally, variants were identified by mapping the reads to the assembly (Bowtie2 and SAMtools) and comparing the assembly to the closest reference (MUMmer) (Langmead and Salzberg, 2012, Kurtz et al., 2004, Li et al., 2009). A detailed report comprising the gene presence, its completeness as well as the presence of SNPs/indels and their effects was then generated.

SRST2 and ARIBA were used either with custom databases or with the public databases ResFinder, PlasmidFinder, vfdb and INTEGRALL to *in silico* detect the presence of antibiotic resistance genes, plasmids, virulence genes and integrons, respectively (Zankari et al., 2012, Carattoli et al., 2014, Chen et al., 2016, Moura et al., 2009).

2.2.9 Blast Ring Image Generator (BRIG)

Blast ring image generator (BRIG) enabled us to display comparisons of multiple plamids in Chapter 3. The tool relies on BLAST which performed the alignment against the reference sequence, then a circular comparison image was generated. Each query was represented as a ring, and a colour gradient reflected the percentage identity to the reference (Alikhan et al., 2011).

2.3 Statistical analyses

Chi-squared tests (Chapter 3) were performed to test the null hypothesis that two categorical variables were independent by measuring how well the observed distribution of the data fitted the expected distribution if the variables were independent. The Chi-squared tests were carried out using the chisq.test function in the core “stats” package of R (R Core Team, 2019) with the simulate.p.value option; 2,000 Monte Carlo permutations were performed for each test to obtain empirical p-values.

The Mann-Whitney U test and the Fisher’s exact test conducted in Chapters 3 and 4 were performed using GraphPad Prism v5.04. The Mann-Whitney U test was performed to test the null hypothesis that the distribution of two unpaired groups were identical, while the Fisher’s exact test was carried out to determine whether there was a non-random association between two categorical variables.

In Chapter 4, permutation tests comprising 5,000 simulations were performed in R (R Core Team, 2019) to test the null hypothesis that the statistically significant difference in the patristic distance between groups of isolates was independent from the distribution of *sopE* across the phylogeny. The p-value was calculated as how many times the simulated difference exceeded the observed one.

3 Molecular epidemiology and phenotypic characterisation of *S*. Typhimurium and *S*. 4,[5],12:i:- from Irish pig farms

3.1 Introduction

Pork and pork products represent a major source of salmonellosis (ECDC, 2018). The colonisation of pigs represents a major risk factor for the contamination of pig meat. Historically, *S*. Typhimurium was the most common *S*. *enterica* serovar associated with pig production (Miller et al., 2011, Visscher et al., 2011, Gebreyes et al., 2006). Nevertheless, MDR monophasic variant of S. Typhimurium (*S*. 4,[5],12:i:-) ST34 is currently the predominant *Salmonella* serovar isolated from pigs and pig meat in the EU (Anonymous, 2010, ECDC, 2018).

The acquisition of a novel 80 kb genetic element encoding resistance to multiple heavy metals, termed *Salmonella* genomic island 4 (SGI-4), contributed to the emergence of *S*. 4,[5],12:i:- ST34 (Petrovska et al., 2016). In particular, SGI-4 enhances resistance to copper, used as a growth promoter and antimicrobial in swine husbandry, suggesting it confers a fitness advantage in the ecological niche represented by pigs (Branchu et al., 2019, Petrovska et al., 2016). Despite its high clonality, a certain level of microevolution has been described for *S*. 4,[5],12:i:- ST34. The acquisition of the *sopE* virulence gene following lysogenic conversion of a novel temperate phage (mTmV) contributed in generating variation in *S*. 4,[5],12:i:- ST34 (Petrovska et al., 2016). SopE is a guanine nucleotide exchange factor (GEF) which activates Rac-1 and Cdc42, promoting host cell invasion and gut inflammation (Friebel et al., 2001, Lopez et al., 2012). The *sopE* gene, rarely identified in *S*. Typhimurium, was previously reported in epidemic clones of *S*. Typhimurium DT204 suggesting a selective advantage conferred by the gene (Mirold et al., 1999).

S. 4,[5],12:i:- ST34 replaced the previous dominant epidemic clone *S*. Typhimurium DT104. In Ireland, a rapid clonal replacement of *S*. Typhimurium DT104 occurred in pig herds between 2010 and 2012, when *S*. 4,[5],12:i:- ST34 was first isolated, and accounted for over the 40% of total *Salmonella* isolates (Rowe et al., 2003, Burns, 2015). An additional MDR clone epidemic in pigs, but sporadically isolated in other livestock animals or responsible for human salmonellosis, is S. Typhimurium U288, which emerged in the UK in 2001 (Hooton et al., 2013).

A common characteristic of the epidemic clonal groups is resistance to multiple antimicrobials. *S*. Typhimurium DT104, which was epidemic in Europe between 1990-2010, exhibited resistance to ampicillin, chloramphenicol, streptomycin, sulphonamide and tetracycline (ACSSuT) on *Salmonella* genomic island 1 (SGI-1), which was integrated in the chromosome (Lucarelli et al., 2010). *S*. 4,[5],12:i:- is commonly characterised by the tetra-resistant pattern ASSuT, the AMR genes are located in two regions (RR1 and RR2) on a composite transposon integrated in the phase 2 flagellar antigen (Petrovska et al., 2016). The ACSSuTTm (including trimethoprim) resistance profile associated with *S*. Typhimurium U288 is conferred by two independent insertions into a pSLT-like plasmid (Hooton et al., 2014).

In order to increase the food safety of pork products, practices aimed at reducing the prevalence of *Salmonella* in pigs on farms are performed, including cleaning and disinfection between production cycles (Andres and Davies, 2015). Although cleaning procedures are effective in reducing contamination, *Salmonella* can still persist in the farm environment (Martelli et al., 2017). As biofilms confer protection to a variety of environmental stressors, biofilm-forming ability is considered important for the environmental persistence (Vestby et al., 2009). The majority of *S*. 4,[5],12:i:- isolated in Portugal from multiple sources (clinical, animal and environmental samples) were able to produce biofilms on Congo red agar, plastics and glass (Seixas et al., 2014). Notably, the ability of the isolates to produce biofilms increased over time suggesting a possible selection for this phenotype. However, the belonging of the isolates to the *S*. 4,[5],12:i:- ST34 clonal complex was not assessed.

Here I performed a molecular epidemiology study to determine the phylogenetic relationship of *S*. 4,[5],12:i:- and *S*. Typhimurium isolates from Irish pig farms with epidemic clones circulating in the UK and in Europe. The persistence of *Salmonella* on pig farms was studied by evaluating the phylogenetic relatedness of the bacteria isolated from consecutive production cycles and the occurrence of microevolution events in persistent clones was investigated. The biofilm-forming ability of the isolates was also assessed to test if any correlation with environmental persistence existed.

3.2 Results

3.2.1 Collection of *S*. Typhimurium and *S*. 4,[5],12:i:- isolates from Irish pig farms and feed mills

The isolates used in the study comprised 104 *S*. 4,[5],12:i:- and 34 *S*. Typhimurium isolated from a previous investigation (Burns, 2015) on pig farms located across three provinces of the Republic of Ireland and feed mills (Table 3.1 and Appendix 1). Briefly, ten farrow-to-finish pig farms (farms A-J) with a history of high *Salmonella* prevalence (>50% of the herd was seropositive for *Salmonella*) were sampled twice between March 2012 and June 2013. Samples were collected across all production stages and from faeces, feed, water and the farm environment including floors, walls, water drinkers and troughs. *S*. Typhimurium or *S*. 4,[5],12:i:- was isolated from all farms except farm E. Therefore, isolates from nine farms were included in this study. Isolation from feed from three commercial feed mills (mills B, C and D) and one home compounder (mill E) were described previously (Burns et al., 2015). The susceptibility of the isolates to antimicrobials was assessed performing the broth dilution method of the Clinical and Laboratory Standards Institute (formerly NCCLS) (CLSI, 2008). The most common antimicrobial resistance (AMR) profiles are reported in Table 3.1.

Table 3.1 Farm/feed mill, serotype, production stage, source of isolation and antimicrobial resistance profile of the *S*. 4,[5],12:i:- (mSTm: monophasic *S*. Typhimurium) and *S*. Typhimurium (STm) isolates used in the study. The table reports the number of isolates of each serotype for each the farm, the stage of production and the number of isolates from pig faeces, environmental samples (env), feed and water as well as the most common AMR profiles. The stages of pig production included: farrowing (FW), gilts (G), 1st stage weaner (W1), 2nd stage weaner (W2), dry sow (D) and finisher (F). Environmental samples comprised swabs from the pen, water drinkers, feed troughs and feed bins. AMR genes abbreviations: ampicillin (A), azithromycin (Az), ceftriaxone (Ce), chloramphenicol (C), ciprofloxacin (Cp), gentamycin (Gm), nalidixic acid (Na), streptomycin (S), sulphonamide (Su), tetracycline (T), trimethoprim (Tm).

Farm	Serotype (n)	Stage	Faeces	Env	Feed	Water	AMR profile
A	mSTm (41)	FW,W1,W2, G,F	14	22	3	2	ASSuT (20) ASSuTTm (12) Others (9)
	STm (0)	-	-	-	-	-	-
B	mSTm (34)	W1,W2, F	8	19	2	5	ASSuTTm (23) ASSuT (5) Others (6)
	STm (0)	-	-	-	-	-	-
C	mSTm (2)	F	-	2	-	-	ASSuT (2)
	STm (0)	-	-	-	-	-	-
D	mSTm (1)	D	-	1	-	-	T (1)
	STm (5)	FW,W1,G,F	-	2	2	1	ACSSuT (2) ACSuT (2) Other (1)
F	mSTm (0)	-	-	-	-	-	-
	STm (1)	W2	-	1	-	-	ASSuTNa (1)
G	mSTm (3)	G,F	-	1	-	2	T (3)
	STm (0)	W1,W2,D,G, G	3	12	1	2	Susceptible (11) ASSuT (3) Others (4)
H	mSTm (1)	W1		1			ASSuTAzt (1)
	STm (0)	-	-	-	-	-	-
I	mSTm (0)	-	-	-	-	-	-
	STm (7)	G,F	-	5	-	2	T (7)

J	mSTm (17)	W1,W2,G	6	8	-	3	ACSSuTTmGm (8) ASSuT (4) Others (5)
	STm (3)	G,F		2		1	ACSSuTTm (2) Other (1)
mill B	mSTm (1)	-	-	-	1	-	ACSSuTTmGm (1)
mill C	mSTm (1)	-	-	-	1	-	ASSuT (1)
mill D	mSTm (2)	-	-	-	2	-	ACSSuTTmGm (1) ACSTCpCe (1)
mill E	mSTm (1)	-	-	-	1	-	TGm (1)

3.2.2 Farm-specific genotypes of *S*. Typhimurium and *S*. 4,[5],12:i:-, which was related to the UK epidemic, persisted across production cycles

A molecular phylogenetic analysis was performed in order to test the relatedness of *S*. Typhimurium and *S*. 4,[5],12:i:- circulating in Irish pig farms and feed mills with epidemic clones of *S*. Typhimurium and *S*. 4,[5],12:i:- and to investigate the population structure of *Salmonella* on pig farms. The whole genome sequences of 34 *S*. Typhimurium and 104 *S*. 4,[5],12:i:- isolates were determined and used to generate a maximum-likelihood phylogenetic tree with RAxML v8.2.12 (Stamatakis, 2014) based on 2,427 SNPs in the core genome identified in reference with *S*. Typhimurium SL1344 (Figure 3.1). The nucleotide substitution model used was the GTR and the log likelihood score obtained was -15986.1.

Sequences representative of the epidemics *S*. Typhimurium DT104 (strain NCTC13348) and U288 (strain S01960-05) and *S*. 4,[5],12:i:- DT193/120 (strain S04698-09) were included to provide context. Five major clades were identified (A-E). All the 104 *S*. 4,[5],12:i:-clustered in a single clade (clade A) and were closely related to S04698-09, a reference strain of the current pandemic *S*. 4,[5],12:i:- clone (Petrovska et al., 2016). At least one isolate from eight of the nine pig farms was in clade A.

In contrast, *S*. Typhimurium isolates were more diverse. Three isolates from two farms were related to the *S*. Typhimurium representative of U288 phage type (clade D), which is currently epidemic in pigs in the UK (Hooton et al., 2013), whereas six isolates from two farms were related to reference strain of the previous pandemic clone *S*. Typhimurium DT104 (clade E) (Threlfall, 2000,

Mather et al., 2013). The remaining *S*. Typhimurium formed two closely related clades (clades B and C), each containing isolates from a single farm.

The observation that isolates from some farms were more closely related than isolates from other farms potentially indicated the presence of farm-specific genotypes. In several cases, closely related isolates (<5 SNPs different) were isolated from multiple sources, comprising voided faeces, feed, water and pen environment, from multiple stages of pig production suggesting the widespread contamination of the same clones across the farm. In six farms related *S*. Typhimurium or *S*. 4,[5],12:i:- were isolated up to nine months apart, from successive production cycles, suggesting the persistence or the reintroduction of the same clone from a common source. However, examples of closely related isolates (< 5 SNPs in the core genome) isolated from distinct farms or feed mills indicating either contamination from a common source or transmission events were observed. Three *S*. 4,[5],12:i:- isolates from two feed mills (B and D) were closely related with an isolate from farm J. An isolate from mill C was closely related with three isolates from farm A. Identical isolates were isolated from farm D and mill E and from farms A and G. Similarly, indistinguishable isolates were isolated from farms A and B.

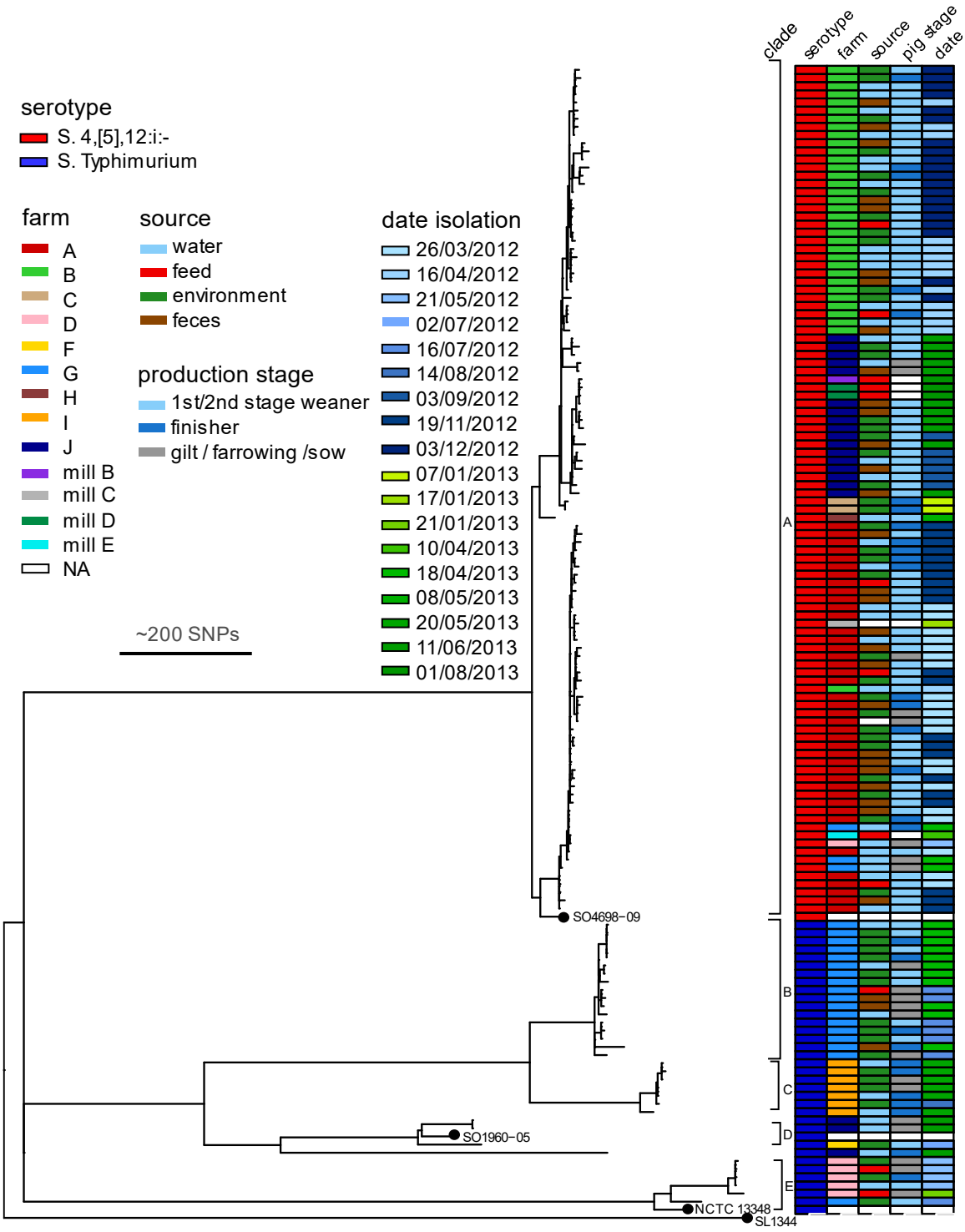

Figure 3.1 Maximum-likelihood phylogenetic tree of 34 *S*. Typhimurium and 104 *S*. 4,[5],12:i:- from Irish pig farms and feed mills based on 2,427 SNPs identified using *S*. Typhimurium SL1344 as a reference. The representative strains S04698-09 (*S*. 4,[5],12:i:- ST34), NCTC 13348 (*S*. Typhimurium DT104) and S01960-05 (*S*. Typhimurium U288) were included in the analysis. The serotype, farm, source of isolation, stage of pig production and date are reported.

3.2.3 Antimicrobial resistance (AMR) of *S*. 4,[5],12:i:- and *S*. Typhimurium

To investigate the repertoire and the flux of AMR genes in the Irish porcine *S*. Typhimurium and *S*. 4,[5],12:i:- isolates, the whole genome sequences were assessed for the presence of AMR genes and plasmid replicons genes. Sequences homologous to those in the ResFinder, INTEGRALL and PlasmidFinder databases (Zankari et al., 2012, Moura et al., 2009, Carattoli et al., 2014) were detected using SRST2 (Inouye et al., 2014).

A general concordance between the presence of AMR genes and the resistance profile previously determined *in vitro* was observed. Most sequences encoded multiple AMR genes (Figure 3.2). Only thirteen S. Typhimurium isolates, all in clade B, were pan-susceptible to all tested antimicrobials. The distribution of AMR genes associated with the phylogenetic clades (Chi-squared tests p-value 0.01 for *aadA1*, *aadA2*, *aph*, bla_{CARB}, bla_{TEM}, *cmI*, *dfrA1*, *dfrA12*, *dfrA14*, *mefA*, *sul1*, *sul2*, *tet(A)*, *tet(C)* and *tet(G)*; p-value 0.02 for *floR*, *mefB*, *msrD*, *strA*, *strB*, *sul3*, *qacE (Delta 1)*, *qacH*, *tet(B*)) suggesting that their acquisition occurred once, followed by clonal expansion (Figure 3.2 and Figure 3.3).

Isolates that were closely related to *S*. Typhimurium DT104 NCTC 13348 harboured the *aadA2* (streptomycin/spectinomycin), bla_{CARB} (ampicillin), *floR* (chloramphenicol/florfenicol), *sul1* (sulphonamides) and *tet(G)* (tetracycline) genes responsible for the ACSSUT profile common in DT104 clones (Boyd et al., 2002). The isolates also encoded the *qacE (Delta 1)* gene, conferring resistance to quaternary ammonium compounds (Boyd et al., 2002). Isolates closely related to S01560-05 (*S*. Typhimurium U288) were positive for the genes conferring resistance to ampicillin (bla_{TEM}), chloramphenicol (*cmI*) streptomycin (*aadA1*, *aadA2*, *strA*, *strB*), sulphonamide (*sul2*, *sul3*), tetracycline (*tet(A)*), trimethoprim (*dfrA12*), consistent with the ACSSuTTm profile (Hooton et al., 2014). In addition, these organisms harboured the *qacH*, *aph* and *mefB* genes, previously not described in U288 clones, involved in the resistance to quaternary ammonium compounds, aminoglycosides and macrolides. *S*. Typhimurium isolates from clades B and C were more susceptible to antimicrobials. The majority of isolates from clade B were pan-susceptible, with the exception of three isolates encoding resistance against ampicillin (bla_{TEM}), sulphonamides (*sul2*), streptomycin (*strA*/*strB*) and tetracycline (*tet(A)*). All the isolates from clade C encoded the resistance only to tetracycline (*tet(C)*). Most of S. 4,[5],12:i:- exhibited the tetra-

resistant profile ASSuT, conferred by *bla*$_{TEM}$, *strA*/*strB*, *sul2* and *tet(B)*, respectively, associated with the incQ replicon integrated in composite transposon on the chromosome, typical of this clonal group (Petrovska et al., 2016). However, a number of isolates partially lost the resistance possibly due to mobilisation of IS26 elements in the RR1 and RR2 regions located adjacent to the AMR genes (Lucarelli et al., 2012).

In conclusion, AMR was widespread and clade-specific in *S*. 4,[5],12:i;- and *S*. Typhimurium lineages previously (DT104) or currently associated with epidemics (*S*. 4,[5],12:i;- and U288).

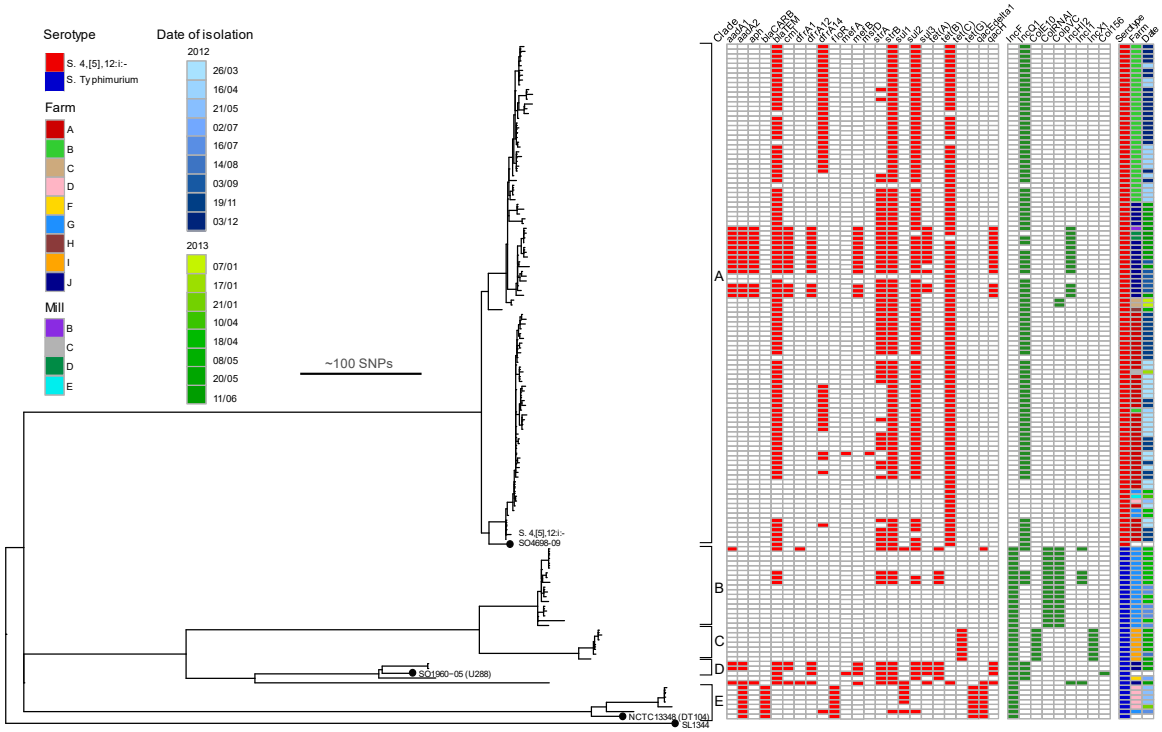

Figure 3.2 Presence of antimicrobial resistance genes and plasmid replicons in *S*. Typhimurium and *S*. 4,[5],12:i:- isolates. The presence of AMR genes (red blocks) and plasmid replicons (green blocks) was shown in the phylogenetic context of the isolates (the maximum-likelihood tree is the same described in Figure 3.1). The serotype, farm and date of isolation are also reported.

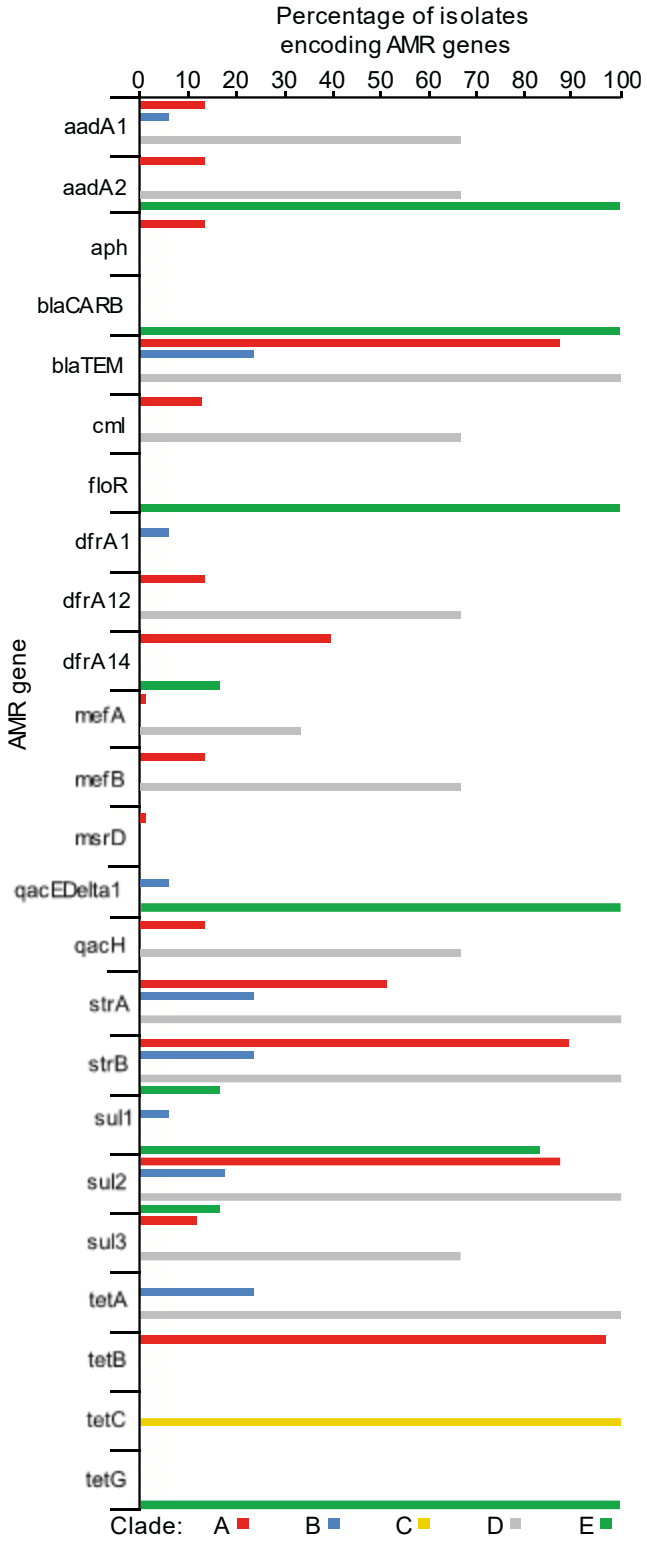

Figure 3.3 Distribution of AMR genes in the clades of the phylogenetic tree showed in Figure 3.2. Horizontal bars indicate the proportion of isolates from each clade positive for each AMR gene. Clade A: *S*. 4,[5],12:i:- (n=104), Clade B: *S*. Typhimurium (n=17); clade C: *S*. Typhimurium (n=7); clade D: *S*. Typhimurium U288 (n=3); clade E: *S*. Typhimurium DT104 (n=6). Genes identified were *aadA1/aadA2*: aminoglycosides resistance gene; *aph*: aminoglycosides resistance gene; *blaCARB*: carbenicillin/ampicillin resistance; *blaTEM*: penicillins/early cephalosporins resistance;

cmI: chloramphenicol resistance; *floR*: florfenicol/chloramphenicol resistance gene; *dfrA1*/*dfrA12*/*dfrA14*: trimethoprim resistance; *strA*/*strB*: streptomycin resistance genes; *sul1*/*sul2*/*sul3*: sulphonamide resistance genes; *tetA*/*tetB*/*tetC*/*tetG*: tetracycline resistance; *mefA*/*mefB*: macrolide, lincosamide and streptogramin B resistance; *msrD*: macrolide resistance; *qacEDelta1*/*qacH*: quaternary ammonium compounds.

3.2.4 The acquisition of plasmids expanded the AMR profile

The presence of AMR genes in two groups of isolates in clades A and B correlated with the detection of plasmid replicons. The absence of the plasmid in the related isolates, which were susceptible to these antimicrobials, suggested that the additional resistance was plasmid-borne. In the case of *S*. Typhimurium in clade B, there was evidence from the tree structure for the acquisition having occurred on the farm.

Clade B comprised sixteen *S*. Typhimurium isolates from a single farm, most of which were pan-susceptible, with the exception of four isolates encoding the bla_{TEM}, *strA*, *strB*, *sul2* and *tet(A)* AMR genes that correlated with the presence of IncI/IncQ replicons (Figure 3.2; Chi-squared test p-values: 0.001 for IncI/IncQ replicons and bla_{TEM}, *strA*, *strB* and *tet(A),* 0.005 IncI/IncQ and *sul2*). The fact that the MDR isolates were isolated on the second sampling date and were descendants of pan-susceptible *S*. Typhimurium isolated on the first sampling date suggested that a persisting clone acquired the plasmid-encoded resistance on the farm site.

In clade A, thirteen related *S*. 4,[5],12:i:- from farm J and two feed mills acquired an IncHI2 plasmid which correlated with the presence of eight additional resistance genes conferring resistance to aminoglycosides (*aadA1*/*aadA2*, *aph*), chloramphenicol (*cmI*), trimethoprim (*dfrA12*), macrolides (*mefB*), sulphonamides (*sul3*), and quaternary ammonium compounds (*qacH*) (Chi-squared test p-values 0.0005). Notably, a single *S*. Typhimurium isolate (from farm J) harboured an IncHI2 plasmid which may suggest the transfer of the plasmid between the distinct genotypes on the farm.

The sequences of the IncHI2 plasmids from the fourteen *S*. 4,[5],12:i:- and *S*. Typhimurium were then isolated and analysed in order to assess whether the additional resistance genes were plasmid-encoded. The alignment of the assembled plasmid sequences in GenBank revealed their homology with the IncHI2 plasmid pSTM6-275 reported in a *S*. 4,[5],12:i:- isolated from pigs in

Australia (Dyall-Smith et al., 2017) (Figure 3.4A). The IncHI2 plasmids in the Irish isolates harboured up to ten AMR genes and genes conferring resistance to heavy metals, such as copper (*pco* genes), silver (*sil* genes) and tellurium (*ter* genes) (Billman-Jacobe et al., 2018).

The relatedness of the plasmids in *S*. 4,[5],12:i:- and *S*. Typhimurium was then investigated in order to test the hypothesis of a possible transmission between the two genotypes. A maximum-likelihood phylogenetic tree based on 162 core-genome SNPs with reference to the pSTM6-275 plasmid was generated with RAxML v8.2.12 using the GTR model of nucleotide substitution under the Gamma model of rate heterogeneity (Figure 3.4B). The log likelihood score obtained was -748.130932.

The phylogenetic tree topology suggested that the IncHI2 plasmids were closely related; however, it did not reveal any direction of plasmid transmission from one genotype to the other, rather the presence of a recent common source.

The IncHI2 plasmids were capable of self-transmission *in vitro* by conjugation at a frequency of 10^{-3} per recipient. The acquisition of the IncHI2 plasmid was demonstrated to confer resistance to aminoglycosides, chloramphenicol, trimethoprim and sulphonamides. Transconjugants were subjected by PCR using oligonucleotides targeting sequences of IncHI2 replicons (data not shown) (Carattoli et al., 2005).

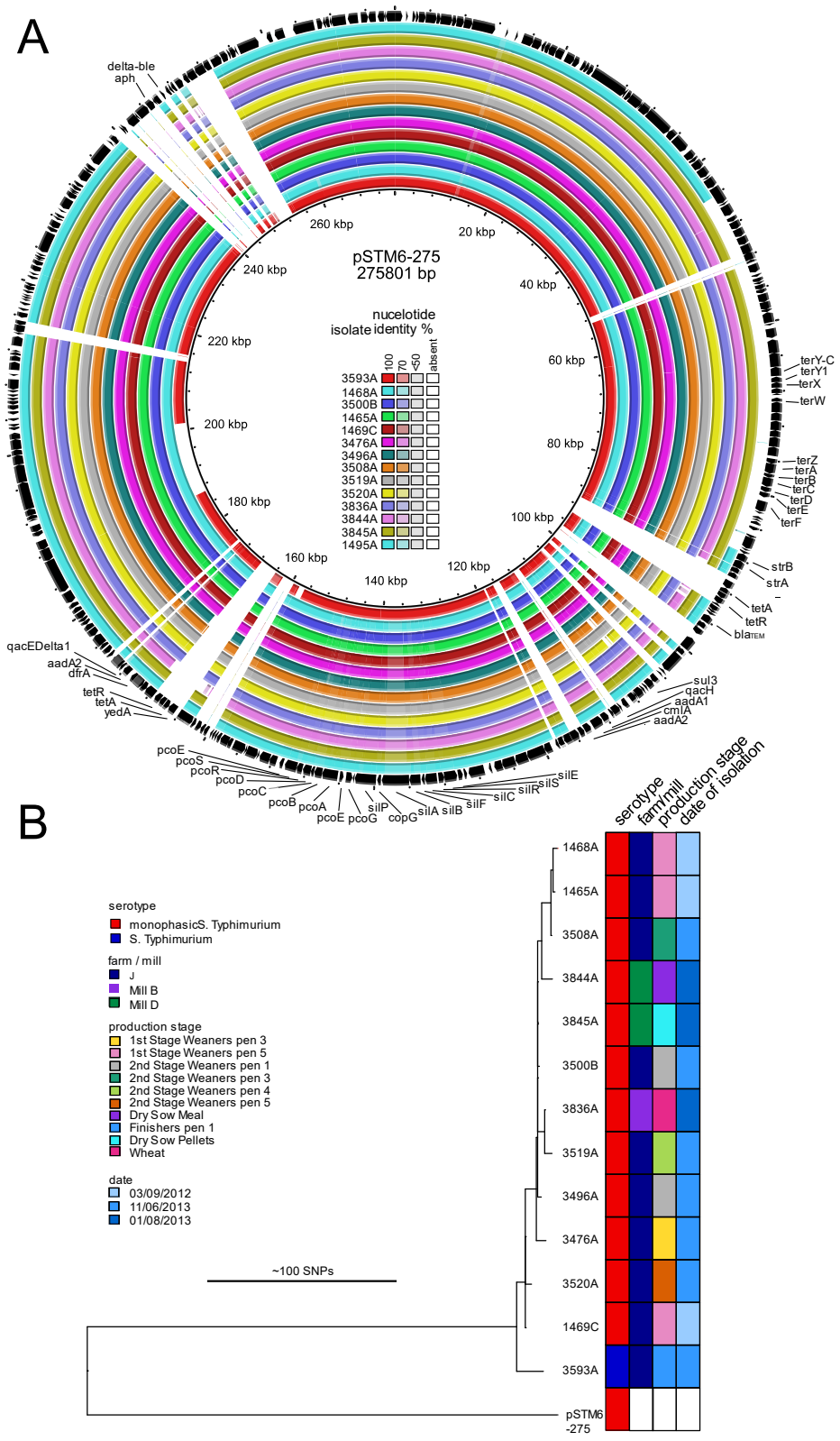

Figure 3.4 Sequence similarity and phylogenetic relatedness of IncHI2 plasmids from Irish *S*. 4,[5],12:i:- and *S*. Typhimurium isolates with the pSTM6-275 plasmid. A: A comparative genomic analysis of the assembled contigs of IncHI2 plasmids from Irish isolates with pSTM6-275 plasmid was performed using BLAST and visualised with BRIG (Alikhan et al., 2011). Each circle represents the plasmid from a single isolate. The colour gradient reflects the percentage of nucleotide identity. The coding sequences of

the reference plasmid pSTM6-275 are represented by arrows (outer circle). Genes conferring resistance to antibiotics and heavy metals are annotated. B: Maximum-likelihood phylogenetic tree of IncHI2 plasmids in Irish *S*. 4,[5],12:i:- and *S*. Typhimurium based on 162 core-plasmid SNPs identified using pSTM6-275 as a reference. Due to the high proportion of missing data, the plasmid from isolate 1495C was excluded from the phylogenetic analysis.

3.2.5 SGI-4, fljAB locus deletion and sopE gene flux in *S*. 4,[5],12:i:-

Most of *S*. 4,[5],12:i:- harboured the SGI-4 inserted in the *yidC* gene encoding for the resistance to multiple heavy metals (Petrovska et al., 2016, Branchu et al., 2019). The stable maintenance of the SGI-4, with only infrequent and sporadic loss, suggested an associated selective advantage (Petrovska et al., 2016). The totality of *S*. 4,[5],12:i:- from Irish pig farms carried the SGI-4, further supporting the role of the genetic element in the success of the clones (Figure 3.5).

Another main characteristic of the pandemic *S*. 4,[5],12:i:- ST34 is the lack of the phase 2 flagellar antigen encoded by the *fljB*. Multiple deletions of the *fljB* gene and neighbouring coding sequences generated further genotypic variation (Petrovska et al., 2016). The *fljB* gene and flanking regions were inspected in the Irish isolates by assessing gene presence/absence with SRST2 (Inouye et al., 2014) using the coding sequences of *S*. Typhimurium LT2 as a reference (accession number AE006468.2). All 104 *S*. 4,[5],12:i:- isolates lacked *fljB*. However, two main patterns of deletion reflecting the tree topology were present (Figure 3.5). Approximately half of the isolates only lacked the *fljB and hin* DNA invertase genes, whereas the remaining *S*. 4,[5],12:i:- lacked thirteen genes spanning from STM2760 (DNA-binding protein) to *hin* as in the UK representative *S*. 4,[5],12:i- S04698-09. However, two unrelated isolates exhibited a deletion extending from ORF STM2752, encoding for a PTS sorbitol transporter, to *hin* encompassing several genes involved in cellular metabolism. A single isolate lacked 25 genes spanning from STM2760 to the *tctE* gene, including the virulence genes *iroBCDEN*, *pipB*, *virK* and *mig-14*. Distinct patterns of gene deletion in the sequences flanking *fljAB* operon represent a source of variation which may affect the coding capacity of *S*. 4,[5],12:i:- to produce virulence factors and/or proteins involved in cellular metabolism.

The acquisition of the *sopE* virulence gene mediated by the mTmV bacteriophage further contributed to the microevolution of *S*. 4,[5],12:i:- from UK and Italy

(Petrovska et al., 2016). The presence of the *sopE* gene and mTmV prophage sequence in the raw sequence reads on Irish farms was screened using SRST2 software (Inouye et al., 2014). Fifty-five *S*. 4,[5],12:i:- isolates were positive for *sopE*, all harbouring mTmV sequences indicating that the prophage was responsible for its acquisition as for UK *S*. 4,[5],12:i:-. The distribution of the positive isolates within the phylogenetic context is consistent with a single event of *sopE* acquisition followed by clonal expansion (Figure 3.5).

Finally, none of the Irish *S*. 4,[5],12:i:- harboured the virulence plasmid pSLT, common in *S*. Typhimurium, in agreement with previous reports (Petrovska et al., 2016). The lack of the virulence plasmid was suggested by the lack of the IncF replicon associated with it (IncF) (Figure 3.3) and of the virulence genes encoded by the pSLT (data not shown).

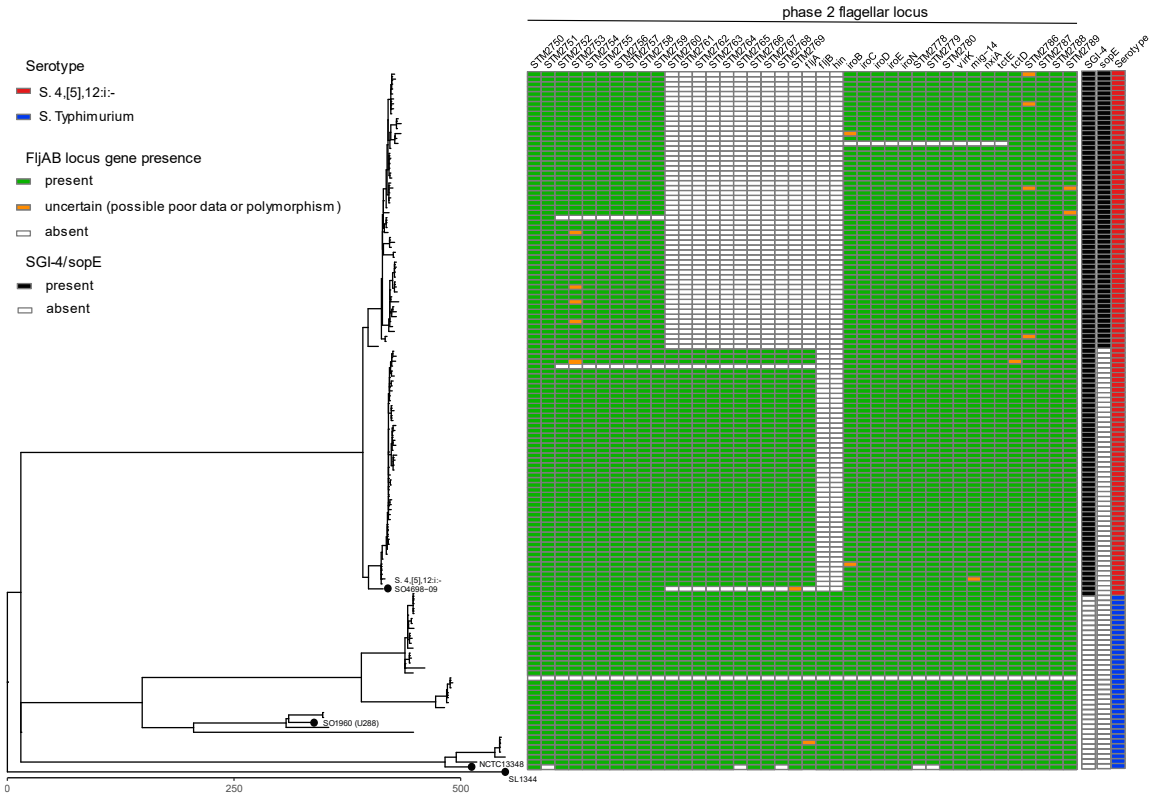

Figure 3.5 Variation in presence/absence of the *fljAB* locus and surrounding genes, SGI-4 and *sopE*. The presence of the genes encoding the second phase flagellum FljB and the flanking genes (green/orange boxes), the presence of SGI-4 and of *sopE* (black boxes) are shown within the phylogenetic context of *S*. Typhimurium and *S*. 4,[5],12:i:-. Minimum gene coverage: 90%.

3.2.6 Phylogenetic signature of biofilm formation of *S*. 4,[5],12:i:- and *S*. Typhimurium isolates

The formation of biofilms promotes the survival and persistence of some bacteria in the environment enhancing their resilience to physical, chemical and biological threats (White et al., 2006). In the present study, I addressed the hypothesis that biofilm forming capacity is variable within *S*. Typhimurium and *S*. 4,[5],12:i:- isolated from Irish pig farms and that successful clones have distinct capacities to form biofilms.

To test the ability of the isolates to produce cellulose and curli fimbriae, two major components of the extra-cellular matrix of *Salmonella* biofilms (Romling, 2005), the red, dry and rough (RDAR) morphotype of the colonies was assessed by plating on LB agar supplemented with Congo red dye. The majority of both *S*. 4,[5],12:i:- (98.1%) and *S*. Typhimurium (76.7%) isolates exhibited the (RDAR) phenotype suggesting the production of both cellulose and curli fimbriae (Romling, 2005) (Figure 3.6). The remaining isolates produced a pale and smooth (PAS) phenotype which denoted a lack of curli fimbriae but normal cellulose production (Baugh et al., 2012).

To further study the ability of the isolates to produce biofilms, the formation of biofilms on the air-liquid interface and plastic surfaces was assessed. *Salmonella* can produce pellicles (a film of cells grown on the air-liquid interface) when incubated in LB broth without salts without shaking (Scher et al., 2005). While most of *S*. 4,[5],12:i:- formed strong pellicles (70%), the majority of *S*. Typhimurium produced fragile pellicles (72%) (Figure 3.6C and D). Pellicles were suggested to be formed by isolates with rugose morphology on Congo red agar plates (Anriany et al., 2006, White and Surette, 2006). While RDAR *S*. 4,[5],12:i:- isolates tended to produce a strong pellicle, most of RDAR *S*. Typhimurium formed a fragile pellicle (Figure 3.6E). The PAS phenotype was associated with the production of a fragile pellicle regardless of the genotype.

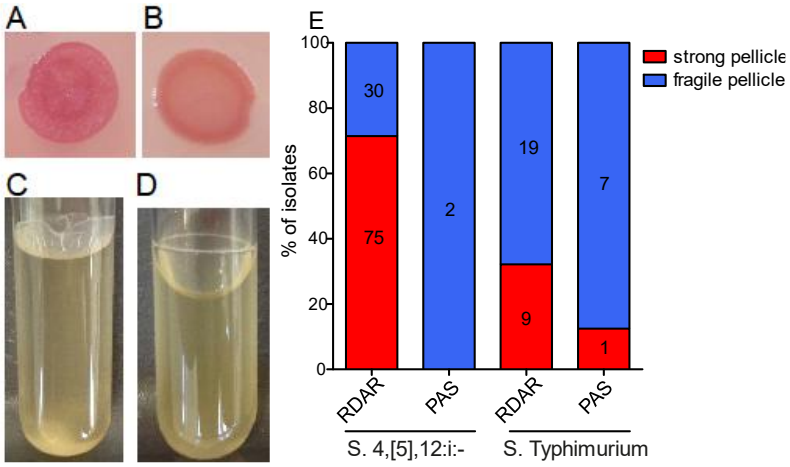

Figure 3.6 Colony morphotype on Congo red agar plates and pellicle formation in the air-water interface. A and B) Morphology of the colonies grown on LB agar plates supplemented with the Congo red dye. Red, dry and rough (RDAR) phenotype and pink and smooth (PAS). C and D) Formation of strong and rigid (C) or fragile (D) pellicle in the air-liquid interface. E) Correlation between the phenotype of the colonies displayed on Congo red agar plates and the strength of the pellicle produced in the air-liquid interface. The proportion (%) of RDAR/PAS isolates forming a strong or fragile pellicle is shown for *S*. 4,[5],12:i:- and *S*. Typhimurium isolates. The absolute number of isolates is reported within each bar.

The microtiter plate assay was then performed to study biofilm formation on a polystyrene surface. The biomass of the biofilms was measured as optical density (OD_{595nm}) following the staining of the adherent cells with the Crystal violet dye. *S*. 4,[5],12:i:- and *S*. Typhimurium formed biofilms to the same extent after 24 and 48 hours of incubation at 22°C (Figure 3.7A). To test the hypothesis that phylogenetically distinct groups varied in their ability to form biofilms, the OD_{595nm} values of the isolates were analysed in relationship with the tree topology. A phylogenetic signature with significant differences in biofilm forming ability between the groups was observed (Figure 3.7B and C). All the *S*. Typhimurium related with the DT104 representative strain were high biofilm formers (clade E). The remaining *S*. Typhimurium were generally poor biofilm producers, with the exception of two isolates in clade B. By contrast, all *S*. 4,[5],12:i:- were moderate/strong biofilm producers. Notably, a group of eight closely related isolates (sub-clade A1) produced considerably more biofilms than the rest of *S*. 4,[5],12:i:- (sub-clade A2) despite their high relatedness (Figure 3.7B and C). A single non-synonymous SNP in the *edd* gene encoding for a phosphogluconate

dehydratase differentiated the *S*. 4,[5],12:i:- with enhanced biofilm forming ability from the *S*. 4,[5],12:i:- isolates in sub-clade A2.

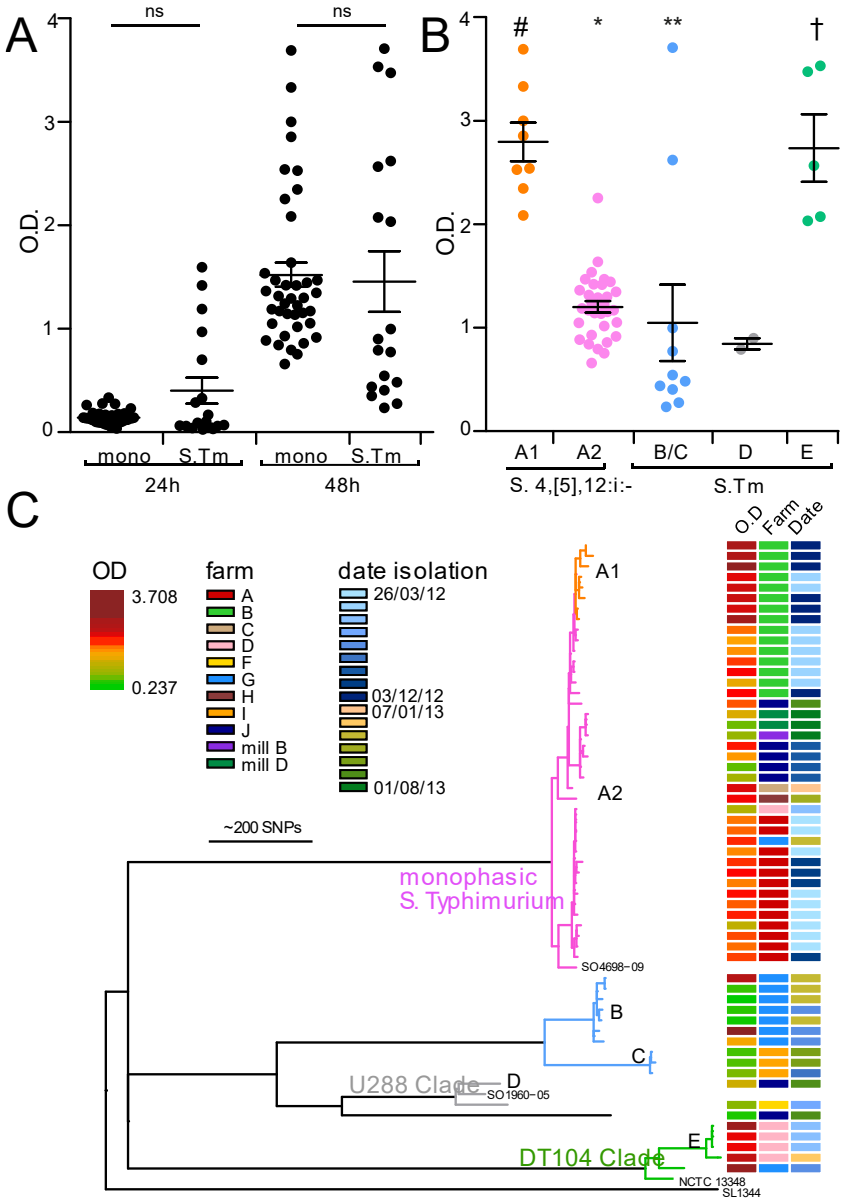

Figure 3.7 Biofilm formation of *S*. 4,[5],12:i:- and *S*. Typhimurium on polystyrene correlates with the phylogeny. A) Scatterplot of the OD_{595nm} values measured for *S*. 4,[5],12:i:- and *S*. Typhimurium after 24 and 48 hours of incubation at 22°C. Non-significant difference of biofilm formation between *S*. 4,[5],12:i:- and *S*. Typhimurium after 24h (p-value 0.7) and after 48h (p-value: 0.1) at 22°C. B) Comparison of the OD_{595nm} measured after 48 hours of incubation at 22°C of isolates from each clade of the phylogenetic tree shown in C. Pairwise Mann-Whitney test (95% confidence interval) was performed and p<0.05 were indicated;(*): $p < 0.05$ for all except clade E isolates (DT104-like) (#),: $p < 0.05$ for all clades (**),: $p < 0.05$ for all except clade A1 (†). C) Maximum-likelihood phylogenetic tree based on 2,328 informative SNPs in reference with *S*. Typhimurium SL1344. The OD_{595n} values measured after incubation at 22°C for

48 hours was reported as heatmap, together the farm and date of isolation of the isolates. The designation of the clades (A1, A2, B and C, D and E) is reported.

An association between production of biofilms on polystyrene after 48 hours of incubation at 22°C and the formation of a pellicle in the air-liquid interface was observed for both *S*. 4,[5],12:i:- and *S*. Typhimurium isolates (Figure 3.8).

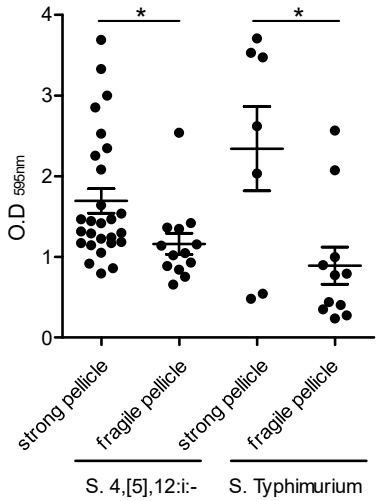

Figure 3.8 Correlation between pellicle formation and biofilm formation on plastics. Scatterplot showing the OD_{595nm} values in relationship with the strength of the pellicle produced in the air-liquid interface by *S*. 4,[5],12:i:- and *S*. Typhimurium. The Mann-Whitney test (95% C.I.) was performed, p-values: 0.0086 (*S*. 4,[5],12:i:-) and 0.0297 (*S*. Typhimurium).

In conclusion, variation in biofilm formation was observed in the production of biofilms in the air-liquid interface and on plastics.

While the majority of *S*. Typhimurium produced a fragile pellicle in the air-liquid interface, most of *S*. 4,[5],12:i:- produced a strong pellicle; the ability to produce pellicles may confer competitive advantage in the extra-host environment.

On plastics, the formation of biofilms at 22°C after 48h of incubation correlated with the phylogenetic clustering of the isolates. Importantly, isolates belonging to epidemic clades (S. 4,[5],12:i:- and S. Typhimurium DT104) formed significantly more biofilms than non-epidemic isolates suggesting that biofilm-forming ability may confer selective advantage to the isolates and contribute to their epidemiological success. Additionally, a group of closely related S. 4,[5],12:i:-

exhibited an increased biofilm-formation ability over a short period of time indicating selection towards this phenotype.

3.3 Discussion

S. 4,[5],12:i:- was the predominant *Salmonella* serovar from pig farms in the Republic of Ireland, present in eight out of the ten farms originally sampled and in four feed mills (Burns, 2015). Irish *S*. 4,[5],12:i:- were closely related to the UK epidemic S. 4,[5],12:i:- characterised by sequence type ST34 and phage DT193/DT120 currently pandemic (Petrovska et al., 2016). The *S*. Typhimurium isolates, by contrast, were more diverse and were present on only six farms.

Each farm was colonised by a specific genotype which, in most of the cases, was widespread across the farm site. Highly related (<5 SNPs) *S*. 4,[5],12:i:- and *S*. Typhimurium were isolated from faeces and a number of environmental samples, including feed, water, pen walls and floor, trough and water drinkers. In five farms *S*. 4,[5],12:i:- and *S*. Typhimurium isolates from weaners and finishers were closely related to isolates from breeding pigs suggesting vertical transmission of the pathogen. However, a recent study showed that the transmission of *Salmonella* from sows/gilts to the progeny was uncommon and that environmental contamination posed a greater risk for the acquisition of the bacteria (Lynch et al., 2018). However, examples of highly related isolates from distinct farms and feed mills were observed, suggesting either contamination from a common source, such as feed, or transmission. Transmission cannot be excluded as movement of the animals between farms was not known. Contaminated feed is a well-established source of *Salmonella* on farms (Burns, 2015). The close relatedness of a group *S*. 4,[5],12:i:- from farm A with an isolate from mill C, which supplied the farm, indicated a recent transmission from contaminated feed.

Most of the genotypes persisted for up to 9 months on farms, across multiple cycles of pig production. Clones on each farm underwent sequence variation, generally differing by 0-12 SNPs. Based on the mutation rate of epidemic *S*. Typhimurium of 1-2 SNPs per genome per year (Mather et al., 2013, Okoro et al., 2012), the MRCA of each farm-specific clone existed in the past 2-7 years. Analogous findings were previously described using sub-typing methods (Gebreyes et al., 2006, Weaver et al., 2017, Lim et al., 2011). The persistence of the clone on the farm site may be promoted by ineffective biosecurity measures (such as cleaning and disinfection) or by re-introduction from a common source such as feed or wild animals (Andres-Barranco et al., 2014, De Lucia et al., 2018).

Additionally, the formation of biofilms has been associated with enhanced environmental persistence of *Salmonella* (Vestby et al., 2009).

Most of the Irish *S*. 4,[5],12:i:- and *S*. Typhimurium produced both cellulose and curli fimbriae, two of the principal components of the extracellular matrix of *Salmonella* biofilms, but differences in the production of biofilms were observed between distinct genotypes. Contrary to *S*. Typhimurium, most of *S*. 4,[5],12:i:- formed strong and rigid biofilms in the air-liquid interface (pellicles). The ability to produce pellicles may confer an ecological advantage over other species which do not occupy this niche, but also confers protection against chemical agents such as chlorine (Scher et al. 2007). By contrast, the formation of biofilms on polystyrene surfaces was analogous for *S*. 4,[5],12:i:- and *S*. Typhimurium. However, biofilm forming ability greatly varied between phylogenetically distinct clusters especially for *S*. Typhimurium. Isolates related to the epidemic clone *S*. Typhimurium DT104 (clade E) produced significantly greater biofilm mass than U288-related isolates (clade D) and non-epidemic *S*. Typhimurium from clades B and C. *S*. 4,[5],12:i:- were all moderate to strong biofilm formers, although a cluster of isolates (sub-clade A1) exhibited greater biomass production than other *S*. 4,[5],12:i:-. Isolates in sub-clade A1 differed from the remaining *S*. 4,[5],12:i:- for a single non-synonymous SNP in the core genome in the *edd* gene encoding for an enzyme in the Entner-Doudoroff (ED) metabolic pathway (Peyru and Fraenkel, 1968) which has been previously implicated in biofilm formation in *Campylobacter jejuni* (Vegge et al., 2016). The greater formation of biofilms by isolates associated with the previous dominant epidemic *S*. Typhimurium DT104 and with the current pandemic *S*. 4,[5],12:i:- suggested that this phenotype may have contributed to the success of the clones. However, the ability to produce biofilms was not essential for the environmental persistence of *Salmonella*, since poor biofilm-forming isolates were isolated on two different occasions on farm G.

AMR was frequent in Irish *S*. 4,[5],12:i:- and *S*. Typhimurium, in accordance with previous reports (ECDC, 2018). The resistance profiles of isolates related to *S*. 4,[5],12:i:- DT193/DT120, S. Typhimurium DT104 and U288 agreed with the resistance patterns reported for these clones (Lucarelli et al., 2012, Hooton et al., 2014, Petrovska et al., 2016), with occasional deletion of one or more genes as previously described (Mulvey et al., 2006). However, two groups of isolates exhibited further AMR following the acquisition of plasmids. The acquisition on

the farm site of plasmid sequences associated with IncI/IncQ replicons conferred tetra-resistance to four *S*. Typhimurium (clade B). Thirteen *S*. 4,[5],12:i:- acquired a large IncHI2 plasmid related to pSTM6-275, previously reported in Australia, encoding multiple AMR and heavy metal resistance genes (Billman-Jacobe et al., 2018). The plasmids identified here differed from the IncHI2 plasmids reported in *S*. 4,[5],12:i:- in Vietnam, Portugal, China and in a number of countries from patients with a travel-associated history to Asia (Mather et al., 2018, Campos et al., 2016, Biswas et al., 2019). Eleven of the *S*. 4,[5],12:i:- positive for the plasmid were isolated on a single farm on two separate occasions, whereas three related isolates were isolated on two distinct feed mills. Feed may represent a possible source of the plasmid. Noteworthy was the presence of a related IncHI2 plasmid in a single *S*. Typhimurium isolate on the same farm, indicating the potential dissemination of the IncHI2 plasmid on farms. IncHI2 plasmids may contribute to the environmental persistence of the isolates.

Besides AMR, further genotypic variation was described in the phase 2 flagellar locus *fljAB*, in agreement with previous reports (Petrovska et al., 2016). Despite the clonality of the isolates, multiple patterns of gene deletions flanking the locus were observed. The impact of the deletion of the ORFs, which encoded metabolic functions and virulence genes, on the fitness of the cells is unknown. Additionally, over half of *S*. 4,[5],12:i:- encoded the virulence gene *sopE*, whose acquisition was mediated by mTmV (Petrovska et al., 2016) and followed by clonal expansion, suggesting the advantage conferred by it.

3.4 Conclusion

In conclusion, specific and persistent clonal groups were described on each pig farm. Nevertheless, highly clonal genotypes potentially varied in virulence, flagella antigen expression, and in antimicrobial and heavy metal resistance. Additionally, the ability to produce biofilms correlated with genotype, suggesting that the effectiveness of intervention strategies aimed at *Salmonella* prevalence on pig farms may vary depending on the prevalent genotype.

4 The molecular epidemiology of the *sopE* virulence gene in pandemic *S*. 4,[5],12:i:- ST34

4.1 Introduction

S. Typhimurium is a diverse serovar which comprises several distinct pathovariants characterised by different host range and pathogenic potential (Branchu et al., 2018). Only few MDR *S*. Typhimurium variants have been responsible for epidemics in humans over the past 50 years in Europe, namely DT9, DT204 complex, DT104 complex and, currently, the *S*. 4,[5],12:i:- DT193/DT120 ST34 complex (Rabsch et al., 2011, Petrovska et al., 2016). The genetic factors that lead to emergence and clonal expansion are not known but a reassortment of the repertoire of virulence genes mediated by mobile genetic elements, such as plasmids and temperate bacteriophages, has been proposed to contribute to the success of epidemic clones (Figueroa-Bossi et al., 2001).

Several bacteriophage-encoded virulence factors have been described in *Salmonella*. One of these, SopE, is an effector secreted by the SPI-1-encoded T3SS that has guanine nucleotide exchange factor activity for the RhoGTPases CDC42 and Rac1 (Friebel et al., 2001). Upon translocation into enterocytes, SopE promotes microbial internalisation into intestinal cells through rearrangement of the cytoskeleton through its guanine nucleotide exchange factor activity (Hardt et al., 1998a). This activity is detected by the inflammasome and contributes to the proinflammatory response that remodels the intestinal environment promoting the expansion of *Salmonella* in the inflamed gut (Winter et al., 2010a, Muller et al., 2009). Additionally, SopE may be involved in the initial formation of *Salmonella*-containing vacuole in macrophages (SCV) (Vonaesch et al., 2014).

sopE has only been reported in a limited number of *S*. Typhimurium lineages. Most *S*. Typhimurium isolates from a collection of representative isolates from Germany (SARA collection) were negative for *sopE*, with the exception of isolates belonging to phage types DT68 and DT175, and the DT204 complex (DT204/DT49/DT44/DT204c) (Mirold et al., 1999). Similarly, in a collection of human and animal *S*. Typhimurium isolates from the UK, *sopE* was identified exclusively in *S*. Typhimurium phage types DT49, DT204 and 204c, but also DT29 and DT44 (Hopkins and Threlfall, 2004). The observation that *sopE* was present in phage types responsible for major epidemics, namely DT49, 204 and

204c which accounted for a large proportion of infections in cattle and humans in the United Kingdom and former East Germany across the 1970s and 1980s, led to the hypothesis that SopE contributed to the emergence of epidemic clones (Mirold et al., 1999, Hopkins and Threlfall, 2004).

More recently, *sopE* was reported in the current MDR pandemic *S*. 4,[5],12:i:- ST34, predominantly characterised by phage types DT193/120 (Petrovska et al., 2016). *S*. 4,[5],12:i:- ST34 emerged in the UK in the mid 2000s in pigs, rapidly replacing *S*. Typhimurium DT104 and becoming the predominant clone (Petrovska et al., 2016). Lysogenic conversion by a novel bacteriophage, termed mTmV, was responsible for the acquisition of *sopE* in *S*. 4,[5],12:i:- (Petrovska et al., 2016). Multiple events of *sopE* acquisition by *S*. 4,[5],12:i:- together with clonal expansion of the isolates, which resulted in increased frequency of *sopE* positive isolates over time, suggested an advantage conferred by its acquisition (Petrovska et al., 2016). The presence of *sopE* in a large proportion of human and animal *S*. 4,[5],12:i:- isolated in Europe after 2010 (48%; 350/727) further indicated the contribution of *sopE* to the success of the *S*. 4,[5],12:i:- ST34 clones (Palma et al., 2018).

In order to test the hypothesis that the acquisition of *sopE* was positively selected in *S*. 4,[5],12:i:-, the dynamics of *sopE* acquisition were investigated in pig isolates and placed in the temporal context with the emergence of *S*. 4,[5],12:i:-. Additionally, the epidemiological impact of *sopE* acquisition on human infections was investigated in a collection of *S*. Typhimurium and *S*. 4,[5],12:i:- from clinical infections in the UK.

4.2 Results

4.2.1 The *sopE* virulence gene is sparsely distributed in *S*. Typhimurium, but common in the current pandemic *S*. 4,[5],12:i:- ST34 clade

To test the hypothesis that *sopE* associates with epidemic *S*. Typhimurium clones, the distribution of *sopE* in a collection of whole genome sequences of 142 strains representative of the genotypic diversity within *S*. Typhimurium distinguished by phage typing was determined. The collection comprised 128 isolates previously published (Branchu et al., 2018), 12 food and environmental isolates, and isolates 14028 and SL1344 (Appendix 2). The presence of *sopE*, detected using ARIBA-2.8.1 (Hunt et al., 2017a), was investigated within the phylogenetic context of the isolates. A maximum-likelihood phylogenetic tree based on 8,663 SNPs in the core genome identified was generated with RaxML v8.2.10 using the GTR+Γ substitution model, *S*. Typhimurium SL1344 was used as a reference (Figure 4.1). The log likelihood score obtained was -69641.7.

The analysis suggested association of the *sopE* virulence gene with limited number of *S*. Typhimurium lineages associated with major epidemics which occurred across the 1970-80s, namely DT204 clonal complex, or are currently ongoing, *S*. 4,[5],12:i:- ST34 complex (chi-squared p-value <0.001). In particular, *sopE* was detected in *S*. Typhimurium phage types DT204, DT44, DT49, but also U310, which clustered in a single complex (DT204 complex) consistent with previous reports (Hopkins and Threlfall, 2004, Mirold et al., 1999), as well as in the current pandemic *S*. 4,[5],12:i:- ST34 clade (Mirold et al., 1999, Hopkins and Threlfall, 2004, Petrovska et al., 2016). Despite the diverse genotypes included in the analysis, none of the *S*. Typhimurium outside the two clades above contained the *sopE* gene. As previously published, the virulence gene was absent from the previous dominant epidemic *S*. Typhimurium DT104, and from *S*. Typhimurium U288, which is commonly isolated from pigs as *S*. 4,[5],12:i:- (Hooton et al., 2011).

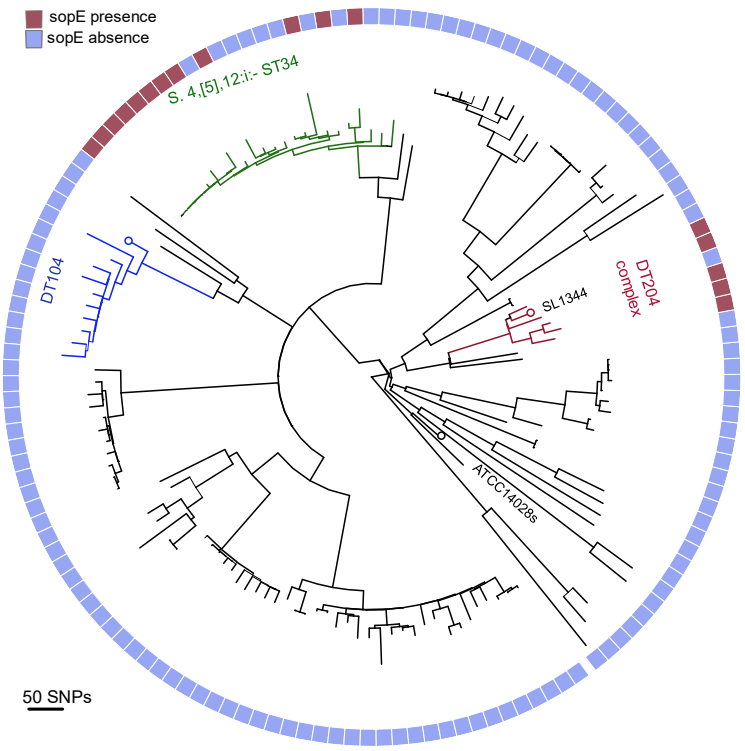

Figure 4.1 Phylogenetic distribution of *sopE* in representative *S*. Typhimurium and *S*. 4,[5],12;i;- isolates. Maximum-likelihood phylogenetic tree based on 8,663 core-genome SNPs with reference to *S*. Typhimurium SL1344. The presence/absence of *sopE* is indicated by filled boxes as indicated in the embedded key. *S*. Typhimurium DT204 and DT104 and *S*. 4,[5],12:i- ST34 clonal complexes are highlighted.

4.2.2 Acquisition of *sopE* occurred early during the emergence of *S*. 4,[5],12:i:- in pig farms and on multiple occasions

The distribution of the *sopE* gene in *S*. 4,[5],12:i:- ST34 could be the result either of a single event of acquisition followed by subsequent multiple deletions or of multiple acquisitions. While the former hypothesis would imply a certain selection against *sopE*, the second would suggest that the acquisition of the gene was favoured. To test these hypotheses, ancestral state reconstruction was performed on the phylogenetic structure of 62 representative *S*. 4,[5],12:i:- isolates isolated in pigs in the UK between 2006 and 2015 (Appendix 3). The distribution of the number of isolates per year is shown in Figure 4.2. The analysis was limited to *S*. 4,[5],12:i:- in pigs in order to focus the analysis on a defined host population and therefore limit confounding factors that may arise from consideration of multiple parallel epidemics. The WGS of the isolates partly came from a previous study (Petrovska et al., 2016), and partly were generated at the

Earlham Institute from isolates obtained by the Animal Plant Health Agency (APHA).

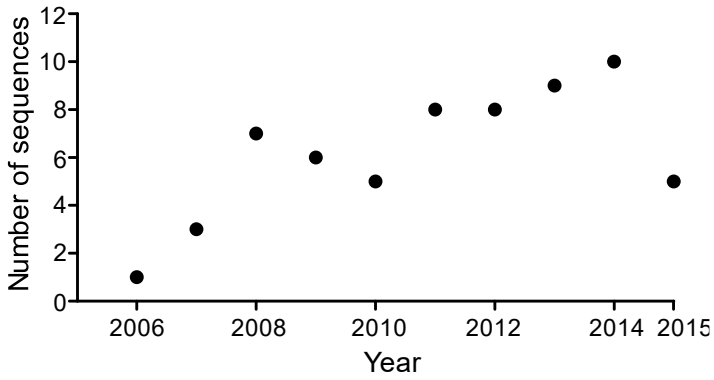

Figure 4.2 Plot showing the number of pig isolates per year included in the analysis.

In order to interpret the evolutionary events, a maximum-likelihood phylogenetic tree was generated using 1,095 informative SNPs with reference to *S*. Typhimurium strain SL1344 and rooted using an outgroup (*S*. Typhimurium LT2) (Figure 4.3A). The phylogenetic tree was generated with RAxML v8.2.12 with the GTR+Γ substitution method and the log likelihood score obtained was -9806.1.

sopE was detected in the whole genome sequences of 39 isolates with SRST2 (Inouye et al., 2014). Discrete clades of *sopE*-positive isolates were present across the phylogeny similar to that previously described (Petrovska et al., 2016). Ancestral state reconstruction was performed with PastML marginal posterior probabilities approximation (MPPA) method with the F81 substitution model (Ishikawa et al., 2019). The analysis suggested that *sopE* was acquired multiple times, at least seven (Figure 4.3B).

The most recent common ancestor (MRCA) of the isolates and the most deeply rooted nodes lacked *sopE*, which was instead present in multiples clusters of the descendants branches consistent with multiple gene acquisition events. The finding was in agreement with the literature (Petrovska et al., 2016).

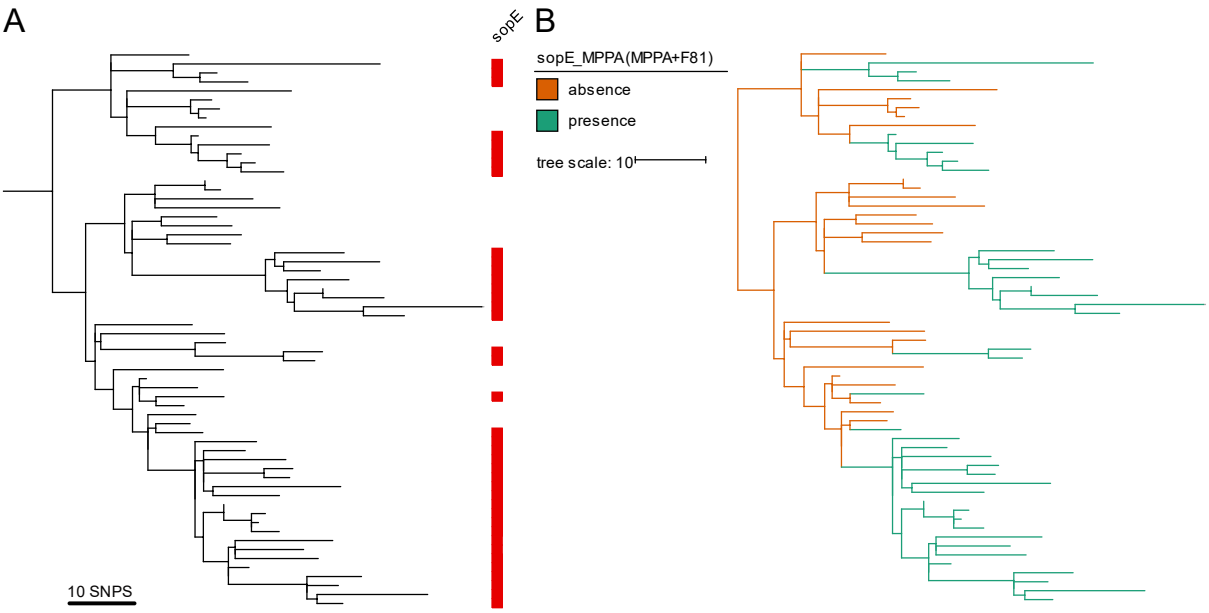

Figure 4.3 *sopE* was acquired multiple times by *S*. 4,[5],12;i;- from UK pig farms. A) Maximum likelihood phylogenetic tree of porcine S. 4,[5],12:i:- based on 1,565 SNPs in reference with *S*. Typhimurium SL1344 (not shown) and rooted using S. Typhimurium strain LT2 as outgroup (not shown). The red blocks represent the presence of *sopE*. B) Ancestral character reconstruction for the *sopE* gene generated using the marginal posterior probabilities approximation (MPPA) method with the F81 substitution model.

The acquisition of *sopE* was then investigated from a temporal perspective in relationship with the dynamics of the epidemic to test the hypothesis that it contributed to the emergence of *S*. 4,[5],12:i:- ST34. To this end, a time-dependent phylogenetic tree was constructed using the BactDating software (Didelot et al., 2018). The presence of a temporal signal within the core genome sequence of the isolates was determined by performing a root-to-tip linear regression based on the topology of the phylogenetic tree in Figure 4.4A and the isolation date of the strains. A good temporal signal was observed (Figure 4.4B). Generally, the earlier the isolates were isolated, the closer to the root they tended to be, and vice versa. There was a statistically significant correlation (p-value <1.00e-4) between the number of SNPs accumulated and the year of isolation (R^2 = 0.34) (Figure 4.4). The molecular clock rate was estimated to be 4.9 substitutions per site per year or 2.46 mutations per genome per year, slightly higher than the 1-2 SNPs per genome per year substitution rate previously reported for *Salmonella* epidemics (Okoro et al., 2012, Mather et al., 2013).

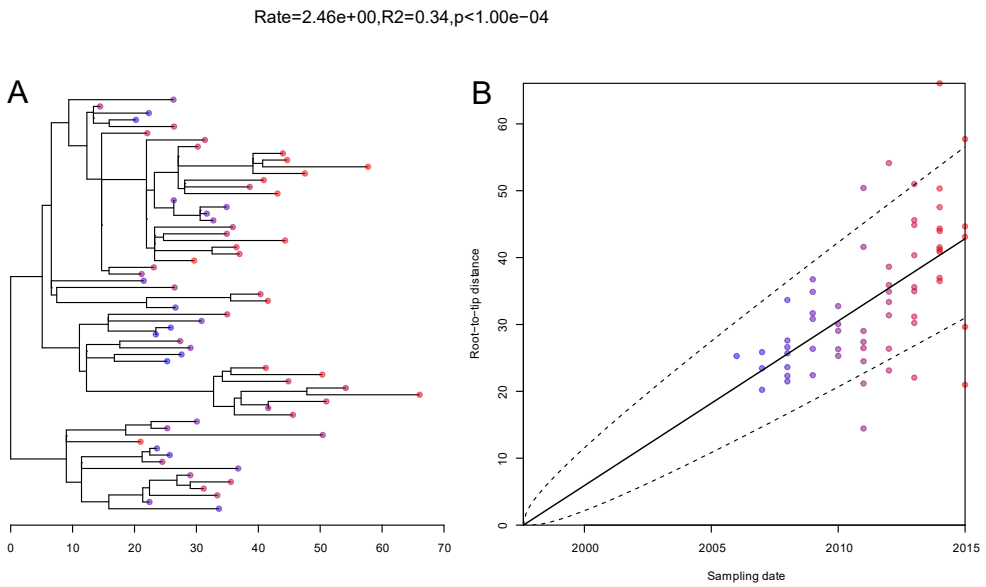

Figure 4.4 Analysis of the temporal signal in the ML phylogenetic tree of *S.* 4,[5],12;i;- from UK pig farms by root-to-tip regression analysis. A: Rooted maximum-likelihood phylogenetic tree used for the root-to-tip analysis (same as Figure 4.3A). The bar indicates the number of SNPs from the root. The tips colour corresponds to the year of isolation: the older isolates are blue, whereas the most recent are red. B: Linear regression of root-to-tip (y-axis) versus sampling dates (x-axis). The dashed lines represent the 95% confidence interval.

The most recent common ancestor (MRCA) dated back in 1994, although there was a large margin of about 15 years of uncertainty (Figure 4.5). The date is consistent with first detection of *S*. 4,[5],12:i:- in pigs in the UK (Mueller-Doblies et al., 2018).

The earliest events of *sopE* acquisition occurred in a temporal window concomitant with the emergence of *S*. 4,[5],12:i:- ST34. Two major subclades (A and B), which correspond to subclades (A and B, respectively) of the phylogeny previously published (Petrovska et al., 2016), were identified which emerged in the late 1990s. *sopE* was predicted to have been acquired at least seven times between 2000 and 2012, during the clonal expansion of the isolates (Figure 4.5). The earliest events of acquisition of the gene were estimated to have occurred between 2000-2005 and 2004-2007, in subclade B, and 2001-2007, in subclade A. These temporal windows were consistent with the detection of the first *sopE*-positive pig isolate in the UK in 2007 (Petrovska et al., 2016). Multiple events of *sopE* acquisition, concomitant with the temporal emergence of *S*. 4,[5],12:i:-

ST34, followed by the clonal expansion indicated that *sop*E was positively selected.

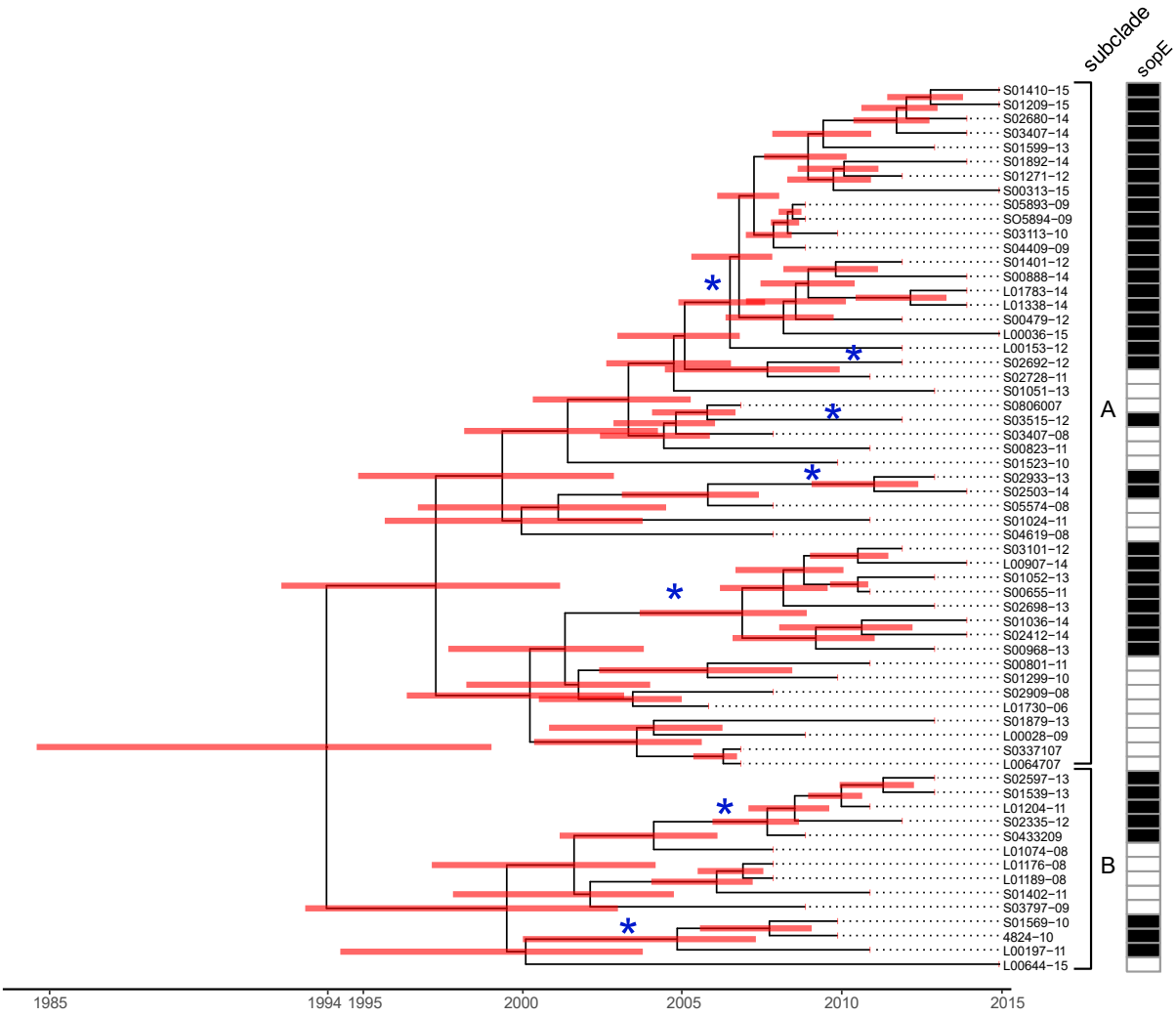

Figure 4.5 Time-dependent phylogenetic tree of *S*. 4,[5],12;i;- pig isolates from the UK showing the events of *sopE* acquisition. The red bar represents the confidence interval (C.I.=95%) of the inferred ancestral dates. The presence of *sopE* presence was shown. The events of sopE acquisitions are marked with blue asterisks.

4.2.3 High frequency of *sopE* in clinical isolates of *S*. 4,[5],12:i:- sequence type ST34

To investigate whether the acquisition of *sopE* played any role in the transmission of *S*. 4,[5],12:i:- ST34 from animals to humans, the frequency and distribution of *sopE* was first investigated in a collection of whole genome sequences of 1,697 clinical *S*. 4,[5],12:i:- and *S*. Typhimurium from UK. The isolates were isolated and sequenced by Public Health England (PHE) between November 2012 and

January 2016 (Table 4.1 and Appendix 4). Isolates between April 2014 and March 2015 comprised all the *S*. Typhimurium received by PHE over the twelve-month period.

Table 4.1. Number and relative percentage of sequences of *S*. 4,[5],12:i:- and S. Typhimurium from each year.

Year	N. of isolates (%)
2012	20 (1.2%)
2013	4 (0.2%)
2014	1,150 (67.8%)
2015	516 (30.4%)
2016	7 (0.4%)

The distribution of the *sopE* gene was determined within the phylogenetic context of 1,697 clinical *S*. 4,[5],12:i:- and *S*. Typhimurium from UK. A maximum-likelihood phylogenetic tree generated using 16,408 informative SNPs with reference to *S*. Typhimurium SL1344 was used for the analysis (Figure 4.6).
The phylogenetic tree was generated with RAxML v8.0.22 using the GTR+Γ substitution model, the log likelihood score obtained was -167205.3.
Isolates belonging to pandemic *S*. 4,[5],12:i:- ST34 clones were identified based on the presence of the *Salmonella* genomic island-4 (SGI-4) and the sequence type (ST34) (Branchu et al., 2019). The presence of *sopE* and of the SGI-4 were assessed with SRST2 (Inouye et al., 2014) and ARIBA (Hunt et al., 2017b) using the reference sequence of *S*. 4,[5],12:i:- S04698-09 (Petrovska et al., 2016), while the sequence type was previously determined by PHE.

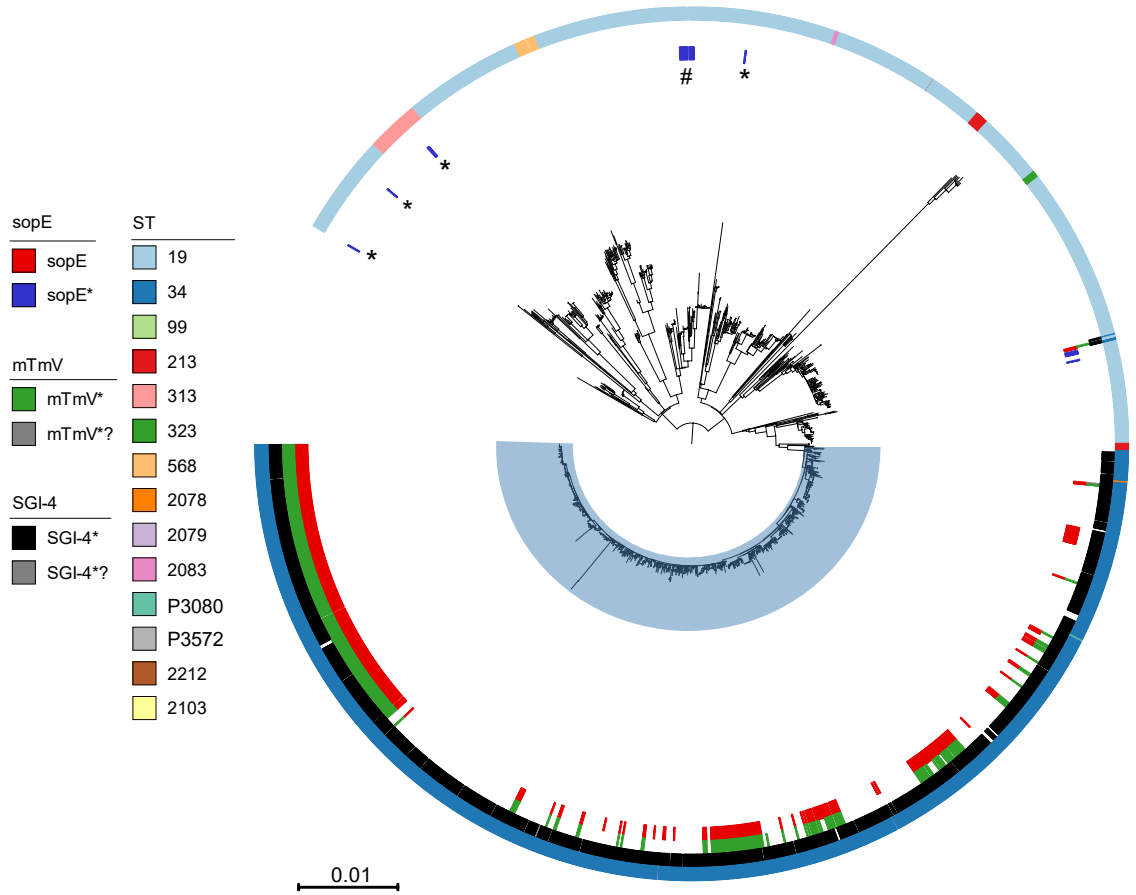

Figure 4.6 Maximum-likelihood phylogenetic tree showing the distribution of *sopE*, mTmV prophage, SGI-4 and sequence type (ST) of clinical *S.* Typhimurium and *S.* 4,[5],12:i:- isolates. The phylogenetic tree was based on 16,408 SNPs in the core genome identified with reference to S. Typhimurium strain SL1344. The presence of *sopE*, mTmV prophage, SGI-4 and the sequence types of the isolates are shown (from innermost to outermost circle). Polymorphic sequences of *sopE*, mTmV and SGI-4 are marked with an asterisk (*), whereas uncertainty due to polymorphisms and poor read depth is marked with a question mark (?). The epidemic *S.* 4,[5],12:i:- clade is highlighted in light-blue. The S. Typhimurium DT204 clonal complex is marked with a hashtag (#); S. Typhimurium genotypes which have not previously been associated with *sopE* are marked with an asterisk (*).

sopE was strongly associated with the *S.* 4,[5],12:i:- ST34 clade and predominantly associated with the mTmV prophage sequence. Distinct allelic variants distinguished *sopE* in *S.* 4,[5],12:i:- ST34 and in *S.* Typhimurium suggested a distinct source of acquisition.

S. 4,[5],12:i:- ST34 isolates formed the largest clade, containing 923 isolates, accounting for the 54.4% of all the sequences. Almost the totality of *S.* 4,[5],12:i:- were ST34, with just four monophasic isolates of other sequence types (ST2078,

ST3080, P3572, ST2079). Similarly, most of the *S*. 4,[5],12:i:- isolates carried SGI-4, although some had lost it (n=42). The majority of *S*. 4,[5],12:i:- belonged to phage types DT193/DT120 (65.7%), while the remaining isolates were characterised by a variety of less common phage types, such as U311 (7.3%), U323 (4.3%), U329 (2.0%) and DT195 (1.4%), in accordance with previous publications (Petrovska et al., 2016).

The majority of *sopE*-positive isolates belonged to the pandemic *S*. 4,[5],12:i:- ST34 clade (376 isolates out of 397) accounting for the 40.7% of *S*. 4,[5],12:i:- isolates. Only twenty-one *S*. Typhimurium sequences were positive for the virulence gene, accounting for 2.7% of *S*. Typhimurium. In *S*. 4,[5],12:i:- the presence of the *sopE* virulence gene did not correlate exclusively with a single phage type, but mirrored the prevalence of the phage types in the clade. Most of the *sopE*-positive *S*. 4,[5],12:i:- isolates were DT193 (60.9%) and DT120 (10.9%), but also U323 (6.4%) and U311 (1.1%); the remaining isolates belonged to eight minor phage types.

The distribution of *sopE* in *S*. Typhimurium was consistent with previous reports (Mirold et al., 1999). For example, ten related *S*. Typhimurium ST19 isolates with phage types DT49, DT44, PT U281 and none/RDNC and an additional, unrelated isolate with phage type DT44, had the *sopE* gene (# in Figure 4.6). However, *sopE* was also detected in *S*. Typhimurium phage types which had not previously been associated with the virulence gene. For example, three *S*. Typhimurium ST19 isolates with phage type DT120, two *S*. Typhimurium ST19 phage type DT193 related with each other and three unrelated *S*. Typhimurium ST19 phage types DT203, DT40 and phage type unknown (* in Figure 4.6) all carried *sopE*. Additionally, two related *S*. Typhimurium ST313 phage type DT193 and two *S*. Typhimurium isolates of ST34 (but not part of the main ST34 clade) were also positive *sopE*.

Remarkably, all *sopE*-positive *S*. 4,[5],12:i:- carried an identical *sopE* allele, whereas five allelic forms were identified in *S*. Typhimurium which differed by 1-37 SNPs from that in *S*. 4,[5],12:i:- (Table 4.2). The two *S*. Typhimurium strains H144820727 and H143300402 of ST34, but outside the main ST34 clade, shared the same *sopE* allele of *S*. 4,[5],12:i:-, suggesting that the polymorphism associated with the sequence type and the presence of *sopE* may be related to a common recombination event. Also, the presence of the same polymorphisms

(synonymous mutation nucleotide 27 A/G) in the gene sequence of the ten isolates with phage types DT49, DT44, PT U281 and none/RDNC with a single, unrelated isolate phage type DT44 (strain H150820681) suggested either transmission or acquisition of *sopE* from a common source. Similarly, the three DT 120 isolates carried the same *sopE* allele as the ST313 isolates (SNP in nucleotide 27 as above and non-synonymous SNP nucleotide712 A/T, Asn238Tyr).

Table 4.2 *S*. Typhimurium Isolates positive for *sopE*. The sequence types, phage type and number of isolates is reported. In addition, the presence of SNPs in the *sopE* sequence in comparison with the reference from *S*. 4,[5],12:i:- S04698-09 is reported. The isolates in bold were used for the *sopE* phylogenetic analysis in Section 4.2.4.

Strain	**Sequence type**	**Phage type**	**Number of isolates**	**SNPs (nucleotide position)**
H142540304 H142920573 H143140489 H143400441 H143420589 H143520444 **H143740716** H144660699 H153020667 H154780417	ST19	DT49, DT44, PT U281 and none/RDNC	10	1 SNP (nt 27)
H143660450 H143880705 H141840480	ST19	DT120	3	2 SNPs (nt 27 and 712)
H150820681	ST19	DT44	1	1 SNP (nt 27)
H142520441	ST19	DT203	1	18 SNPs
H142780372	ST19	DT40	1	37 SNPs
H142940536	ST19	None	1	29 SNPs
H150620545 H150620544	ST313	DT193	2	2 SNPs (nt 27 and 712)
H144820727 H143300402	ST34	PT21 and U323	2	No SNPs

In *S*. 4,[5],12:i:- S04698-09 the acquisition of the virulence gene was mediated by the mTmV prophage (Petrovska et al., 2016). To test whether *sopE* was associated with mTmV in clinical *S*. 4,[5],12:i:-isolates, the presence of the prophage sequence was investigated using SRST2 (Inouye et al., 2014). The phage genome was detected in 342 isolates, exclusively in epidemic *S*. 4,[5],12:i:- isolates, with the exception of non-epidemic *S*. Typhimurium ST34 strains H144820727 and H143300402, indicating a remarkable association of mTmV with *S*. 4,[5],12:i:- (Figure 4.6). The detection of the mTmV prophage correlated with the presence of *sopE* with only two exceptions, suggesting that the acquisition of the virulence gene was likely mediated by lysogenic conversion of mTmV. However, a proportion of *sopE*-positive *S*. 4,[5],12:i:- isolates (9.8%) did not harbour mTmV. This may be due to sequence variation of the mTmV prophage over the threshold set in SRST2 (10% of variation over total length of the prophage sequence) or to the presence of a distinct *sopE* bacteriophage (Chapter 6).

4.2.4 A distinct *sopE* allelic variant was associated with *S.* 4,[5],12:i:-

sopE in *S*. Typhimurium SL1344 was the only allele described in the *S*. Typhimurium serovar and represented the prototype of *sopE* on which a large part of the previous studies were based (Schlumberger et al., 2003, Buchwald et al., 2002, Hardt et al., 1998a, Humphreys et al., 2012, Vonaesch et al., 2014, Muller et al., 2009, Mirold et al., 2001b, Lim et al., 2014, Friebel and Hardt, 2000, Ehrbar and Hardt, 2005). To test the relatedness of *sopE* from *S.* 4,[5],12:i:- with previously described allelic forms of *sopE* in *S*. Typhimurium SL1344 and other serotypes, a phylogenetic tree based on synonymous substitutions in the coding sequence of the gene was generated with SNAP (www.hiv.lanl.gov) as previously described (Mirold et al., 2001a). The same *sopE* sequences of a previous study were included (Mirold et al., 2001b) as well as a single gene from each of the allelic variants described in *S*. Typhimurium isolates from human infection in the UK (strains H143740716, H143660450, H142780372, H142520441 and H142940536) .

The *sopE* gene present on the *S.* 4,[5],12:i:- ST34 genomes was highly related, but distinct from the prototypical *sopE* of *S*. Typhimurium SL1344 previously described. The phylogeny of *sopE* was concordant with that previously described

consisting of three *sopE* sequence clusters (Mirold et al., 2001b) (Figure 4.7). Cluster one comprised *sopE* genes from *S.* Typhimurium and *S.* Typhi, cluster II *sopE* from a variety of serotypes (Enteritidis, Dublin, Hadar), whereas cluster III contained sequences from *S. enterica* subspecies IV (SARC9) and VII (SARC16).

The *sopE* gene from *S.* 4,[5],12:i:- strain S04698-09 was closely related to those from *S.* Typhimurium strain SL1344 and *S.* Typhi strains SARC2 and X3744, but distinct by 1-2 synonymous SNPs. The *sopE* sequences from *S.* Typhimurium strain H143740716 (DT44) and strain H143660450 (DT120) were identical to the gene sequences of *S*. Typhimurium strain SL1344 and *S*. Typhi (strains SARC2 and X3744), respectively, which differed by no more than a single SNP. The *sopE* gene from *S.* Typhimurium strain H142780372 (DT40), which had 36 SNPs compared with the sequence of *S*. Typhimurium SL1344, was closely related to the sequences from *S*. Enteritidis strain ATCC BAA-708 and *S*. Hadar strain X3230, in cluster II, from which differed by 1 SNP. This suggested the phage-mediated transfer of the *sopE* allele between phylogenetically distant serovars. The *sopE* gene from *S.* Typhimurium strain H142520441 (DT203) and strain H142940536 (no phage type determined) had 19 SNPs and 28 SNPs relative to *S*. Typhimurium SL1344 respectively, and were present on distinct branches, unrelated from previously described *sopE* genes.

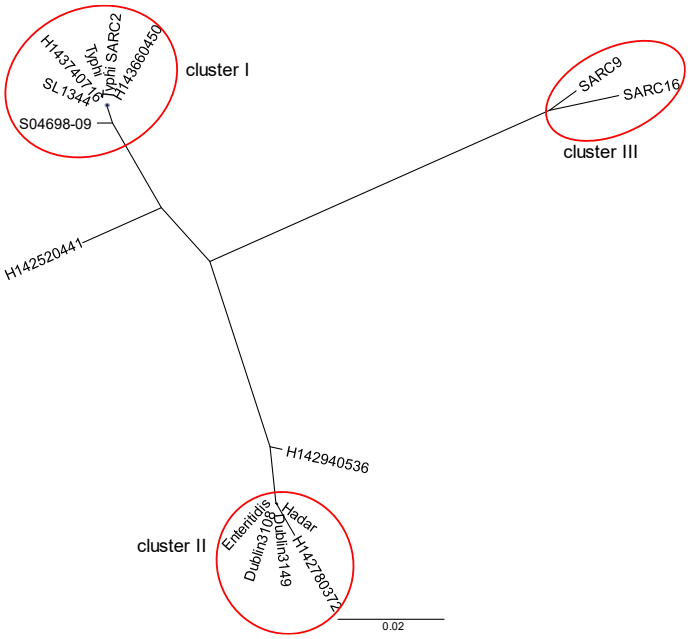

Figure 4.7 Unrooted neighbor joining phylogenetic tree of *sopE* based on SNPs at synonymous sites of the coding sequence. The bar shows the relative genetic distance.

4.2.5 Evidence for the impact of *sopE* acquisition for *S*. 4,[5],12:i:- transmission

The frequent acquisition and increasing frequency of *sopE* in *S*. 4,[5],12:i:- from human clinical infections suggested the advantage associated with the virulence gene. The proportion of clinical *sopE*-positive *S*. 4,[5],12:i:- was investigated over time to test whether *sopE* was maintained and conferred an epidemiological advantage to the isolates reflected in an increasing proportion of isolates with *sopE*.

To avoid potential biases, the analysis was performed on the 737 *S*. 4,[5],12:i:- isolated between April 2014 and March 2015, the period for which all the clinical isolates received by PHE were available. The *sopE* gene was present in 40.7% of the *S*. 4,[5],12:i:- isolates considered, in which it was stably maintained over time. The frequency of *sopE*-positive isolates generally ranged between 26.4% to 46.4%, with the exception of an increase to 70% in March 2015 that was atypical (Figure 4.8). A linear regression analysis suggested a positive trend in the proportion of *sopE*-positive isolates over time, however, the data poorly fit to the regression (R^2=0.2) with no statistically significant association between the percentage of *sopE*-positive isolates and time (p-value: 0.138).

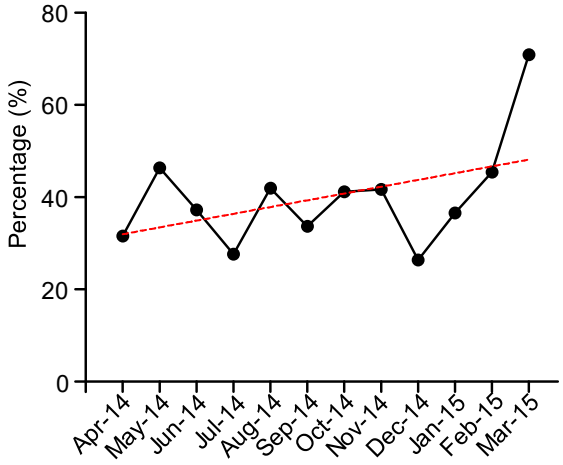

Figure 4.8 Percentage of *S*. 4,[5],12:i:- isolates from UK human salmonellosis cases positive for *sopE* over time. The linear regression is shown as a red dashed line (R^2=0.2, p-value: 0.138).

I then tested whether *sopE* enhanced the success, therefore, the selection of *S*. 4,[5],12:i:- resulting in a clonal expansion of the isolates, which on a phylogenetic tree can be visualised as a cluster of closely related isolates. Assuming that

clonally expanded isolates are distinguished by a lower genetic distance than those which did not undergo clonal expansion, the hypothesis was tested by comparing the mean patristic distance (pairwise genetic distance between tips of a phylogenetic tree) of *S*. 4,[5],12:i:- isolates with and without *sopE*. As human isolates represent a sample of the population of *S*. 4,[5],12:i:- circulating in animals, the analysis of the genetic distance may have implications on the transmission rate of genotypes. If a genotype is associated with increased transmission then closely related isolates belonging to that genotype are expected to be observed on the phylogenetic tree, resulting in lower patristic distance. Thus, the comparison of the patristic distance between *sopE*-positive and *sopE*-negative genotypes can potentially provide lines of evidence in the impact of *sopE* on the transmission of *S*. 4,[5],12:i:- between distinct hosts. Nevertheless, the lower patristic distance may reflect the clonality of the *S*. 4,[5],12:i:- population in pig farms.

Four possible scenarios regarding the impact of *sopE* in animal and human *S*. 4,[5],12:i:- populations were hypothesized (Figure 4.9). If *sopE* is under neutral selection (meaning it does not confer positive nor negative advantage to the isolates) in both pig and human hosts, a similar level of expansion, therefore patristic distance, would be expected for *sopE*-positive and *sopE*-negative lineages in pig isolates and in human isolates as a result of an analogous rate of transmission of the two genotypes. However, if *sopE* is under neutral selection in pigs, but promotes transmission, an enhanced frequency of *sopE*-positive clones is expected to be isolated in humans, therefore a lower patristic distance in human isolates. In the case that *sopE* is positively selected in pigs, a lower patristic distance would be expected as a consequence of the clonal expansion; however, even if *sopE* does not enhance the transmission to humans, a lower patristic distance reflecting the pig population would be expected, and even lower patristic distance if the transmission is enhanced. Therefore, the analysis of the pig population would be required in order to differentiate the scenario.

Also, in order to confidently correlate the observation of a lower patristic distance with enhanced transmission a large number of observations would be required. For this reason the analysis was not performed with pig isolates due to the lack

of a large cohort of sequences representative of the *S*. 4,[5],12:i:- population in pigs. By contrast, we assumed that the collection from PHE, comprising of all the 737 clinical *S*. 4,[5],12:i:- isolates from a 12-month period, would be representative of *S*. 4,[5],12:i:- in human.

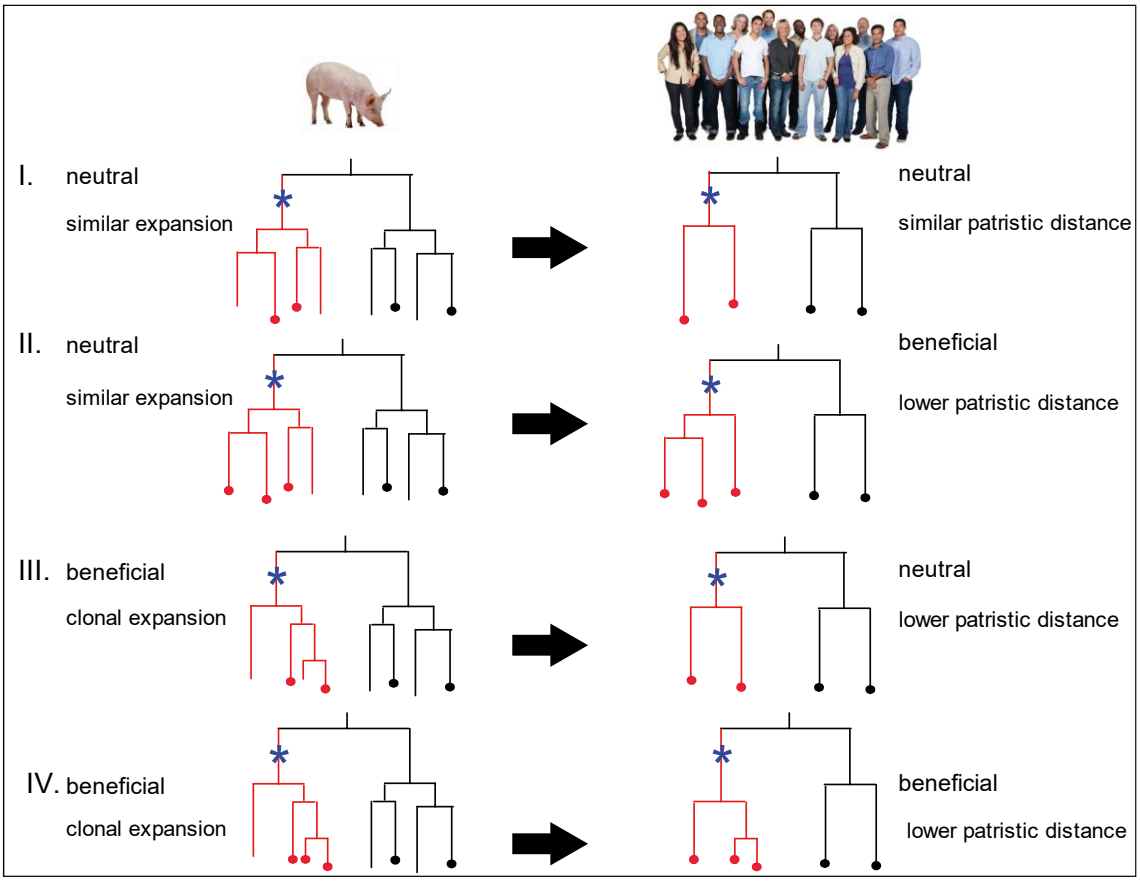

Figure 4.9 Possible impact of *sopE* on fitness of *S*. 4,[5],12:i:- in pigs and on *S*. 4,[5],12:i:- transmission from pigs to humans inferred from the structure of the phylogenetic trees. Four possible scenarios (I-IV) were hypothesized regarding the impact of *sopE* acquisition on the populations of *S*. 4,[5],12:i:- in pigs and in humans. The phylogenetic relationships of the isolates were represented, the acquisition of *sopE* is displayed as an asterisk in red lineage, the isolates of the phylogenetic tree that are transmitted from pigs to humans are indicated with a filled dot. Case I. *sopE* is under neutral selection (does not confer positive nor negative advantage to the isolates) in both pig and human hosts resulting in similar patristic distance. Case II. *sopE* is under neutral selection in pigs (equal patristic distance) but promotes transmission (lower patristic distance) Case III. If *sopE* is positively selected in pigs but does not enhance selection, a lower patristic distance would be expected for isolates of both hosts. Case IV. If *sopE* promotes fitness of *S*. 4,[5],12:i:- in pigs and transmission, even lower patristic distance is expected

The patristic distances were computed using the R package ape (Paradis and Schliep, 2018) on a novel maximum-likelihood phylogenetic tree which contained exclusively the sequences of *S*. 4,[5],12:i:- isolated between April 2014 and March 2015 (Figure 4.10). The mean patristic distance of isolates encoding *sopE* was then compared with that of the isolates lacking the virulence gene.

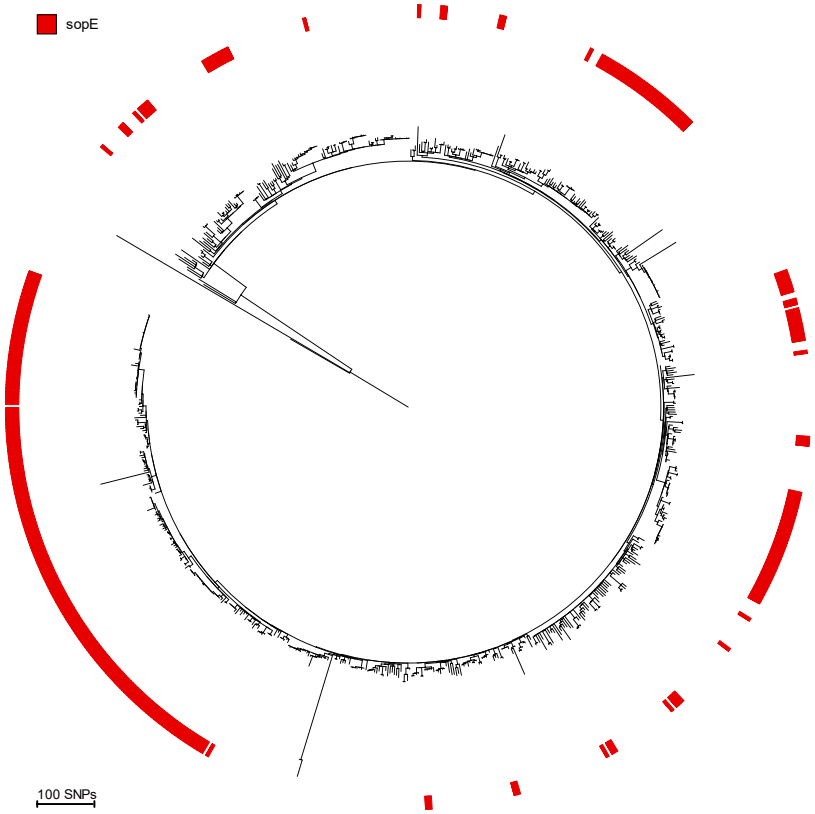

Figure 4.10 Distribution of *sopE* in *S*. 4,[5],12:i:- isolated from human clinical infections in the UK between April 2014 and March 2015. Maximum-likelihood phylogenetic tree of clinical *S*. 4,[5],12:i:- based on 4,423 core genome variant sites in reference to *S*. Typhimurium SL1344 was generated with RAxML v8.2.12 using GTR+Γ substitution model (log likelihood score: -67394.6). The tree was rooted using *S*. Typhimurium strain LT2 as outgroup. The bar indicates the number of SNPs.

The presence of *sopE* correlated with significantly lower patristic distance of the isolates compared with *sopE*-negative isolates supporting potential increased selection and transmission. The isolates harbouring *sopE* exhibited a mean patristic distance of 6.2 SNPs, whereas the mean patristic distance of *sopE*-negative isolates was almost twice (11.8 SNPs) (Figure 4.11). The mean

difference between the two groups of isolates was statistically significant (Mann-Whitney test, p-value<0.0001).

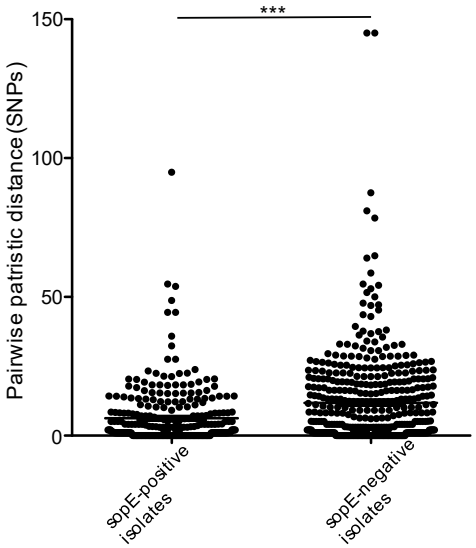

Figure 4.11 Distribution of patristic distances of *sopE*-positive and *sopE*-negative clinical *S*. 4,[5],12:i:- isolates. The patristic distance is expressed in SNPs in accordance with the scale of the phylogenetic tree. The mean patristic distances between isolates with and without *sopE* gene are statistically significant (p-value <0.0001).

The significant difference in mean patristic distance strictly depended on the distribution of *sopE* and was not observed by chance as suggested by a randomisation test with 5,000 replicates which was performed by changing the distribution of *sopE* across the phylogeny but conserving the proportion *sopE*-positive isolate. The difference between the average patristic distance of *sopE*-negative and *sopE*-positive *S*. 4,[5],12:i:- isolates did not exceed 4 SNPs (Figure 4.12). The p-value, calculated as how many times the simulated difference exceeded the observed one, was zero.

The lower patristic distance of isolates encoding *sopE* indicated that the acquisition of the gene was associated with clonal expansion, therefore, it was beneficial for *S*. 4,[5],12:i:- circulating in the human population. The results suggested that scenarios II, III and IV depicted in Figure 4.9 are plausible; however, in order to define the most likely scenario the analysis of the population structure of *S*. 4,[5],12:i:- circulating in pigs would be necessary.

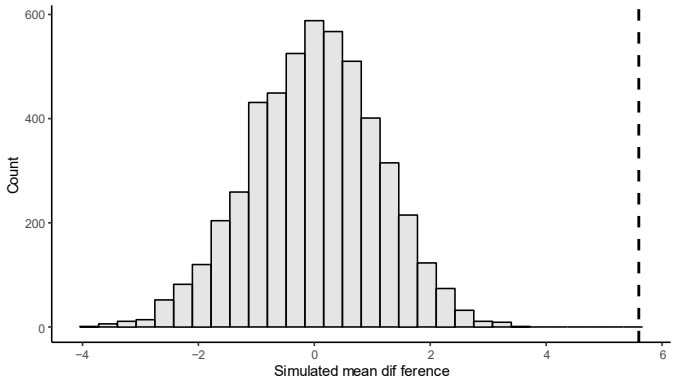

Figure 4.12 Distribution of simulated mean difference of patristic distance of *sopE*-negative and *sopE*-positive *S*. 4,[5],12;i;- isolates. The dashed line indicates the difference in the mean patristic distances observed in the real data between *sopE*-negative and *sopE*-positve *S*. 4,[5],12:i:- isolates. 5,000 permutations were performed. p-value: 0.

As the methodology was tree-dependent and, therefore, relied on the accuracy of the phylogenetic reconstruction, a tree-independent method was performed to further investigate the degree of similarity of *sopE*-positive and *sopE*-negative *S*. 4,[5],12:i:- isolates. The snp-dists v-0.2 software (Seemann, 2017) was used to identify the number of pairwise SNPs identified from a multiple sequence alignment previously generated with snippy v3.0 (Seemann). The average pairwise SNP-distance between *S*. 4,[5],12:i:- isolates encoding *sopE* was 3.4 SNPs compared with an average of 7.3 SNPs for *sopE*-negative isolates, a difference that was statistically significant (Mann-Whitney test, p-value < 0.0001). Over the 70% of *sopE*-positive isolates had no SNP difference between each other, whereas only the 33% of *sopE*-negative isolates were undistinguishable (Figure 4.13). The results of the analysis were concordant with the analysis of the

patristic distances of *sopE*-positive and *sopE*-negative isolates.

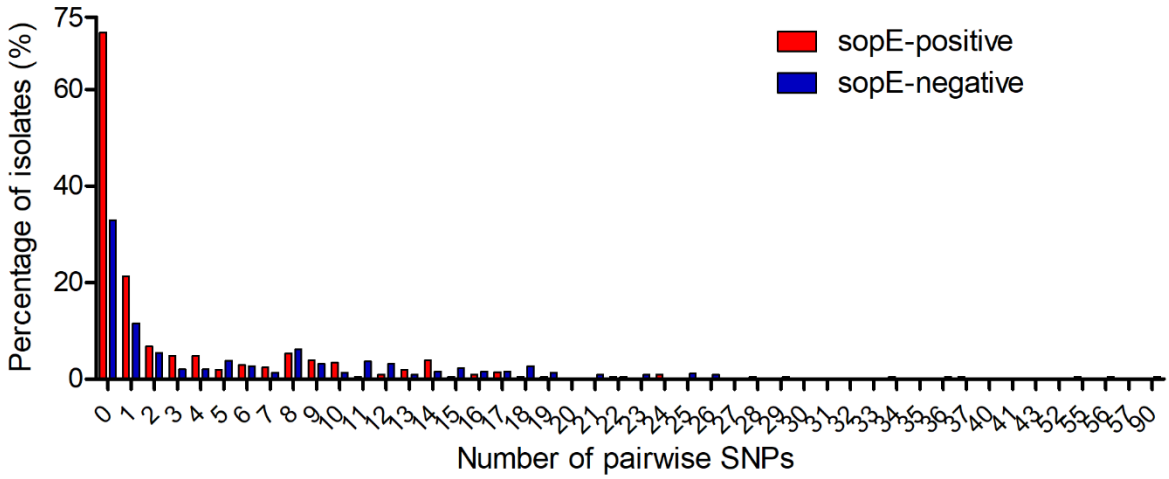

Figure 4.13 Distribution of pairwise SNP differences of *sopE*-positive and *sopE*-negative *S*. 4,[5],12:i:- isolates. Bar chart showing the percentage of *sopE*-positive and *sopE*-negative isolates for each pairwise SNPs distance value.

To further study the epidemiological impact associated with the acquisition of *sopE*, the tendency of *sopE*-positive and *sopE*-negative genotypes to be transmitted as part of a cluster was compared. I hypothesized that if *sopE*-positive isolates display an enhanced transmission rate compared with *sopE*-negative isolates, they will have an enhanced tendency to be transmitted as part of a transmission cluster, namely a group of isolates that are epidemiologically related. As microorganisms within a transmission cluster are more similar to each others than those outside the cluster, the genetic relatedness of the isolates can be used for the identification of the transmission clusters. Therefore, putative transmission clusters were defined using the SNP address generated by PHE.

The SNP address method defines the relatedness of the isolates to a common reference strain at the SNP level, identified by single linkage hierarchical clustering (Dallman et al., 2018). Clusters constituted by four or more isolates which did not differ by more than 5 single linkage SNPs, the same cut-off used by PHE to define potential epidemiologically linked cases, were used for the analysis. Overall, 34 clusters were identified, of which 21 associated with *sopE*-negative *S*. 4,[5],12:i:- and the remaining 13 with *sopE*-positive *S*. 4,[5],12:i:-. The size of the transmission clusters associated with *sopE*-positive isolates was greater than those associated with *sopE*-negative isolates, but not significantly different (Figure 4.14; Mann Whitney U test p-value 0.08).

However, the proportion of *sopE*-positive isolates transmitted as part of a cluster was significantly higher than that of isolates without *sopE* (56.2% and 44.1%, respectively; Fisher's exact test p-value 0.001) suggesting that clones of *S*. 4,[5],12:i:- with *sopE* had enhanced transmission fitness.

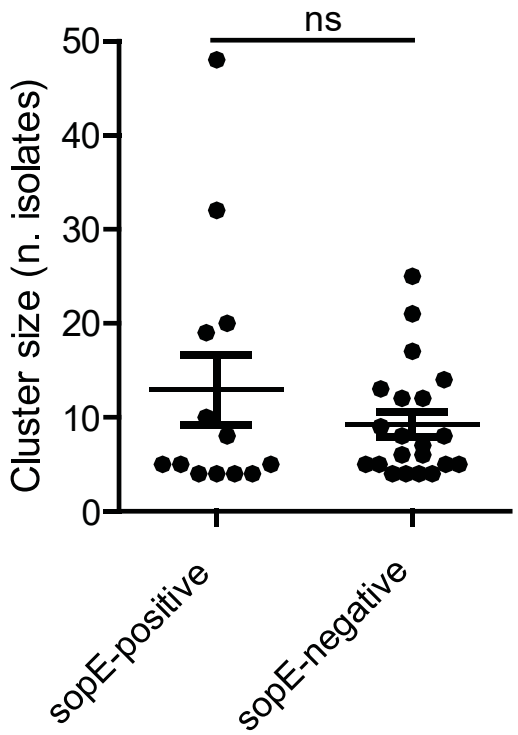

Figure 4.14 Scatterplot showing the distribution of the size of the transmission clusters for *sopE*-positive and *sopE*-negative *S*. 4,[5],12;i;- isolates from UK human infections. Transmission clusters were defined as groups with at least 4 of isolates sharing the same SNP address at a 5 SNPs level. The size of the clusters of the two groups was not significant different (Mann Whitney U test p-value 0.08).

In conclusion, the higher degree of relatedness of *sopE*-positive isolates together with their greater tendency to be part of potential transmission clusters suggested that the presence of *sopE* is associated with enhanced transmission fitness of *S*. 4,[5],12;i;- to humans.

4.3 Discussion

sopE is sporadically distributed and associated with few lineages of *S*. Typhimurium which were or are currently epidemic. In a collection of representative *S*. Typhimurium spanning a large variety of phage types isolated over a wide range of years and from various sources, the virulence gene was associated with only two lineages, namely with the previously described *S*. Typhimurium DT204 complex epidemic in the 1970-80s as well as with the current pandemic *S*. 4,[5],12:i:- sequence type ST34 (Mirold et al., 1999, Petrovska et al., 2016).

Since 2005 *S*. 4,[5],12:i:- ST34 progressively emerged from UK pigs rapidly becoming epidemic (Petrovska et al., 2016). It is currently the most common *Salmonella* serovar circulating in swine: in 2017 *S*. 4,[5],12:i:- ST34 accounted for 57.2% of total *Salmonella* isolates from pig samples in the UK (APHA, 2018b). At least seven distinct events of *sopE* acquisition followed by clonal expansion occurred in porcine *S*. 4,[5],12:i:- from the UK, in accordance with the previous literature (Petrovska et al., 2016), indicating that the acquisition of the gene was positively selected. The fact that one of the first events of acquisition of *sopE* was estimated to have occurred between 2000 and 2005 suggested that the acquisition of the gene may have contributed to the emergence of *S*. 4,[5],12:i:- ST34 from the pig chain.

The beneficial impact conferred by *sopE* was also suggested by the high proportion of human *S*. 4,[5],12:i:- isolates received by PHE positive for the virulence determinant (40.7%), which was stably detected over time. The chromosomal integration of the mTmV prophage was responsible for the acquisition of the virulence gene in 90% *S*. 4,[5],12:i:- isolates. By contrast, only a small proportion of clinical *S*. Typhimurium isolates (2.7%) carried *sopE*. Notably, the virulence determinant was not exclusively found in *S*. Typhimurium related to the epidemics previously described (Mirold et al., 1999), but also in *S*. Typhimurium types which have not previously been associated with *sopE*. Nonetheless, these were sporadic cases constituted by 1-2 isolates.

sopE associated with *S*. 4,[5],12:i:- differed from the prototype of *S*. Typhimurium SL1344 by a single synonymous SNP at nucleotide 27. While all the *S*. 4,[5],12:i:- harboured the same *sopE* allele, *S*. Typhimurium isolates carried up to five distinct *sopE* alleles. To the best of my knowledge, only a single form of *sopE*

identical to that of *S*. Typhimurium SL1344 had been described in *S*. Typhimurium, so far (Hopkins and Threlfall, 2004, Mirold et al., 2001b), indicating that a repertoire of *sopE* wider than that previously described in the literature is circulating in *S*. Typhimurium. The *S*. Typhimurium isolates belonging to phage types previously not associated with *sopE* were those carrying polymorphic forms of *sopE* compared to the prototypical one of SL1344. The impact of sequence variation of the functionality of the gene is not known.

The high proportion of *sopE* in clinical *S*. 4,[5],12:i:- over time indicated a beneficial effect of *sopE* acquisition on the success of *S*. 4,[5],12:i:-. This was further supported by the finding that *S*. 4,[5],12:i:- isolates with *sopE* were more closely related than *sopE*-negative *S*. 4,[5],12:i:- isolates as measured both by patristic distance and pairwise SNP-distance. In absence of selection, a similar level of genetic relatedness would have been expected for *sopE*-positive and *sopE*-negative isolates. Importantly, the observed difference in clonality was strictly associated with *sopE* distribution within the *S*. 4,[5],12:i:- population.

The higher genetic relatedness of human isolates encoding *sopE* suggests that they were sampled more frequently from the *S*. 4,[5],12:i:- population circulating in animals, namely they were transmitted at a higher rate. The use of patristic distance as a metric of evolutionary distances between isolates has been historically used to infer transmission in viruses which evolve quickly but more recently applied to microbial phylogenies thanks to the enhanced resolution achieved by SNP-based phylogenies (Poon, 2016). However, the lower patristic distance in *S*. 4,[5],12:i:- from human infections may be the reflection of the fitness advantage that *sopE* confers to the microorganisms in the animal reservoir. The analysis of the population structure of *S*. 4,[5],12:i:- circulating in livestock may help us understanding whether *sopE* either promotes the fitness of the bacteria in the animal reservoir or confers transmission fitness from animals to humans, or both.

Nevertheless, the analysis of the potential transmission clusters identified basing on the SNP address suggested that *S*. 4,[5],12:i:- encoding *sopE* were significantly more prone to be transmitted as part of a transmission cluster suggesting that individual *S*. 4,[5],12:i:- clones carrying *sopE* were more commonly responsible for multiple cases of human salmonellosis than those lacking *sopE*.

4.4 Conclusion

This study showed the strong association of a novel allelic variant of *sopE* with the current MDR pandemic *S*. 4,[5],12:i:- ST34. The contribution of the virulence gene to the emergence of the pandemic was suggested by the acquisition early after *S*. 4,[5],12:i:- emergence in pig farms and by the multiple events of *sopE* acquisition followed by clonal expansion. The benefit associated with *sopE* was reflected by the high proportion of *sopE*-positive clinical isolates together with the maintenance of the virulence gene over time. The acquisition of *sopE* can potentially promote the transmission of *S*. 4,[5],12:i:- to humans suggesting the contribution of *sopE* to the dissemination of *S*. 4,[5],12:i:-. The investigation of the genetic elements responsible for the mobilisation of *sopE* will then be important to understand how the dissemination of the gene occurs.

5 The dissemination of the *sopE* bacteriophage mTmV to diverse *S. enterica* serotypes

5.1 Introduction

Temperate bacteriophages are major drivers of microbial evolution (Brussow et al., 2004). Because of their dual life cycle, they establish both a parasitic and mutualistic relationship with the host cell (Harrison and Brockhurst, 2017). As parasites, they exploit microbial cell machinery to replicate and produce new virions with the consequent lysis of the host cell (lytic cycle). Alternatively, they can integrate into the genome of the host as a prophage and be transmitted vertically to the progeny cells following cell replication (lysogenic cycle). The lysogenic conversion by phages harbouring cargo genes (genes not required for the phage cycle) may increase the fitness and the virulence of the host cells (Brussow et al., 2004). Nevertheless, the establishment of the lysogeny can influence the biology of the host in other ways, such as gene disruption following phage integration or genomic rearrangements resulting from prophages recombination (Brussow et al., 2004).

Multiple examples of phage-encoded toxins have been described, including the *Vibrio cholerae* toxin, the Diphteria toxin acquired by *Corynebacterium diphtheriae* and the neurotoxin encoded by *Clostridium botulinum* whose phage-mediated acquisition transforms non-pathogenic ancestors to highly virulent bacteria (Brussow et al., 2004). In *S.* Typhimurium the phage-encoded guanine nucleotide exchange factor (GEF) SopE was associated with the clonal complex DT204 epidemic in the 1970-80s and the current pandemic *S*. 4,[5],12:i:- ST34 (Chapter 4). The *sopE* gene is rare in *S*. *enterica* but has been reported in organisms from various subspecies and *S*. *enterica* subspecies I serovars (Hardt et al., 1998b, Mirold et al., 2001a). A phylogenetic analysis of *sopE* from different subspecies and serovars clustered the sequences in three distinct groups (Figure 5.1) (Mirold et al., 2001b). *sopE* described in *S*. Typhimurium SL1344 was highly related to *sopE* in *S*.Typhi but distinct from the gene in various *S*. *enterica* serovars (*S*. Dublin, *S*. Hadar and *S*. Gallinarum); an additional cluster was constituted by *sopE* sequences in subspecies IV and VII.

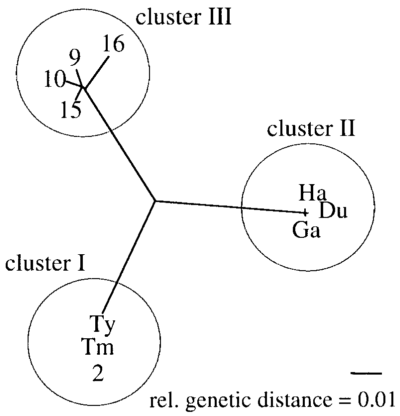

Figure 5.1 Phylogenetic tree based on nucleotide polymorphisms at synonymous sites of the *sopE* coding sequence (Mirold et al., 2001b). The sequences were from *S*. *enterica* subspecies I serovar Dublin (Du), serovar Hadar (Ha), serovar Gallinarum (Ga), serovar Typhimurium (Tm), serovar Typhy (Ty) and serovar Typhi SARC2 (2); *S*. *enterica* subspecies IV strain SARC9 (9) and SARC10 (10); *S*. *enterica* subspecies VII strains SARC15 (15) and SARC16 (16).

Notably, the genetic context of *sopE* from each cluster varied but the regions flanking the genes were conserved suggesting the transfer of the *sopE* cassette between distinct mobile genetic elements by homologous recombination (Mirold et al., 2001b). In *S*. Typhimurium SL1344 *sopE* was harboured by a P2-like phage, termed SopEΦ, following integration at the 3' terminus of *ssrA* which encodes for a tmRNA (Pelludat et al., 2003). In *S*. Dublin, Hadar and Gallinarum *sopE* located in a cryptic lambdoid bacteriophage, whereas in subspecies IV and VII *sopE* was in a chromosomal region not homologous to any known prophage (Mirold et al., 2001b).

In the vast majority (90%) of *S*. 4,[5],12:i:- ST34 isolates from human infections in the UK, *sopE* was carried on the genome of the mTmV prophage (Chapter 4) (Petrovska et al., 2016).The genome of the mTmV prophage, 39,034 base pairs long and encoding for 56 ORFs, was integrated into the threonine tRNA locus (*thrW*) of S04698-09, a well-known integration hotspot (Bobay et al., 2013, Petrovska et al., 2016). Due to the role played by the mTmV prophage in the dissemination of the *sopE* virulence gene in *S*. 4,[5],12:i:- ST34, I characterised mTnV through comparative genomics in order to study its gene content and gain an insight about its biology. Because of the association of mTmV with the clinical isolates of *S*. 4,[5],12:i:- ST34 from the UK, I then tested the hypothesis that

mTmV is able to disseminate in diverse *S*. Typhimurium and isolates of other *S*. *enterica* serovars using both *in vitro* and *in silico* approaches.

5.2 Results

5.2.1 mTmV is related to *Myoviridae* phages

In order to study the gene content and genomic organisation of mTmV, a detailed analysis of the nucleotide sequence of mTmV was performed by inspecting the annotated genome and aligning it to the closest bacteriophage identified in GenBank.

The mTmV prophage is closely related to promiscuous myoviruses described in *Shigella* and *Salmonella* (Figure 5.2). Approximately 75% of the mTmV genome was highly related to *Shigella flexneri* prophage V (SfV) (98.32 % average nucleotide identity). SfV is integrated in the same tRNA locus in *Shigella flexneri* and associated with serotype conversion due to expression of enzymes (*gtrV, gtrB, gtrA*) that modify LPS structure (Allison et al., 2002). A comparison of the genomes of mTmV and SfV using BLAST (Altschul et al., 1990) showed that the two prophages shared genes involved in integration/excision, immunity and regulation, replication, DNA modification (DNA methylase) and recombination, late regulation and lysis as well as DNA packaging genes and tail proteins and assembly (Figure 5.2). SfV carried genes homologous to the repressors cI and Cro in lambdoid bacteriophages which regulate the expression of early genes and are responsible for the establishment of the lysogenic or lytic cycle, respectively (Allison et al., 2002).

The tail fibre genes of SfV, conserved in mTmV, encoded a long contractile tail, typical of the *Myoviridae* (group A1 morphology) (Allison et al., 2002). However, the capsid proteins of SfV were homologous to *Siphoviridae* (Allison et al., 2003). The mTmV prophage differed from SfV by the presence of distinct capsid genes and the lack of O-antigen modification genes, responsible for serotype conversion in *Shigella flexneri* (Allison et al., 2002). Conversely, SfV lacked a 4.5 kb region encoding a tail fibre assembly protein, the *sopE* gene, a *hin* DNA invertase, a gene with sequence similarity to *flxA* and a hypothetical protein (Figure 5.2). The *flxA* gene encodes a 110-amino acid protein of unknown function whose expression is dependent on the sigma factor specific for flagellum FliA (Ide and Kutsukake, 1997).

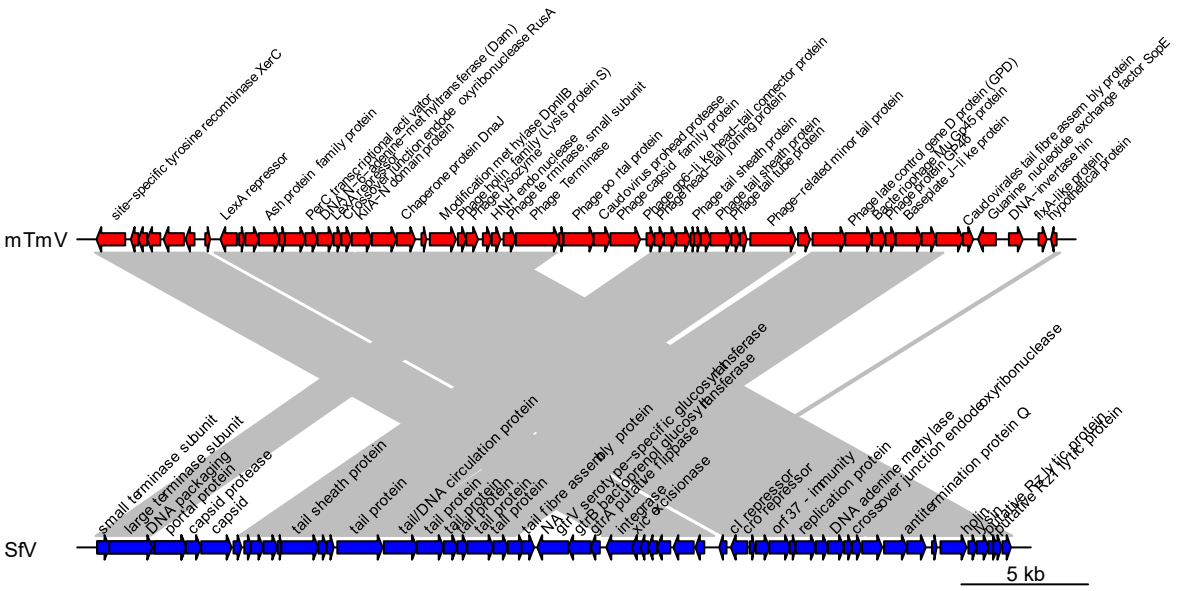

Figure 5.2 Genome comparison between mTmV and SfV phages. The genes unlabelled encode for hypothetical proteins (>90% sequence identity).

In addition, mTmV shared a high degree of homology with prophages present elsewhere in the genome of S04698-09 (Figure 5.3). These prophages were an ST64B-like prophage that had 83.71% nucleotide identity over 51% of the mTmV prophage, and an SJ46-like prophage that had 99.33% nucleotide identity over 75% of the mTmV genome.

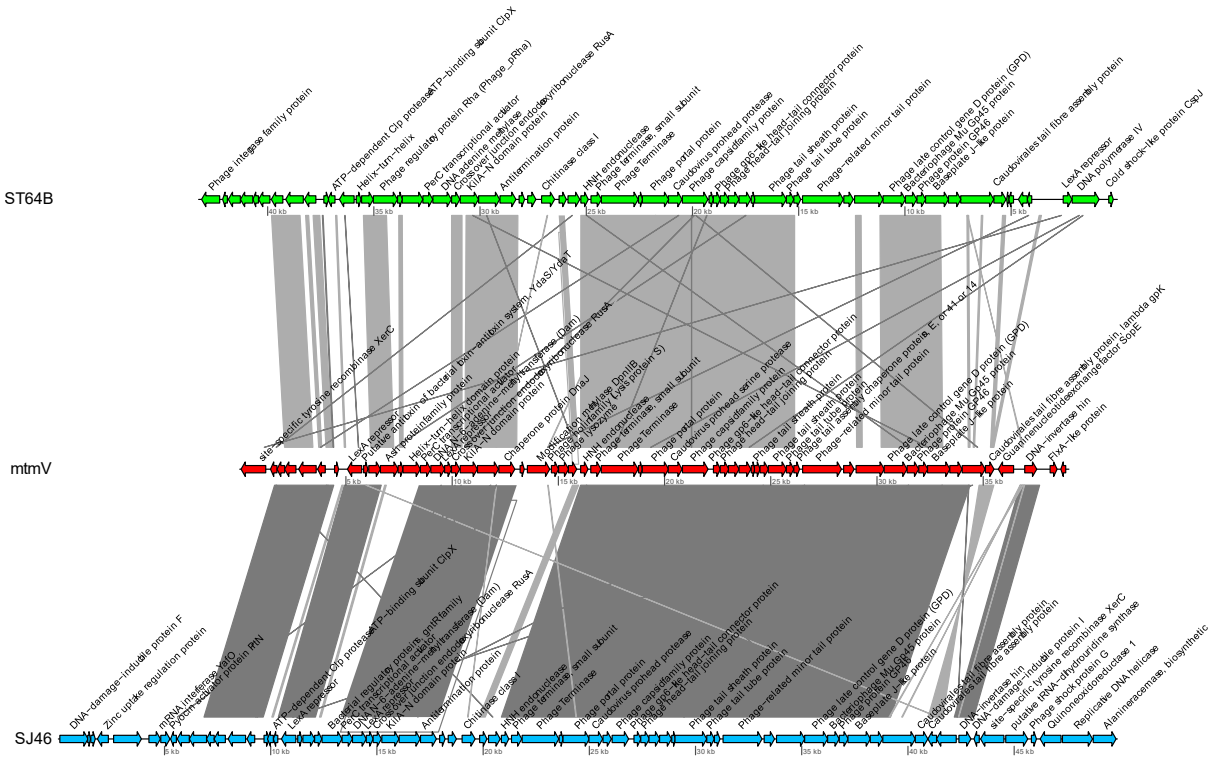

Figure 5.3 Genomic comparison between the genomes of bacteriophages ST64B, mTmV and SJ46 integrated in the chromosome of *S.* 4,[5],12:i:- S04698-09. The gene function is reported; when absent, the gene encodes for a hypothetical protein. Different shades of grey indicate different nucleotide identity. Dark grey: identity >90% nucleotide identity; light grey: 64% - 89% nucleotide identity.

5.2.2 mTmV produced active phage particles able to lysogenise new susceptible host cells *in vitro*

As phylogenomic analyses suggested that mTmV was acquired on multiple occasions by *S.* 4,[5],12:i:- ST34, I tested the ability of mTmV to propagate in the host cell and infect susceptible microbial cells by establishing the lysogeny.

In order to detect the transfer of mTmV, mutants of *S.* 4,[5],12:i:- isolate 0343A were constructed in which the mTmV prophage carried a resistance cassette conferring resistance to either chloramphenicol (*cat* gene) or kanamycin (*aph* gene). The antimicrobial resistance genes were inserted in the intergenic region downstream of *sopE*, between the 3' end of *sopE* and the adjacent gene that encoded a tail fibre protein (Figure 5.4). The two mutant strains generated, termed 0343A mTmV_cat and 0343A mTmV_kan, were resistant to chloramphenicol and kanamycin, respectively. The insertion sites were confirmed by amplification across the region by PCR (data not shown).

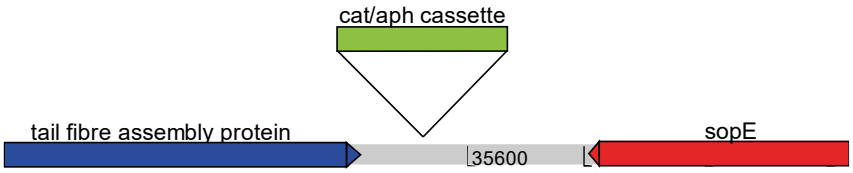

Figure 5.4 Insertion of cassette conferring resistance to either chloramphenicol (*cat*) or kanamycin (*aph*) into mTmV prophage genome. The nucleotide coordinates on the prophage genome are reported.

To test whether the resistance genes of strains 0343A mTmV_cat and 0343A mTmV_kan were transferred to a recipient strain following mTmV excision and virions production, the supernatant from an overnight liquid culture of the strains was filtered with a 0.22 µm pore size membrane to remove bacteria and mixed with *S*. Typhimurium SL1344 that lacked the mTmV prophage. Colonies of *S*. Typhimurium SL1344 lysogenised by the mTmV prophage were selected by plating on LB agar supplemented with chloramphenicol or kanamycin. The prophage integrated into the tRNA-*ThrW* locus as tested by PCR.

These data are consistent with the ability of mTmV prophage to replicate in the host cell and form viral particles which can infect susceptible host cells and establish the lysogeny by integrating into a conserved genomic location. The experiment also showed that mTmV prophage could be spontaneously induced by the 0343A strain without the use of stressors, such as mitomycin C or UV irradiation, which promote the prophage excision (Raya and H'bert, 2009).

5.2.3 Lysogenic conversion with mTmV occurs at variable frequency in diverse genotypes of *S*. Typhimurium *in vitro*

Lysogenic conversion with mTmV occurred at high frequency in *S.* 4,[5],12:i:- ST34 during clonal expansion. However, virtually no *S.* Typhimurium from a collection of diverse genotypes and a collection from UK clinical infections (Chapter 4) harboured the phage suggesting that the mTmV host was restricted to *S.* 4,[5],12:i:- ST34. To test this hypothesis, I performed *in vitro* phage transfer experiments assessing the ability of mTmV to lysogenise diverse *S.* Typhimurium and diverse serovars of *S. enterica* subspecies *enterica* isolates (Appendix 5).

The mTmV phage established lysogeny in all *S*. Typhimurium tested and in two *S*. *enterica* serovars other than Typhimurium. The experimental procedure consisted in mixing the filtered supernatant from overnight cultures of 0343A

mTmV_cat or 0343A mTmV_kan with the recipient. Integration of mTmV into the genome of the recipient strain was monitored by the acquisition of antimicrobial resistance. The rate of lysogenic conversion by mTmV was expressed as the ratio between lysogens per recipient colony units.

The *S.* Typhimurium isolates belonged to distinct genotypes, which may be discriminated by phage typing and sequence typing, in order to encompass the genotypic diversity within the *S*. Typhimurium serovar (Branchu et al., 2018). The strain collection included *S.* Typhimurium SL1344 (DT44), related to the clonal complex DT204 epidemic in the 1970s and 1980s (Mirold et al., 1999), *S.* Typhimurium strain NCTC 13348 (DT104) epidemic in the 1990s (Rabsch et al., 2011) and strain D23580 (ST313), belonging to a human-adapted lineage associated with invasive non-typhoidal salmonellosis in sub-Saharan Africa (Okoro et al., 2012). *S*. Typhimurium 01960-05 (U288), S07292-07 (U308) and 11020-1996 (DT193) belonged to a lineage circulating in pigs as *S.* 4,[5],12:i:- ST34 (APHA, 2018b). Additionally, several animal-associated genotypes were tested, such as *S.* Typhimurium strain L01157-10 (DT8) associated with ducks/geese (Noble et al., 2012), *S.* Typhimurium 03-715 (DT2) and 2882-06 (DT99) associated with wild birds/pigeons (Pennycott et al., 2006), *S.* Typhimurium SO7676-03 (DT56) and 4179-2001 (DT121) associated with passerines (Pennycott et al., 2006).

Although lysogenic cells were obtained for all the *S.* Typhimurium tested, the frequency of lysogenic conversion varied by more than 1000-fold (Figure 5.5 and Figure 5.6). *S.* Typhimurium 11020-1996 (DT193), L01157-10 (DT8), S07292-07 (U308) and NCTC 13348 (DT104) exhibited the lowest rate of lysogenic conversion (10^{-7} lysogens/total recipients), while strain 2882-06 (DT99) had the highest rate (10^{-3}). An intermediate rate of lysogenic conversion comprised between 10^{-6} and 10^{-5} was observed for the remaining organisms.

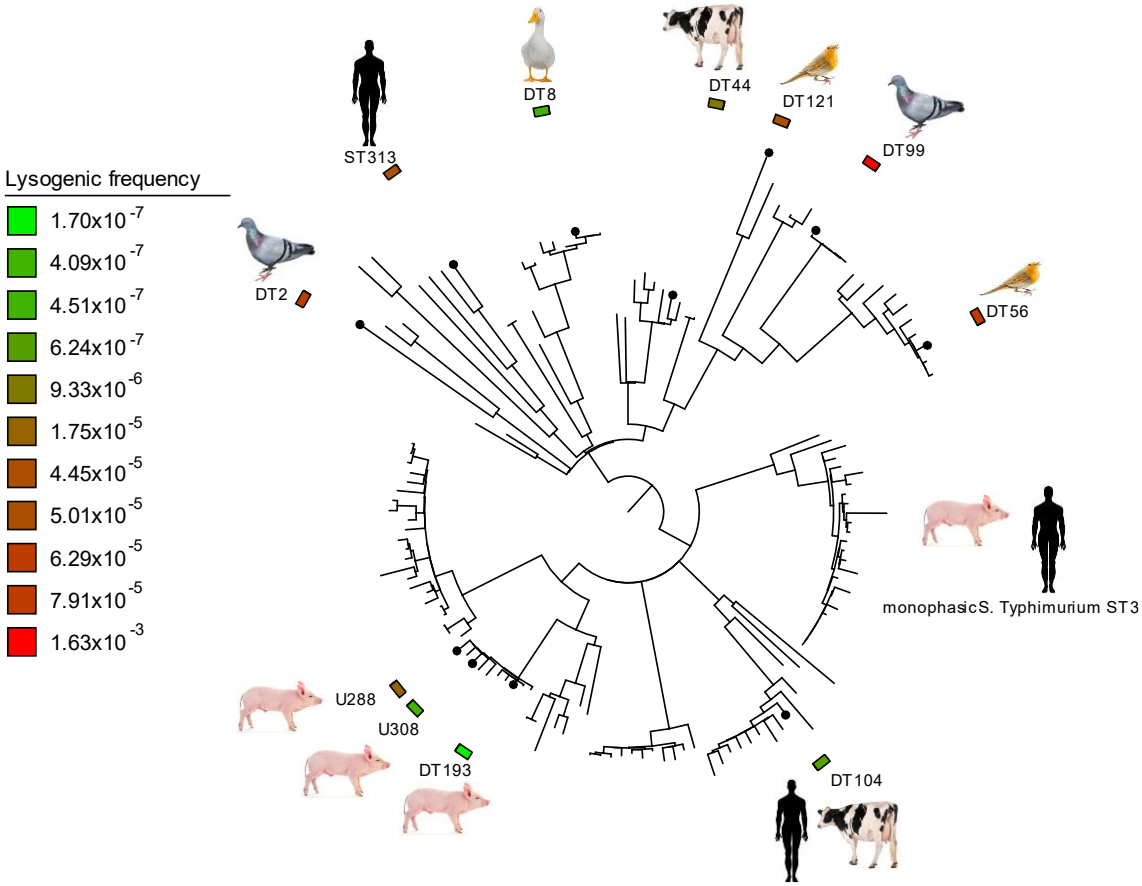

Figure 5.5 Variation of mTmV lysogenic conversion frequency across *S.* Typhimurium phylogeny. The maximum-likelihood phylogenetic tree (same as Figure 4.1) comprised 142 strains representative of the genetic diversity within the serovar. The isolates tested for susceptibility to mTmV, marked with a black dot, belonged to distinct lineages, distinguished by phage type/sequence type and distinct hosts. The colour gradient of the boxes refers to the distinct frequencies of mTmV lysogeny, ranging from green (low frequency) to red (high frequency).

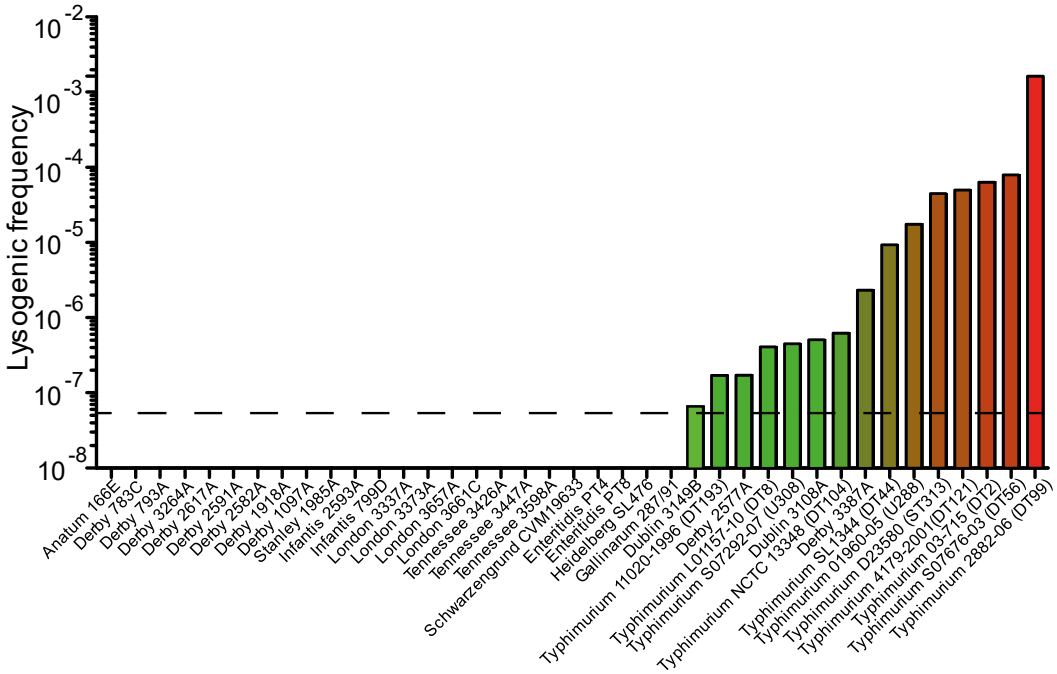

Figure 5.6 Frequency of lysogenic conversion of mTmV in diverse *S. enterica* serovars and *S.* Typhimurium genotypes. The limit of detection is shown as a dashed line.

In addition, the ability of mTmV to infect and integrate in diverse serotypes of *S. enterica* subspecies *enterica* was tested (Appendix 5). Most of the organisms tested were isolated from Irish farms of *S.* 4,[5],12:i:- described in Chapter 3 (Burns, 2015). These included isolates belonging to serovars Anatum (n=1), Derby (n=10), Dublin (n=2), Infantis (n=2), London (n=4), Stanley (n=1) and Tennesse (n=3). In addition, organisms belonging to serovars Enteritidis phage type PT4 (SGSC4901) and phage type PT8 (LK5) (accession number ERS673772), Gallinarum (287/91) (accession number AM933173), Heidelberg (SL476) (accession number NC_011083.1) and Schwarzengrund (CVM19633) (accession number CP001127.1) were also tested.

Of ten *S.* Derby tested, two were lysogenised by mTmV, with a frequency of between 10^{-6} and 10^{-7}. Lysogenic from two *S.* Dublin isolates arose with a frequency of 10^{-7}-10^{-8} lysogens/total cells (Figure 5.6).

The presence of the mTmV prophage and its integration within the *thrW* locus of the recipients was tested by PCR amplification of regions flanking mTmV attachment sites of mTmV (attL and attR). A combination of five oligonucleotides was used (Figure 5.7). A 19-kb island is commonly integrated in the same tRNA locus in *S.* 4,[5],12:i:- and flanking mTmV (Trupschuch et al., 2010). Despite the

island having not been identified in *S.* Typhimurium nor other serovars, so far, an oligonucleotide complementary to it was included in the PCR analysis.

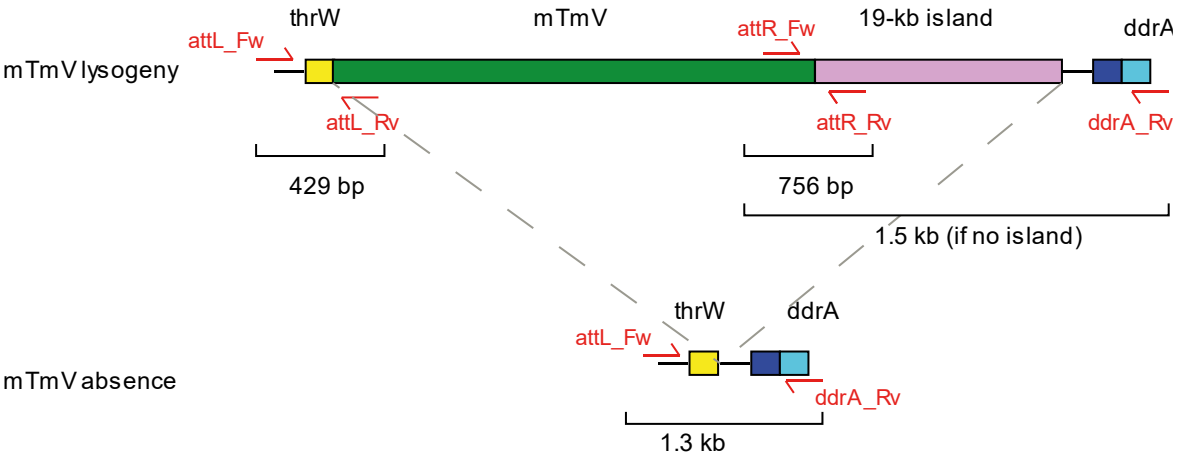

Figure 5.7 Representation of the *thrW* locus with and without the mTmV prophage showing the oligonucleotides binding sites used for the PCR screening and the expected amplicon sizes. The oligonucleotides used to investigate the integration of mTmV into *thrW* of recipient strains are showed as red arrows. The 19-kb genomic island is associated with *S.* 4,[5],12:i:- ST34 isolates (Trupschuch et al., 2010).

The amplification of the empty tRNA locus and attachment sites was variable due to sequence variation across the isolates. Most of the wild type isolates amplified the tRNA locus and all the lysogens amplified the attL sites, indicating that the integration of the phage always occurred in the tRNA for threonine. Conversely, the attR site was not always amplified due to variation downstream *thrW*. Remarkably, a few lysogenic isolates, but not the correspondent wild type, amplified the attR site using the oligonucleotide aligning within the 19-kb *thrW*-associated island (Figure 5.8). This indicated that mTmV may incorporate flanking regions when excising and making new virions, which are then incorporated in the new recipient cell. The extent of the transduced region is not known.

The strains which exhibited this phenomenon were *S.* Derby strain 2617A and *S.* Typhimurium strain 4179-2001 (DT121) (Figure 5.8). Nonetheless, only a few colonies for each lysogen were analysed by PCR, and the occurrence in other isolates cannot be ruled out.

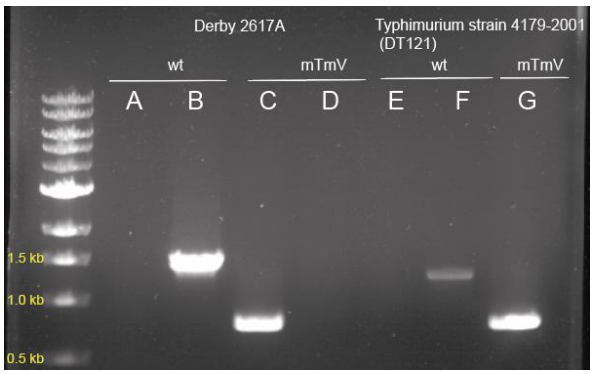

Figure 5.8 Transduction of mTmV flanking regions. Agarose gel probing amplification from within 19-kb thrW-associated island using different pairs of primers. Wells A and E: attL_Fw–attR_Rv (expected product size 750 bp); wells B and F: attL_Fw–ddrA_Rv (expected product ~1.3-1.5 kb); wells C & G: attR_Fw–attR_Rv (expected product size 750 bp); well D: attR_Fw– ddrA_Rv (expected product size ~1.5 kb).

S. London (3373A, 657A and 3661C), *S*. Stanley (1985A) and *S*. Tennessee (3426A) did not form lysogens nor amplified the *thrW* locus, suggesting either lack of the locus, the presence of polymorphisms in the primer binding region or its occupancy by a distinct genetic element. Nevertheless, the presence of a genetic element in the *thrW* site did not prevent integration of mTmV. In the case of *S*. Typhimurium NCTC 13348 (DT104), for instance, the analysis of long-read whole genome sequences showed that the locus was already occupied by the ST104 prophage (accession number NC_005841.1) which was likely replaced by mTmV in the lysogenic cells selected (Figure 5.9). Similarly, *S*. Typhimurium D23580 (ST313) and SO7676-03 (DT56; agarose gel not shown) harboured the BTP1 prophage (accession number FN424405.1) in the *thrW* locus which prevented the amplification of the target sequences, but the candidate lysogens selected for the analysis amplified bands of the predicted size (Figure 5.9).

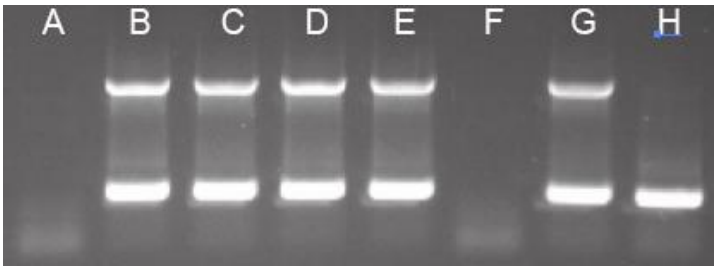

Figure 5.9 PCR targeting the *thrW* locus of wild type (well A) and lysogens (wells B-E) of *S*. Typhimurium DT104 strain NCTC13348 and wild type of *S*. Typhimurium ST313 strain D23580 and lysogens (G-H). The Upper band (approximately 1.5 kb)

represents the amplified right attachment site of mTmV while the lower band (approximately 450 bp) is the amplified left attachment site of mTmV.

The previous occupancy of *thrW* by prophages was not responsible for variation in the frequency of lysogenic conversion. *S*. Typhimurium L01157-10 (DT8) did not harbour a prophage at the *thrW* locus but was lysogenised by mTmV at a frequency as low as that of strain NCTC 13348 (DT104) (approximately 10^{-7} lysogens per recipient cell). Despite harbouring the BTP1 prophage in the *thrW* locus, isolates D23580 (ST313) and SO7676-03 (DT56) formed lysogens at approximately 100-fold higher frequency than strain L01157-10 (DT8) and NCTC 13348 (DT104) (approximately 10^{-5} lysogens per recipient cell). *S*. Typhimurium SL1344 (DT44) whose tRNA-Thr locus was unoccupied, exhibited an intermediate frequency of lysogeny (10^{-6}).

5.2.4 Evidence for transfer of mTmV to distinct *S. enterica* serovars in nature

The mTmV prophage was almost exclusively detected in epidemic *S.* 4,[5],12:i:- from animal and human samples within *S.* Typhimurium (Chapter 4), despite its ability to transfer to other *S*. Typhimurium and other serovars during culture *in vitro*. To investigate evidence for mTmV transfer in nature, the nucleotide sequence of the mTmV prophage from strain S04698-09 was aligned to sequences in the GenBank (accessed on 1st July 2019) using BLAST. Out of the 10,811 genome assemblies, four *S.* Derby from Canada, China and France, and a single sequence of *S*. Brandenburg and *S*. California from Canada and China, respectively, harboured sequences homologous to mTmV and *sopE* (Table 5.1). All prophages exhibited less than 0.09% sequence divergence, ranging from 99.91% (*S.* California CD-SL01 and the *S.* Derby CFSA231) to 99.98% (*S.* Derby strain CFSA231) nucleotide sequence identity. The organisms were isolated from pork, an unspecified food product, and chicken.

Table 5.1 Sequences of isolates positive for mTmV prophage in GenBank. The strain name, serovar, country source and accession number are reported together with the percentage of nucleotide identity (Id) of the prophages over whole alignment with mTmV prophage sequence and the nucleotide coordinates of the prophages in the genomes.

Strain	Serovar	Country	Source	Acc. n.	Id (%)	Coordinates
SA20113174	Brandenburg	Canada	-	CP029999.1	99.94%	3,532,397-3,571427
CD-SL01	California	China	Chicken Food	CP028900.1	99.98%	3,610,841-3,649,870
SA20035215	Derby	Canada	-	CP022494.1	99.97%	605,111-644,140
2014LSAL02547	Derby	France	Pork	CP029486.1	99.95%	4,041,042-4,080,071
Sa64	Derby	China	-	CP034250.1	99.91%	3,336,809-3,375,855
CFSA231	Derby*	China	Food	CP033350.2	99.98%	3,611,673-3,650,702

*serovar determined in silico using SeqSero v1.2 software (Zhang et al., 2015)

Sequence variation of mTmV in isolates of diverse *S*. *enterica* serovars mainly affected genes involved in virion structure and assembly. In particular, all mTmV prophages from the isolates in Table 5.1 harboured polymorphisms within two distinct phage tail sheath proteins (ORFs 00361 and 00363), the baseplate protein (ORF 00362) and in a hypothetical protein (ORF 00362) compared with mTmV sequence from strain S04698-09. Other polymorphic genes were present with reference to S04698-09 mTmV but were limited to a single or few mTmV prophages (Appendix 6).

To further study the dissemination of the mTmV phage from *S.* 4,[5],12:i:- to other *S. enterica* serovars, the whole genome sequences were determined for 34 isolates belonging to eight distinct serovars isolated from ten pig farms in Ireland (described in Chapter 3) (Table 5.2 and Appendix 7). The mTmV prophage was detected in four *S.* Derby isolates, of which three originated from farm C, where a mTmV-positive *S.* 4,[5],12:i:- was also isolated. The remaining *S*. Derby was isolated from farm D, where a single mTmV-negative *S.* 4,[5],12:i:- was isolated. Two *S.* Dublin isolates from farm H were negative for mTmV but positive for *sopE*, although the gene differed by 37 SNPs from *sopE* in *S.* 4,[5],12:i:- and was

analogous to the gene previously described in *S*. Dublin (Mirold et al., 2001b). The *sopE* gene was instead located within a lambdoid prophage, as described previously for this serovar (Mirold et al., 2001b). The unrelatedness between the *sopE* phages of *S*. Dublin and *S.* 4,[5],12:i:- was further supported by the ability of mTmV to lysogenise two *S.* Dublin isolates (paragraph 5.2.3), indicating the lack of superinfection mechanisms which affect related phages. In farm H, only a single *S.* 4,[5],12:i:- was isolated, which was negative for both *sopE* and mTmV.

Table 5.2 Number of isolates and farm of isolation of the sequenced *S. enterica* serovars from Irish pig farms.

Serovar	**Number of isolates**	**Farm**	**Number of mTmV-positive isolates**
Anatum	1	A	0
Derby	15	B, C, D, E, G, I, J	4 (Farms C and D)
Dublin	3	H, I	0
Infantis	2	E	0
London	6	I, J	0
Stanley	1	B	0
Tennessee	3	J	0
Typhimurium Copenhagen	3	G, I	0

S. Derby represents the third most common serovar circulating in pigs, which are the main animal reservoir *S*. 4,[5],12:i:-. To study the prevalence of mTmV in the *S*. Derby serovar, the short reads of 116 *S*. Derby were downloaded from Enterobase and investigated for the presence of the prophage genome using the SRST2 software (Inouye et al., 2014, Alikhan et al., 2018). The collection comprised sequences of isolates isolated in the UK from human and food samples between 2014 and 2017 (Appendix 8). Twenty-one sequences had mTmV encoding *sopE*, corresponding to 18% of the total isolates. Two isolates carried a *sopE* gene identical to that of *S.* 4,[5],12:i:- but no mTmV, whereas one isolate had a *sopE* allele with 36 SNPs with reference to *S.* 4,[5],12:i:- S04698-09. Therefore, the analysis suggested that the transfer of mTmV to *S*. Derby isolates did not occur sporadically as mTmV-positive *S*. Derby were detected in

multiple countries and in a relatively high proportion of *S*. Derby sequences collected within the UK.

5.3 Discussion

The acquisition of *sopE* in *S*. 4,[5],12:i:- was mediated by the mTmV prophage (Petrovska et al., 2016), related to phages belonging to the *Myoviridae* family circulating in *Shigella* and *Salmonella*. Similar to many prophages, the genomic organisation of mTmV comprised modules of genes encoding for the same function grouped together (Hatfull and Hendrix, 2011). Most of the genes of mTmV were shared with SfV, with the exception of those encoding for the serotyping-converting functions in SfV, the capsid genes and the *sopE* cassette. Likewise in other toxin-encoding bacteriophages, the *sopE* cargo gene was situated in the late gene region in mTmV (Hendrix et al., 2000).

The sequence homology between SfV and mTmV suggested that two of the three mechanisms of SfV conferring superinfection immunity were conserved in mTmV. One is mediated by the cI repressor, which can act *in trans* in analogous bacteriophages thus inhibiting the phage propagation (Allison et al., 2002). A second mechanism of immunity is mediated by a small RNA, encoded by orf37 in SfV, which acts as a transcription terminator (Roberts et al., 2007). A third immunity mechanism encoded by SfV consisting in modification of the O-antigen, used as a receptor by many bacteriophages, is absent in mTmV (Allison et al., 2002). mTmV and SfV shared the same integration site, namely the *thrW* locus, a conserved hotspot for insertion of mobile genetic elements (Bobay et al., 2013). Additionally, mTmV was related with the prophages ST64B and SJ46 integrated in the genome of S04698-09. The presence of homologous sequences may promote homologous recombination between prophages and give rise to novel, mosaic viruses (De Paepe et al., 2014) (see Chapter 6).

The mTmV prophage, which is an active phage capable of replication, associated with the *S*. 4,[5],12:i:- ST34 clonal complex. Although mTmV was not detected in any of *S*. Typhimurium isolates, it infected and integrated into the *thrW* locus in all the *S*. Typhimurium genotypes tested *in vitro* indicating that the bacteriophage could potentially spread to *S*. Typhimurium. However, the susceptibility to the mTmV bacteriophage varied with the genotype. Factors affecting the susceptibility to mTmV are not known. The occupation of the *thrW* locus by another prophage did not appear to alter the integration of mTmV nor its lysogeny rate. Isolates which already carried a prophage in the tRNA locus formed lysogens of the mTmV phage with an equal frequency to isolates whose locus

was empty. The rate at which mTmV lysogens were formed may be influenced by the lack or different abundance of surface molecules used as receptors by mTmV. Also, the different repertoire of prophages already resident in the recipient isolate can affect the susceptibility to the mTmV phage as temperate bacteriophages evolved superinfection exclusion mechanisms to prevent infection by related viruses (Hofer et al., 1995, Kliem and Dreiseikelmann, 1989, Bebeacua et al., 2013).

S. Typhimurium phage types U308, DT193 and DT104 were among the genotypes which formed lysogens with the lowest frequency. This may help explain why no mTmV-positive isolates from the above genotypes had ever been isolated despite circulating in pigs as *S*. 4,[5],12:i:-. Nonetheless, mTmV has never been reported in the pig-associated *S*. Typhimurium phage type U288 despite forming lysogens 100 times more frequently. The bird-associated phage types DT2, DT56 and DT99 were those which exhibited the greatest susceptibility to the mTmV bacteriophage. *S*. 4,[5],12:i:- was reported in wild birds (De Lucia et al., 2018); therefore, the acquisition of mTmV by the avian-associated types following contact with *S*. 4,[5],12:i:- is plausible.

Outside *S*. Typhimurium, mTmV was able to lysogenise pig isolates belonging to serovars Dublin and Derby with low frequency. Nevertheless, *S*. Derby harbouring mTmV integrated into the genome were isolated from Irish pig farms, in a case together with mTmV-positive *S*. 4,[5],12:i:-, providing evidence for the transfer of the bacteriophage between the two serovars. None of the other seven *Salmonella* serovars isolated from the same pig farms carried the prophage, indicating a preferential entrance of mTmV for *S*. Derby. *S*. Derby was the third most common serovar isolated from pigs in 2016 in the European Union, and it was among the top ten serovars responsible for human infection (ECDC, 2017). *S*. Derby isolates carrying mTmV were isolated in multiple European countries (Ireland, UK, France) and other continents (America, Asia). In addition, approximately 20% of the clinical *S*. Derby isolates from UK had mTmV integrated in the genome, suggesting a positive selection for the prophage. The mTmV prophage was also detected in two further serovars, namely *S*. Brandenburg and *S*. California. *S*. Brandenburg was responsible for human and animal cases, in particular sheep, in New Zealand between late 1990s-early 2000s (Clark et al., 2004). In the European Union, *S*. Brandenburg was among the 20 most common

serovars responsible for human cases in 2017 (ECDC, 2018). In a survey conducted in Italy, the serovar was commonly isolated from pigs (Bonardi et al., 2018). *S*. California was widely isolated in Spanish feed mills between late 1990s-early 2000s and represented one of the most common serovars isolated from pigs in Canada (Mainar-Jaime et al., 2008, Alvarez et al., 2003). However, MDR *S*. California recently emerged in poultry settings in China (Wang et al., 2017). The dissemination of mTmV is not, therefore, limited to serovars which circulate exclusively in pigs.

The mTmV prophages detected in S. Derby, Brandenburg and California carried non-synonymous polymorphisms in the genes encoding for the tail fibre proteins which may have contributed to expanding the host range of the virus. Myoviruses can bind the microbial receptor through the tail proteins before injection of the viral DNA into the host cell (Garcia-Doval and van Raaij, 2013). The acquisition of *sopE* by *S*. *enterica* serovars Derby, Brandenburg and California potentially increase the pathogenic potential and fitness of these serovars and may contribute to the emergence of epidemics as for *S*. Typhimurium where the gene had been detected in lineages which were/are epidemics (DT29 and DT204, *S*. 4,[5],12:i:- DT193/DT120) or precursors of epidemics (DT44 and DT49, precursors of DT29 and DT204, respectively) (Hopkins and Threlfall, 2004, Petrovska et al., 2016). Outside of *S*. Typhimurium, *sopE* was detected in *S*. Enteritidis phage types (PT1, PT4, PT6, P8, PT13 and PT14b) commonly isolated from livestock and responsible for human infections (APHA, 2018a, APHA, 2015), and in *S*. Heidelberg isolates associated with an outbreak (Hopkins and Threlfall, 2004, Hoffmann et al., 2014). However, *sopE* did not necessarily correlate with epidemic isolates.

Beside *sopE* carriage, the integration of the phage itself may be beneficial for the host cell for several reasons. The phage can block secondary infections of related phages through superinfection exclusion mechanisms (Obeng et al., 2016), it represents a source of genetic variation (it was shown to be capable of specialised transduction), and the released virions can act as biological weapons against competing cells (Bossi et al., 2003).

5.4 Conclusion

In conclusion, mTmV played a major role in the dissemination of the *sopE* virulence gene in *S*. 4,[5],12:i:- ST34. The mTmV prophage was virtually absent

in other *S*. Typhimurium despite its ability to infect a variety of types to a different extent, suggesting that its potential spread within this serovar cannot be ruled out. Nonetheless, mTmV disseminated to diverse *S*. *enterica* serovars, which potentially increased their pathogenicity and epidemiological success.

6 *sopE* copy number variation associated with mTmV variants and transfer of *sopE* cassette

6.1 Introduction

The *sopE* virulence gene is uncommon in *S*. Typhimurium but widespread in the current pandemic *S*. 4,[5],12:i:- ST34 clonal group, whose acquisition was predominantly mediated by the mTmV phage in UK human isolates (Chapter 4). However, a proportion of the *sopE*-positive *S*. 4,[5],12:i:- isolates (10%) did not harbour the mTmV sequence suggesting that *sopE* was associated with distinct prophages in some *S*. 4,[5],12:i:-.

Bacteriophages are known to exchange gene modules either via homologous or illegitimate recombination which generates novel, mosaic bacteriophages (Hatfull and Hendrix, 2011). Recently, a variant of mTmV, termed mTmV2, was reported in *S.* 4,[5],12:i:- ST34 isolates from Italy (Palma et al., 2018, Petrovska et al., 2016). Although these organisms were part of the *S*. 4,[5],12:i:- ST34 clonal group, the mTmV2 prophage was only present in a subclade of *S*. 4,[5],12:i:- circulating in Italy and rare or absent in other countries (Palma et al., 2018). The mTmV2 prophage, which was inserted at the *thrW* locus as mTmV, was 42.9 kb long and encoded 63 predicted gene*s*, only half of which were shared with mTmV (Palma et al., 2018). The region of mTmV2 which differed from mTmV comprised genes involved in phage gene regulation, integration-recombination and cell division, prophage repression, cellular lysis and serum resistance (Palma et al., 2018).

The emergence of phage variants can promote the dissemination of *sopE* to novel host cells by circumventing superinfection exclusion mechanisms of the resident prophages. Additionally, the dissemination may be promoted by the transfer exclusively of the *sopE* cassette (comprising the *sopE* coding sequence and the flanking regions) between unrelated phages (Mirold et al., 2001b).

The potential for recombination between prophages and the transfer of *sopE* between unrelated bacteriophages have further implications for the copy number variation of the gene (Mirold et al., 2001b). The polylysogeny of distinct phages carrying *sopE* can occur with subsequent increase of copy number. For instance, the SopEΦ was able to integrate in S. Dublin, Hadar and Gallinarum which already carried *sopE* but within a cryptic lambdoid prophage (Mirold et al., 2001b). Two copies of *sopE* were detected in isolates of *Salmonella* Heidelberg

responsible for a large outbreak in the U.S. harboured by distinct prophages (Hoffmann et al., 2014). Although the impact of multiple copies of *sopE* remains unknown, I tested whether *S.* 4,[5],12:i:- harbouring multiple copies of the gene were present in a collection of isolates previously published (Petrovska et al., 2016). As a group of isolates with *sopE* copy number variation was identified, the genetic context of the distinct copies of the virulence gene was investigated in whole genome long-read sequences (PacBio).

6.2 Results

6.2.1 A cluster of *S*. 4;[5],12:i:- isolates exhibited copy number variation

To test whether any isolates from a collection of 22 *S*. 4;[5],12:i:- epidemic in Italy and UK from a previously published study (Petrovska et al., 2016) carried multiple copies of *sopE*, the gene copy number was estimated by calculating the ratio between the number of reads mapping to the *sopE* coding sequence and those mapping to the *sopE2* gene, which has the same nucleotide length as *sopE* and is always present in a single copy.

The whole genome sequence from most isolates had similar mapped read depth to *sopE* and *sopE2*, with the ratio ranging from 0.87 and 1.24. However, a cluster of five closely related isolates (S01569-10, 4824-10, 4797-10, S3659-10 and H1041406001) exhibited a read ratio above 1, ranging from 1.49 to 3.01 indicating they may encode multiple copies of *sopE* (Table 6.1 and Figure 6.9). In order to investigate the bacteriophages responsible for *sopE* copy number variation, whole-genome, long-read sequences of isolate S01569-10, identified as a candidate for carrying multiple *sopE* copies, were generated using PacBio technology. The organism was isolated from a pig sample in 2010 in the UK.

Table 6.1 Estimated *sopE* copy number in *S*. 4;[5],12:i:- isolates from the UK and Italy (Petrovska et al., 2016).

Strain name	*sopE* n. reads	*sopE2* n. reads	Ratio reads *sopE*/*sopE2*	Subclade	Country
S04797-08	1760	1615	1.09	A	UK
1334-1997	1546	1512	1.02	A	UK
S01364-10	2776	2754	1.01	A	UK
S03113-10	1900	2037	0.93	A	UK
S04698-09	2107	2122	0.99	A	UK
S05893-09	2581	2609	0.99	A	UK
S05894-09	1882	1903	0.99	A	UK
H105280433	3732	3873	0.96	A	UK
S0433209	251	261	0.96	B	UK
S01569-10	3468	1613	2.15	B	UK
4824-10	4993	2867	1.74	B	UK
4797-10	6716	2228	3.01	B	UK
S3659-10	1989	1160	1.71	B	UK

H1041406001	2280	1535	1.49	B	UK
H104680513	1935	1871	1.03	B	UK
H105000301	1723	1391	1.24	B	UK
2200/2	776	847	0.92	C	Italy
2448/2	1653	1798	0.92	C	Italy
1948/2	2226	2217	1.00	C	Italy
1693/1	546	543	1.01	C	Italy
1365/1	2551	2626	0.97	C	Italy
629/2	2125	2439	0.87	C	Italy

6.2.2 Variants of mTmV associated with enhanced copy number of *sopE*

In order to investigate the presence of multiple copies of *sopE*, complete and closed whole-genome sequence assemblies of isolate S01569-10 were generated. The chromosome was 4,973,996 bp and an additional extra-chromosomal 80 kb contiguous assembled sequence was identified. Three copies of *sopE* were identified after inspection of the sequences with the genome browser Artemis (Rutherford et al., 2000), of which two were on the chromosome contiguous sequence, and the third on the 80 kb episome. The presence of three copies of *sopE* was confirmed by PCR (data not shown). To the best of my knowledge, this is the first description of three copies of *sopE* gene in a single isolate. Both chromosomal copies of *sopE* were located on a distinct prophage region homologous to mTmV suggesting they originated from recombination and duplication of mTmV (Figure 6.1)

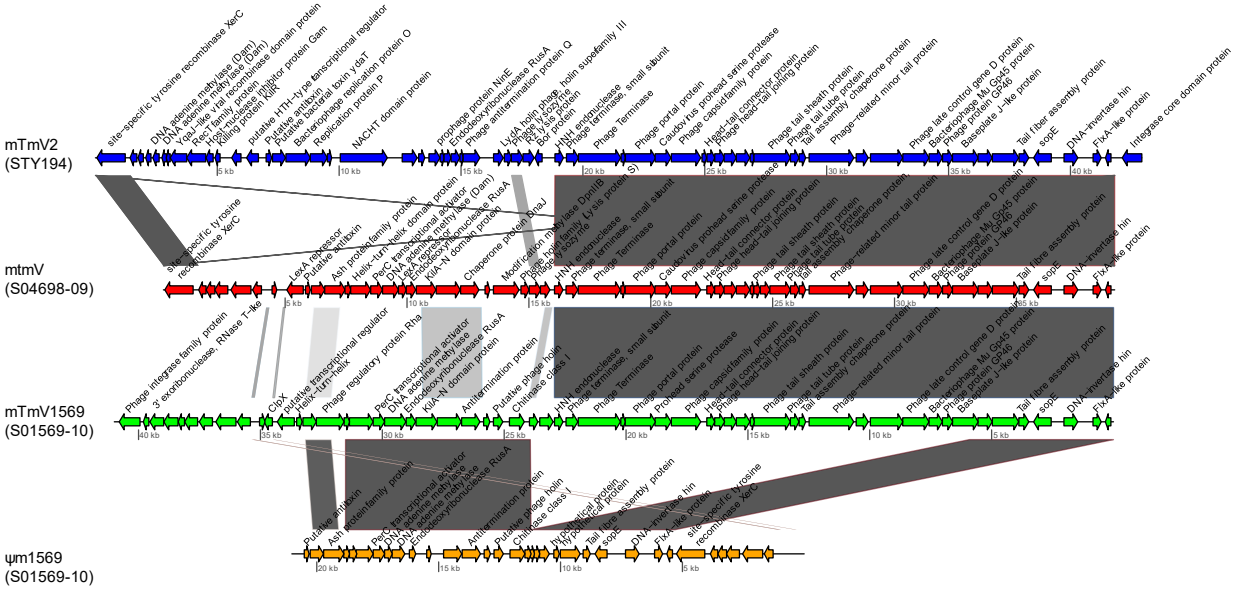

Figure 6.1 Genome comparison between *sopE* bacteriophages in various *S.* 4,[5],12:i:- isolates. The sequence similarity between the genomes of mTmV2 (strain STY194), mTmV (strain S04698-09), mTmV_1569 (strain S01569-10) and Ψm1569 (strain S01569-10) is shown. Different shades of grey display different degrees of nucleotide identity: dark grey: 95-100% identity, mid-grey: 85-94% identity; light-grey: 80-93% identity.

Neither chromosomal copy of *sopE* was encoded on a complete mTmV or mTmV2 prophage element but both contained sections of the mTmV prophage, suggesting that they had emerged as a result of recombination from an ancestral mTmV. One of the chromosomal copies of *sop*E was located on a 40 kb prophage (designated mTmV1569) which shared 99.74% nucleotide identity over 65% of the mTmV genome (Figure 6.1). A 22.9 kb region spanning from the HNH endonuclease to the last ORF of mTmV (PROKKA_0056) was conserved between mTmV, mTmV1569 and mTmV2, indicating the presence of a conserved recombination breakpoint (Figure 6.1). Other shorter and less conserved sequences of mTmV1569 that were homologous to mTmV comprised genes encoding for an Ash protein family (cI repressor) and a flanking hypothetical protein, and partial KilA-domain protein, an antitermination protein and an additional hypothetical protein. The non-homologous regions in mTmV1569 contained genes involved in phage excision/integration, immunity and regulation, replication and lysis. A BLAST analysis of the 18 kb sequence of mTmV1569 which spanned from the integrase to the HNH endonuclease (excluded) showed that it was almost identical (99% nucleotide identity; 100%

query cover) to *Salmonella* phage 118970_sal3 (accession number KU927493.2) suggesting that mTmV1569 may have originated following the recombination between mTmV and sal3 prophages.

In contrast to mTmV and mTmV2, mTmV1569 was not integrated into the *thrW* locus, but instead in the serine tRNA locus*.* In a collection of representative *S*. Typhimurium, the tRNA-Ser locus was occupied by the ST64B prophage. I hypothesized that mTmV1569 resulted from the recombination of an mTmV-like prophage integrated into the *thrW* locus with the ST64B prophage inserted into the tRNA-Ser locus. The *thrW* locus of S01569-10 was, in fact, occupied by a prophage which shared partial homology with mTmV from S04698-09 in the 5' portion of the genome (Figure 6.2). The conserved regions between the prophages in the *thrW* locus and in the tRNA-Ser may have been used as a substrate for the homologous recombination between the prophages with the consequent generation of the mTmV1569 prophage in S01569-10 (Figure 6.2).

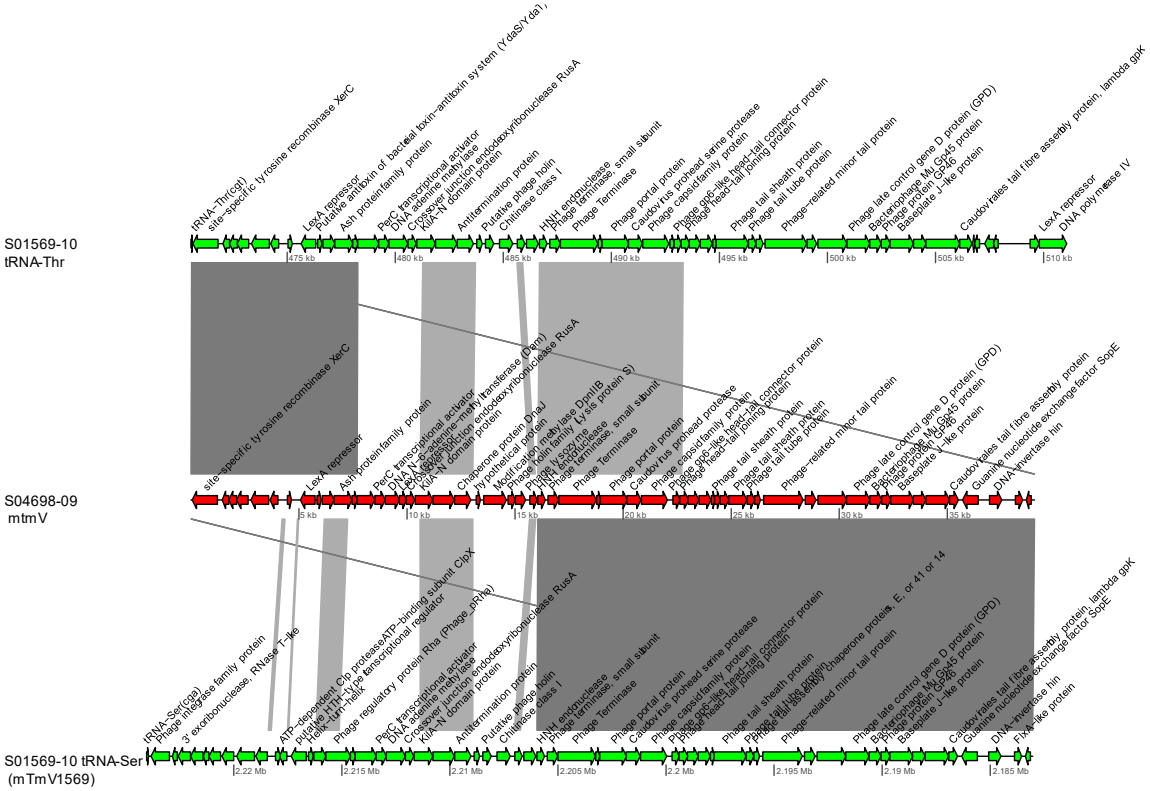

Figure 6.2 Conserved regions between the mTmV prophage and the prophages integrated in the tRNA-Thr and tRNA-Ser sites of S01569-10. Dark grey: 100-90% nucleotide identity; light grey: 89-80% nucleotide identity.

This hypothesis was supported by the alignment of a larger chromosomal region between S04698-09 and S01569-10 which revealed a large inversion (about 1.7

Mb) between the mTmV and the ST64B prophages, leading to the generation of a mosaic ST64-mTmV bacteriophage (Figure 6.3).

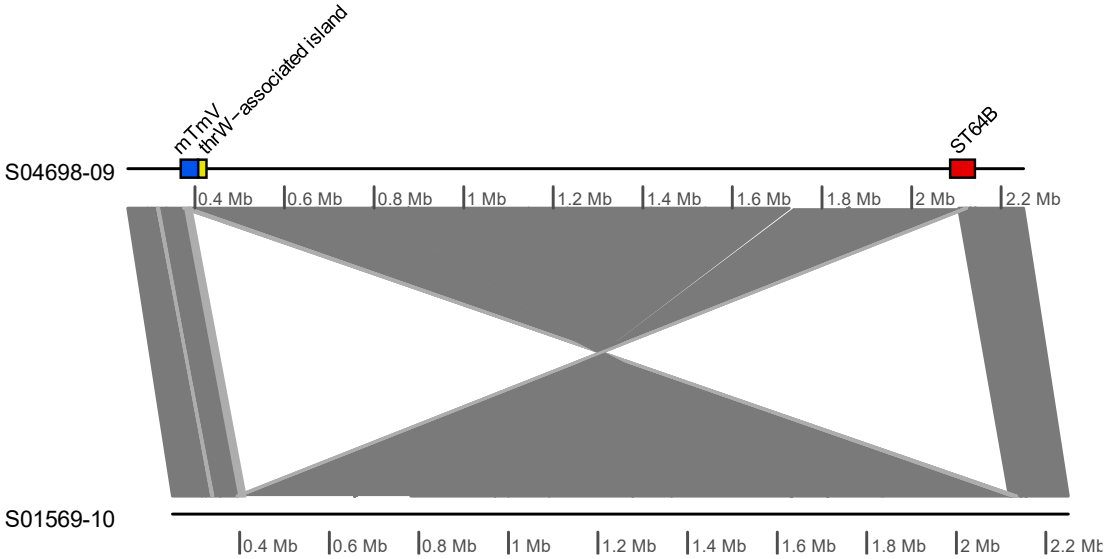

Figure 6.3 Genomic inversion in *S.* 4,[5],12:i:- S01569-10 compared with S04698-09 chromosome. The sequences conserved between a sub-genomic region of *S.* 4,[5],12:i:- S04698-09 and S01569-10 are shown. The regions flanking the inversion are located within the mTmV and ST64B prophages. Dark grey: >90% of nucleotide identity; light grey: 70-89 % of sequence identity.

A second chromosomal copy of *sopE* was located within a region of 21 kb of probably incomplete prophage genome, which was designated Ψm1569 (pseudo-phage mTmV_1569), inserted within the upstream housekeeping gene *thrA*, encoding for a bifunctional aspartokinase I/homoserine dehydrogenase 1 (Figure 6.1). 76% of Ψm1569 prophage shared significant nucleotide identity with mTmV (99.78% nucleotide identity) and appeared to be a mosaic of three, non-contiguous regions of mTmV1569 (Figure 6.1). *sopE* was within a 5 kb fragment homologous to mTmV, in the terminal portion of the genome. The Ψm1569 prophage lacked the genes involved in DNA packaging and virion morphogenesis, therefore it may not be an active phage. Nonetheless, it may be able to use structural genes from other prophage genomes on the chromosome *in trans* to propagate (Briani et al., 2001).

6.2.3 Identification of pS01569, a novel extrachromosomal bacteriophage carrying *sopE*

A third copy of *sopE* in S01569-10 was present on an extra-chromosomal 81.1 kb contig which did not join to the chromosomal assembly. The contig, which the Circlator software (Hunt et al., 2015) predicted to be circular, encoded for 109 ORFs.

To directly test the presence of episomal DNA, separate from the chromosome, DNA was extracted from a culture of strain S01569-10 with the method developed by Kado and Liu (Kado & Liu 1981), followed by size separation by gel electrophoresis. The pSLT plasmid, whose size was 93.3 kb, was used as a control. The gel revealed the presence of DNA that migrated at a faster rate than pSLT, consistent with the presence in S01569-10 of an episome of approximately 80 kb which was termed pS01569 (Figure 6.4).

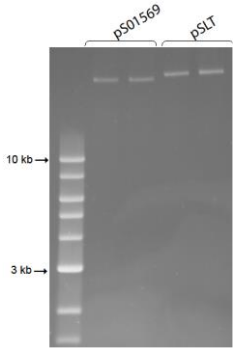

Figure 6.4 Agarose gel electrophoresis of plasmid DNA extracted from *S.* 4,[5],12:i:- S01569-10 and *S.* Typhimurium L01157-10, which carried the pSLT. The pSLT plasmid is approximately 90 kb long.

To further test the hypothesis that pS01569 is circular, amplification across the junctions of the contiguous assembled sequence was performed by PCR using primers facing outward from the assembled sequence. Oligonucleotides binding within an internal region were used as a control (Figure 6.5). The pSLT was included as a negative control. The amplification of the expected fragments indicated that the pS01569 existed as a circular, extra-chromosomal element (Figure 6.6).

Figure 6.5 Linear representation of the pS01569 showing the sites to which oligonucleotides anneal. Green and blue half arrows: outward primers; magenta half arrows: internal primers. The yellow filled arrow corresponds to *sopE*.

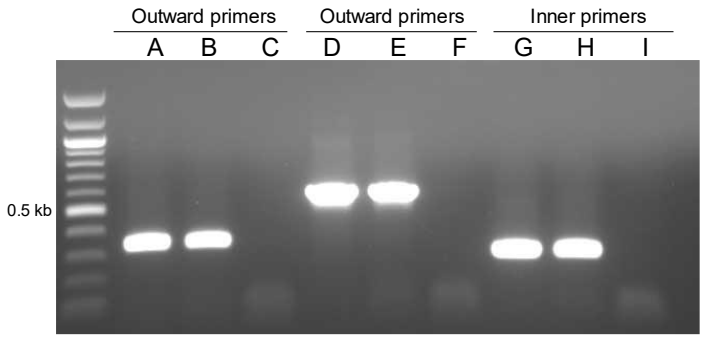

Figure 6.6 Amplicons generated with the oligonucleotides shown in Figure 6.5 suggested that the pS01569 is circular. Agarose gel with the PCR products obtained using the genomic extraction (A, D and G), the plasmid extraction (B,E and H) and the pSLT (C, F, I) with the primers shown in Figure 6.5. The amplicons were of the expected lengths: 331 bp, 583 bp and 295 bp, respectively.

pS01569 was found to be a phage closely related to genetic elements described in *Shigella dysenteriae*. The alignment of pS01569 sequence in GenBank database with BLAST revealed that it shared 89% of the genome with a significant nucleotide identity (99%) with plasmids described in four *S. dysenteriae* (accession number CP026808.1, CP026821.1, CP026816.1 and CP026818.1). As *sopE* was generally associated with bacteriophages, I tested the hypothesis that pS01569 was a bacteriophage with the PHASTER software (Arndt et al., 2016). The software predicted that pS01569 was an intact bacteriophage. The closest subject sequence to pS01569 reported as a bacteriophage in GenBank was the *E*. *coli* bacteriophage D6 (accession number MF356679.1) with which pS01569 shared 90.5% nucleotide sequence identity over 70% of the genome) (Figure 6.7). The D6 phage is a myovirus distantly related to the P1 bacteriophage, which resides in the host cell as a circular plasmid (Gilcrease and Casjens, 2018).

The alignment by BLAST and genome comparison between the pS01569 and D6 genomes showed the presence of conserved genes in pS01569 required for bacteriophage replication (Figure 6.7). The sequences included regulators of viral cycle, genes involved in DNA packaging and virion morphogenesis, lysis, restriction defence, DNA modification, DNA metabolism and recombination, and plasmid stability. In pS01569 the *sopE* cassette replaced the tail fibre proteins of the D6 bacteriophage (Figure 6.7).

As in the D6 phage and the plasmid from *S. dysenteriae* 204/96, the pS01569 carried a region encoding for multiple tRNA loci, namely tRNA-Thr, tRNA-Met and tRNA-Arg (Figure 6.7). The presence of tRNA genes in bacteriophages was hypothesized to be selected either by the codon or by the amino acid usage of the phage (Delesalle et al., 2016). Additionally, tRNA have been hypothesized to increase the phage host range as the phage would be able to propagate in a host with different codon/amino acid usage. An alignment of the tRNA loci of pS01569 in NCBI and tRNADB-CE databases (Abe et al., 2009) showed that the best matches were loci on mobile genetic elements predominantly from *E. coli* and *S. dysenteriae*, but also from *Salmonella* isolates (data not shown), which suggested that the pS01569 and related genetic elements are promiscuous elements circulating in multiple *Enterobacteriaceae* species.

Additionally, two DDE domain proteins encoding for IS6 family transposases according to their alignment in ISfinder (Siguier et al., 2006), were present in pS01569, but not in the D6 phage nor in the *S. dysenteriae* plasmid. The genes were flanking a 35 kb DNA inversion in pS01569 compared with the *S. dysenteriae* plasmid sequence, suggesting their possible role in causing DNA rearrangement (Figure 6.7).

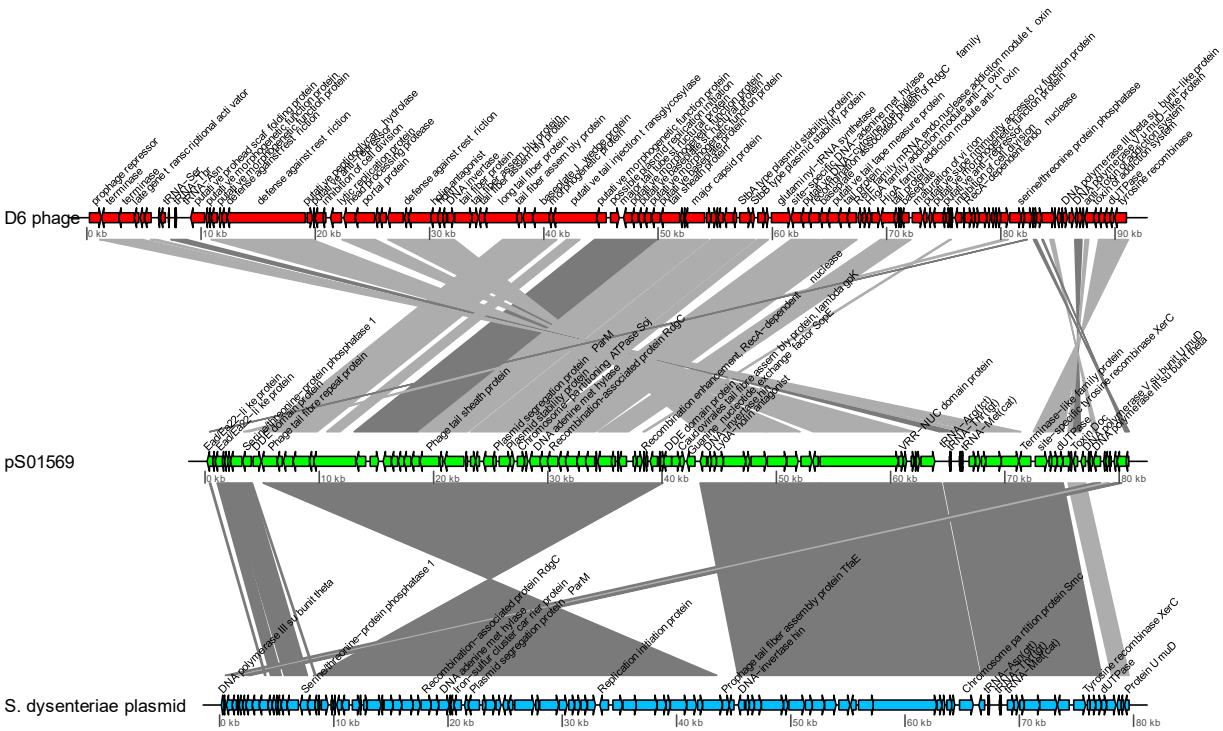

Figure 6.7 Genome comparison between the *E. coli* phage D6, pS01569 and the *S. dysenteriae* plasmid from isolate 204/96. Dark grey: 90-100% nucleotide identity; light grey: 73-89% nucleotide identity.

A genetic element highly related to the pS01569 phage (99% nucleotide identity in 93% of the pS01569 genome) was identified in the sequence of *S*. Typhimurium isolate DA34867 (accession number CP029568.1). The organism, which belonged to sequence type ST34 and was predicted to be *S*. 4,[5],12:i:- using SeqSero-1.2 (Zhang et al., 2015), was isolated from a human urine sample in Sweden and carried the viral genome integrated into a composite transposon described in the *fljAB* locus (Petrovska et al., 2016) (Figure 6.8). In DA34867, the prophage lacked *sopE*, but carried several gene sequences absent in pS01569 encoding two alternative tail fibre genes, a hypothetical protein and a tellurite resistance protein (the whole gene sequence shared 99% nucleotide identity with the *terB* gene in bacteriophage P7, accession number AF503408.1). Tellurium resistance was associated with colicin and bacteriophage resistance in H2 incompatibility plasmids (Taylor and Summers, 1979).

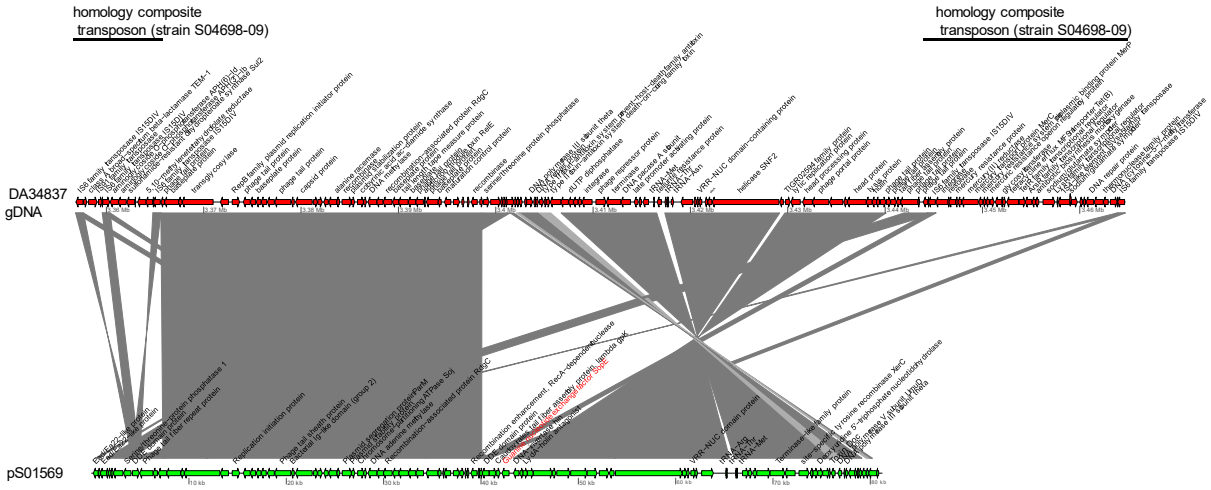

Figure 6.8 Sequence alignment of the pS01569 phage with a related genetic element inserted in the chromosome of *S*. 4,[5],12;i:- DA34837. The sequences of *S*. 4,[5],1:i:- DA34837 homologous with the composite transposon described in S04698-09 are indicated (Petrovska et al., 2016). Light-grey: 70-89% sequence identity; Dark-grey: 90-100% sequence identity.

6.2.4 Distribution of mTmV and its variants and pS01569 in *S*. 4,[5],12:i:- ST34

To test how common the prophage variants of mTmV (mTmV2, mTmV1569 and Ψm1569) and pS01569 were in *S*. 4,[5],12:i:-, their distribution was investigated in a collection of whole genome sequences of *S*. 4,[5],12:i:- from the UK and Italy using the SRST2 software (Inouye et al., 2014).

A phylogenetic signature in the distribution of the genetic elements was observed. The mTmV prophage was predominantly detected in isolates within subclade A, with the exception of a single isolate from subclade B (isolate H104680513) (Figure 6.9). Although all the mTmV sequences detected were marked as polymorphic compared to the reference sequence from isolate S04698-98, their distribution was consistent with the pattern of *sopE* presence indicating that mTmV harboured *sopE*. The sequence identity for mTmV was likely confounded by the presence of prophages homologous to mTmV in their genomes (i.e. ST64B and SJ46, Chapter 5). The mTmV2 variant was detected exclusively in Italian isolates from subclade C (Figure 6.9), consistent with previous reports (Palma et al., 2018). Sequences homologous to mTmV1569 were detected in all the isolates positive for *sopE* across the clade, whereas the Ψm1569 sequence was detected in all *sopE* positive isolates in subclades A and B (Figure 6.9). As for mTmV, their detection may have been confounded by the presence of homologous prophages in the genomes. However, the degree of homology in reference with the sequences of mTmV1569 and Ψm1569 from strain S01569-10 varied among the isolates. mTmV1569 elements identical to the reference were identified in a cluster of 5 closely related isolates in subclade B which comprised strain S01569-10, suggesting, therefore, that the prophage was vertically inherited with clonal expansion of the isolates. Two additional isolates in subclade B (strain H105000301 and S0433209) which encoded *sopE*, but were negative for both mTmV and mTmV2, carried a slightly different variant of mTmV1569 with 2 and 10 SNPs, respectively. The mTmV1569 bacteriophage seemed, therefore, able to move between distinct genotypes within the same subclade. The mTmV1569 sequences identified in the remaining isolates were characterised by a variable number of SNPs which ranged from 21 to 261, and 29 indels, therefore, were not likely corresponding to mTmV1569.

Similarly, Ψm1569 sequences identical to the reference were identified exclusively in the cluster of isolates where S01569-10 was. The remaining isolates had 15-38 SNPs in reference to Ψm1569 sequence. However, as the Ψm1569 prophage region was a mosaic of the mTmV1569 prophage, they were likely false positives. Isolates with polymorphic Ψm1569 harboured a single copy of *sopE* either by mTmV or mTmV1569, which further supported the hypothesis of a false positive result.

Finally, pS01569 was detected only in a cluster of five isolates comprising S01569-10, suggesting that the bacteriophage was likely acquired once followed by clonal expansion.

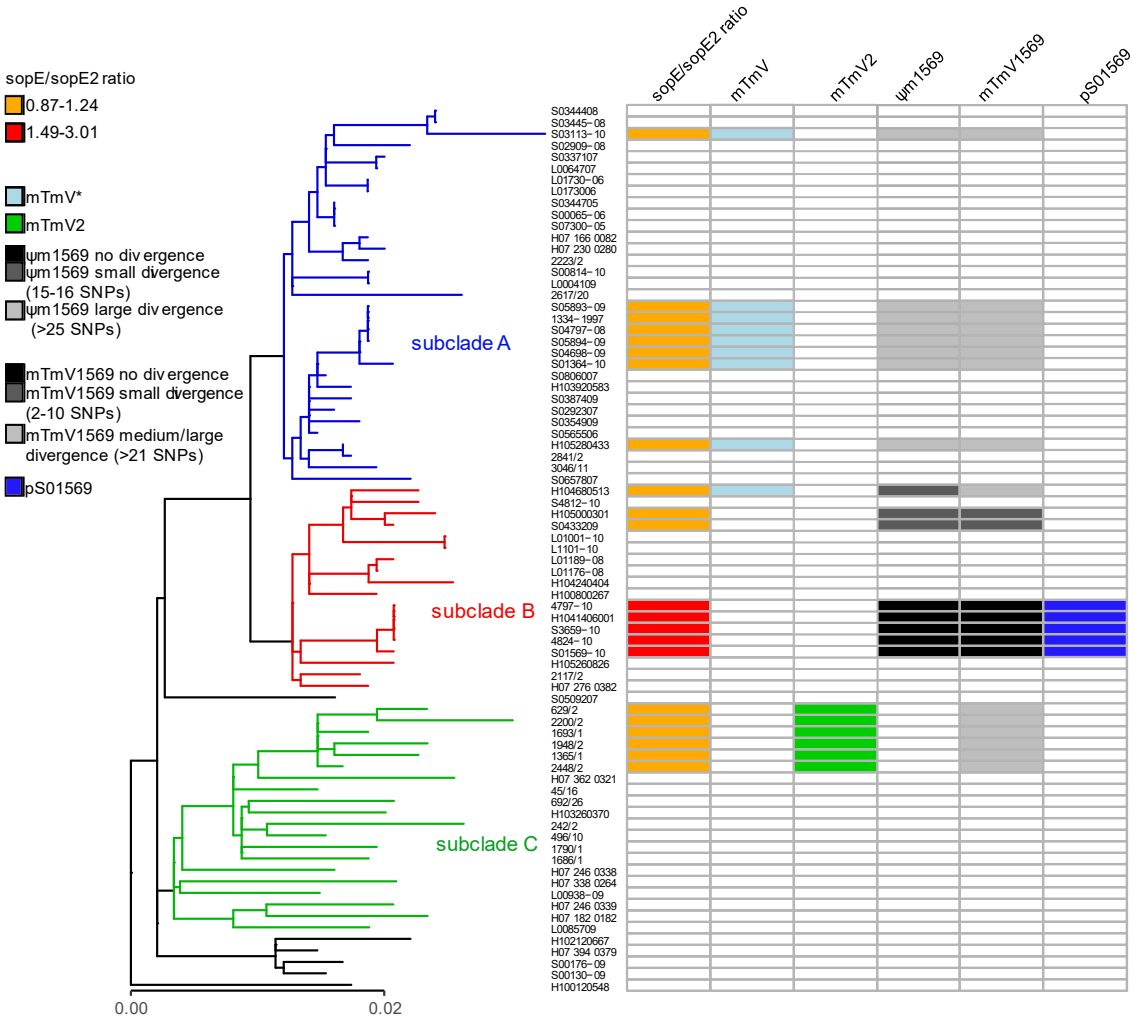

Figure 6.9 Variation of *sopE/sopE2* sequence read ratio and distribution of mTmV, mTmV2, Ψm1569, mTmV1569 and pS01569 elements in epidemic *S.* 4;[5],12:i:- from the UK and Italy (Petrovska et al., 2016). The maximum-likelihood phylogenetic tree, based on 1,558 SNPs identified with reference to *S.* Typhimurium SL1344, was generated with RAxML v8.2.12 using GTR+Γ substitution model (log likelihood score: -9603.9). The presence of mTmV polymorphic sequences is marked with an asterisk (*). The colour scale for Ψm1569 and mTmV1569 displays the different degrees of sequence variation detected in comparison with the reference sequences from isolate S01569-10.

To further test how common was the acquisition of the novel episomal phage pS01569 which may result in enhanced *sopE* dose, its distribution was investigated in a collection of *S.* 4,[5],12:i:- and *S.* Typhimurium from human

salmonellosis in the UK. Only 6 *S.* 4,[5],12:i:- out of 1,697 isolates were positive for pS01569 which were sporadically distributed consistent with multiple acquisitions (Figure 6.9). Nevertheless, two of the isolates carried a phage variant without *sopE*, whereas the remaining four *S.* 4,[5],12:i:- were positive for *sopE*, but also mTmV. The estimate of *sopE* copy number suggested that only two isolates harboured two copies of *sopE* indicating that pS01569 with *sopE* was present in only two isolates (Table 6.2).

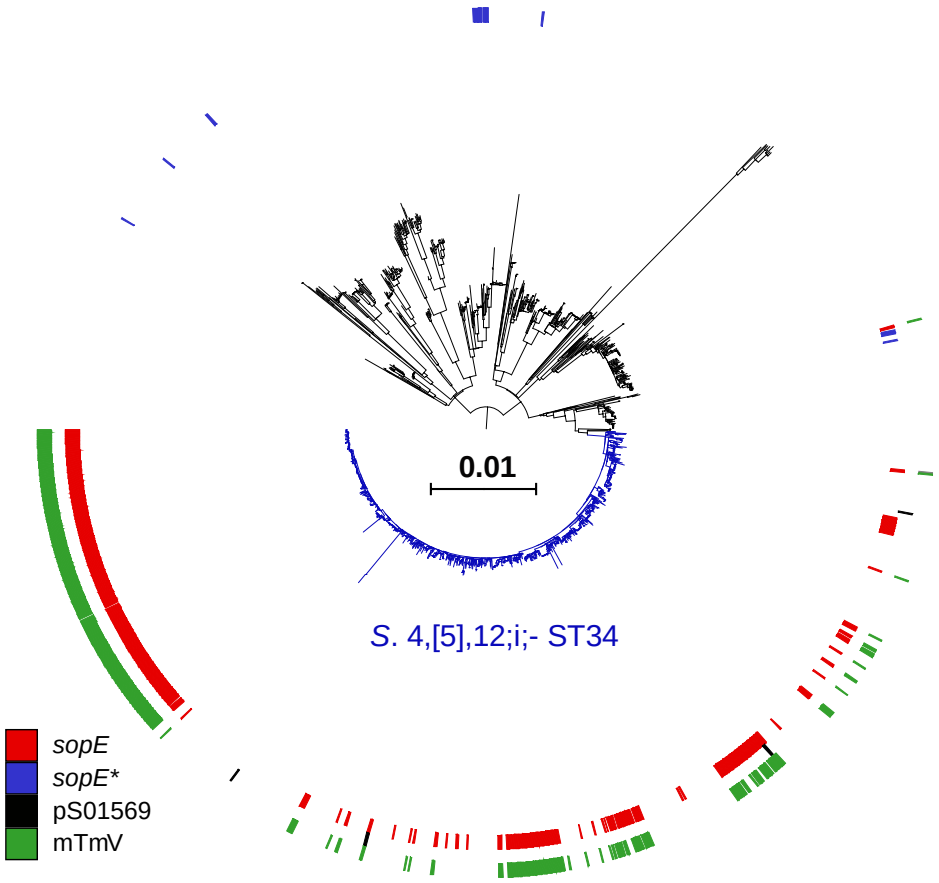

Figure 6.10 Phylogenetic distribution of *sopE*, pS01569 and mTmV in *S.* Typhimurium and *S.* 4,[5],12;i;- isolates from UK human infections. ML phylogenetic tree of *S.* Typhimurium and *S.* 4,[5],12;i;- isolates from human salmonellosis in the UK between 2012 and 2016. The presence of *sopE*, pS01569 and mTmV is shown from the innermost to the outermost circle. *sopE**: presence of polymorphisms in comparison with the reference sequence from S04698-09.

Table 6.2 Comparison of read depth of *sopE* and *sopE2* in *S.* 4,[5],12:i:- isolates positive for both pS01569 and mTmV.

Isolate	*sopE* depth (n. reads)	*sopE2* depth (n. reads)	ratio
H150740519	1709	797	2.14
H151280513	1721	822	2.09
H151420588	735	719	1.02
H151540888	830	843	0.98

6.2.5 Additional prophages harboured the *sopE* allele associated with S. 4,[5],12:i:-

To test whether other, distinct bacteriophages were responsible for the mobilisation of the *sopE* gene in *S.* 4,[5],12:i:-, the sequence of the *sopE* allele specifically associated with *S.* 4,[5],12:i:- (Chapter 4) was aligned in GenBank database using BLAST (date of accession 1st July 2019). Two isolates were positive for the *sopE* allele, but not for mTmV nor its variants, nor for the pS01569 bacteriophage, indicating that distinct genetic elements were responsible for the acquisition of the *sopE* gene in these isolates.

One sequence belonged to *S.* 4,12:i:- isolate 17437230 (accession number CP039593.1) isolated from a human sample in Canada in 2011, in which the *sopE* gene was carried by a bacteriophage which shared a 14.6 kb region of homology with mTmV (Figure 6.11). The second sequence was *S*. Miami ATCC BAA-1586 (accession number CP023468.1) in which *sopE* was located on a Fels-2 like prophage, termed phage ΦBAA which shared only a 2.2 kb fragment with mTmV comprising the *sopE* gene and its flanking sequences (Figure 6.11).

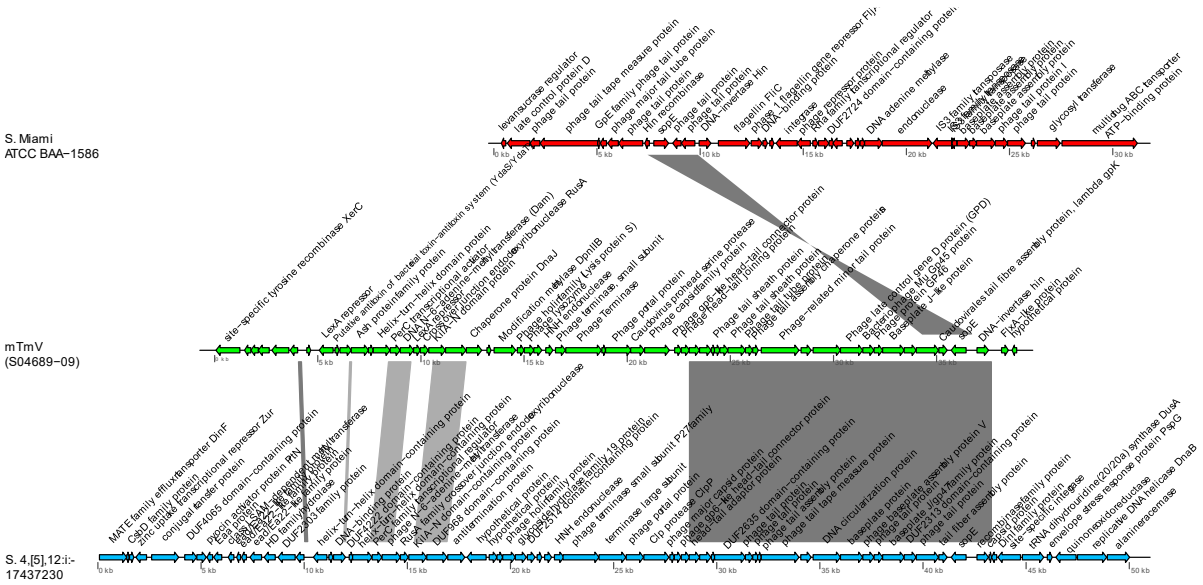

Figure 6.11 Sequence identity between mTmV and the *sopE* prophages in *S.* 4,12:i:- 17437230 and *S.* Miami ATCC BAA-1586. Light-grey: 70-89% sequence identity; dark-grey: 90-100% sequence identity.

6.2.6 The *sopE* allele associated with *S.* 4,[5],12:i:- was mobilised by phylogenetically distinct bacteriophages

The transfer of the *sopE* cassette between unrelated phage families has been previously described (Mirold et al., 2001b). In order to test whether the *sopE* allele circulating in *S*. 4,[5],12:i:- was harboured by unrelated bacteriophages, suggesting a possible transfer of the *sopE* cassette, the phylogenetic relationship of the viruses carrying the *sopE* allele were investigated. To this end a phylogeny based on whole genome sequences of the selected bacteriophages (Table 6.3) was generated using the web-based service VICTOR, which used a Genome BLAST Distance Phylogeny (GDBP) approach to cluster the isolates and build the tree (Meier-Kolthoff and Goker, 2017). In addition, the software used OPTSIL (Goker, Garcia-Blazquez, Voglmayr, Telleria, & Martin, 2009) to perform a taxonomic classification of the viral sequences aimed at clustering of the bacteriophages at the species, genus and family levels.

The prophages/phages included in the analysis comprised mTmV from *S*. 4,[5],12:i:- S04698-09 and its variants (mTmV2, mTmV1569 and Ψm1569), the episome pS01569 and the prophage sequences described in *S*. 4,[5],12:i:- strain 17437230 and *S*. Miami ATCC BAA-1586 (Table 6.3). In addition, the mTmV sequences detected in diverse *S*. *enterica* serovars (Chapter 5) were added in

order to probe their relatedness with mTmV. Prophages carrying *sopE* previously described in the literature were included to test their relatedness with phages harbouring the *sopE* allele associated with *S*. 4,[5],12:i:-. Representative genomes for different bacteriophage families within the Caudovirales order were included to provide context (Table 6.3).

Table 6.3 Phage and prophage genomes included in the phylogenetic analysis. The name of the virus as well as the host they came from and the accession number are reported.

Phage/ prophage	**Name**	**Host**	**Acc. number**
prophage	mTmV	*S.* 4,[5],12:i:- strain S04698-09	NZ_LN999997.1
prophage	mTmV_Brandenburg	*S.* Brandenburg strain SA20113174	CP029999.1
prophage	mTmV_Derby_ SA20035215	*S.* Derby strain SA20035215	CP022494.1
prophage	mTmV_Derby_2014LSAL02547	*S.* Derby strain 2014LSAL02547	CP029486.1
prophage	mTmV_Derby_ Sa64	*S.* Derby strain Sa64	CP034250.1
prophage	mTmV_Derby_ CFSA231	*S.* Derby strain CFSA231	CP033350.2
prophage	mTmV_California	*S.* California strain CD-SL01	CP028900.1
prophage	mTmV1569	*S.* 4,[5],12:i:- strain S01569-10	
prophage	Ψm1569	*S.* 4,[5],12:i:- strain S01569-10	
prophage	φBAA	*S.* Miami strain ATCC BAA-1586	CP023468.1
prophage	sopE_phage_ 17437230	*S.* 4,[5],12:i:- strain 17437230	CP039593.1
prophage	ST104	*S.* Typhimurium phage type DT 104	AB102868.1
prophage	Fels-1	*S.* Typhimurium	NC_010391.1
prophage	Gifsy-1	*S.* Typhimurium strain LT2	NC_010392.1
prophage	Gifsy-2	*S.* Typhimurium strain LT2	NC_010393.1

prophage	Fels-2	*S.* Typhimurium strain LT2	NC_010463.1
prophage	SopEφ	*S.* Typhimurium strain SL1344	NC_016810.1
prophage	sopE_phi_Enteritidis	*S.* Enteritidis strain ATCC BAA-708	CP025554.1
prophage	mTmV2	*S.* 4,[5],12:i:- strain STY194	PRJEB23875
phage	pS01569-10	*S.* 4,[5],12:i:- strain S01569-10	
phage	P22	*S.* Typhimurium LT2	AF217253.1
phage	ε34	*S.* Anatum	EU570103
phage	ST64T	*S.* Typhimurium phage type DT64	AY052766.1
phage	SPN1S	*S.* Typhimurium	JN391180.1
phage	ES18	*S.* Typhimurium	NC_006949.1
phage	ST64B	*S.* Typhimurium phage type ST64B	NC_004313
phage	P2	*E. coli*	NC_001895.1
phage	PsP3	*S.* Potsdam	NC_005340.1
phage	SfV	*Shigella flexneri*	NC_003444
phage	SfII	*Shigella flexneri*	KC736978.1
phage	P27	*E. coli*	NC_003356.1
phage	Mu	*E. coli*	AF083977.1
phage	SJ46	*S.* Indiana J46	NC_031129.1
phage	D6	*E. coli*	MF356679.1
phage	SfI	*Shigella flexneri*	JX509734.1
prophage	Gifsy-1	*S.* Heidelberg strain SA20025921	CP030214.1
prophage	*sopE* Gifsy-2	*S.* Heidelberg strain 41578	NC_021810.1
prophage	*sopE* P2-like	*S.* Heidelberg strain 41578	NC_021810.1

The allelic form of *sopE* associated with *S*. 4,[5],12:i:- was recently transferred between three phylogenetically distinct clusters of phage. The phylogenetic tree was constituted by five distinct clades belonging to distinct families of the Caudovirales order, each associated with a representative bacteriophage (Figure 6.12). The taxonomic classification performed by OPTSIL agreed at the genus level with the five major clades identified (Figure 6.12). The clustering of

reference sequences in accordance with the previous literature supported the topology of the tree generated. For example, the S*iphoviridae Salmonella* phage Gifsy-1, Gifsy-2 and Fels-1 grouped in a single clade (Switt et al., 2015). The *Podoviridae* P22 phage was related with ST64T and ST104 and epsilon34 phages (Switt et al., 2015). The myoviruses PsP3 and Fels-2 were related to the P2 bacteriophage (Kropinski et al., 2007) and P27 clustered with ST64B (Kropinski et al., 2007). The phage carrying *sopE* previously described were either P2-like bacteriophages, such as the SopEΦ from *S*. Typhimurium SL1344 (Pelludat et al., 2003) and the prophage from *S*. Heidelberg 41578 (Hoffmann et al., 2014), or were lambdoid bacteriophages, such as the prophages integrated in *S*. Enteritidis ATCC BAA-708 and in *S*. Heidelberg s 41578 and SA20025921 (Hoffmann et al., 2014, Robertson et al., 2018). Notably, each of the prophages harbouring *sopE* previously reported harboured a distinct allele of the virulence gene which differed by 1 to up 36 SNPs from the allele associated with *S*. 4,[5],12:i:- (Figure 6.12). In contrast, no difference were detected in the *sopE* coding sequences of the phages circulating in S. 4,[5],12;i:- and *S*. Miami suggested a recent event of transfer.

The bacteriophages with the *S*. 4,[5],12:i:- *sopE* allele were found in three distinct clades of the phylogenetic tree. The mTmV prophage from S04698-09 clustered in the P27-like clade with the SfV phage in agreement with the literature (Allison et al., 2002, Petrovska et al., 2016). The mTmV prophages from *S.* Derby SA20035215 and 2014LSAL02547 and in *S.* Brandenburg SA20113174 were virtually identical to mTmV in S04698-09 as they fell in the same cluster at the species level (S32) further suggesting the movement of mTmV between *S*. 4,[5],12:i:- and *S.* Derby (Figure 6.12). The mTmV prophages from *S.* California CD-SL01 and *S.* Derby Sa64 were identical (same species cluster S31). While mTmV from these two species clusters, S32 and S31, were more related with SfV, mTmV2 was more closely related to the SfI phage (Figure 6.12). The two mTmV-like prophages in S01569-10, mTmV_1569 and Ψm1569, as well as the bacteriophage from *S.* 4,[5],12,i:- 17437230, which were diverse from each other, represented the *sopE* phages most divergent from mTmV in the cluster. The pS01569 phage was related with the D6 myovirus, as predicted from the genome comparison (Figure 6.7 and Figure 6.12). Finally, ΦBAA from *S*. Miami ATCC BAA-1586 was related with the P2-like phage Fels-2 and the SopEΦ.

In conclusion, the phylogenetic analysis showed that the *sopE* allele associated with *S.* 4,[5],12,i:- was associated with three phylogenetically distinct groups of bacteriophages suggesting a recent transfer of the *sopE* cassette.

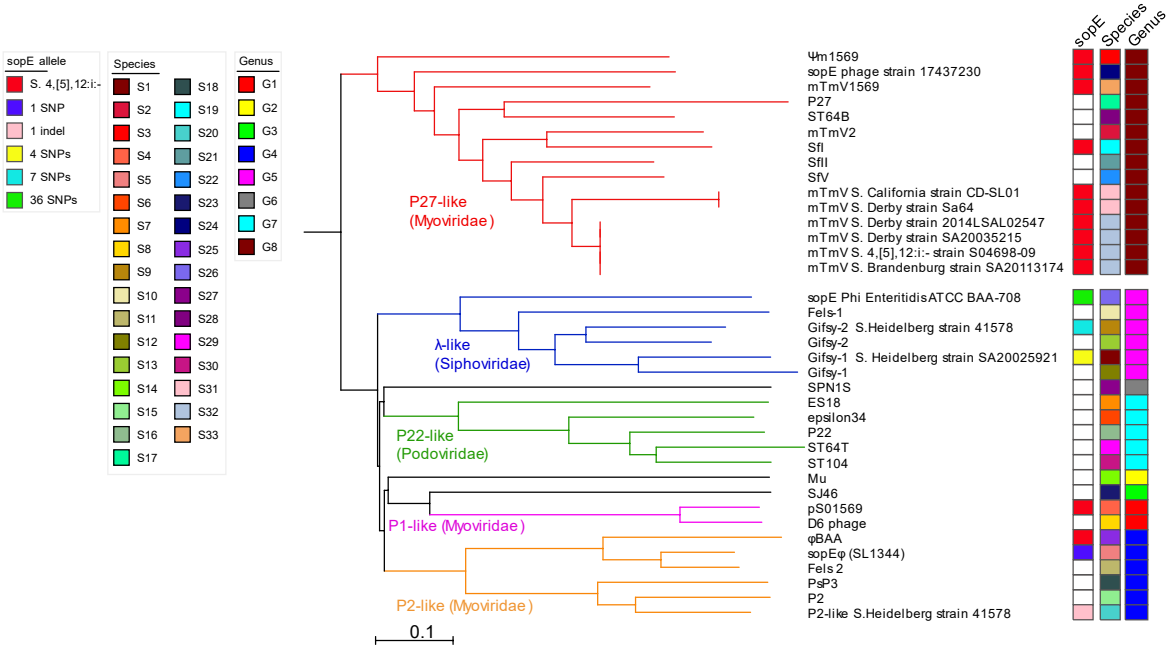

Figure 6.12 The *sopE* allele associated with *S*. 4,[5],12:i:- was harboured by three unrelated phylogenetic groups of phages. Phylogenetic tree inferred from whole-genome sequences of bacteriophages listed in **Table 6.3**. The presence of the *sopE* allele of S. 4,[5],12:i:- or *sopE* variants compared with the reference sequence from strain S04698-09 is reported. In addition, the species (S) and genus (G) identified by VICTOR are reported.

6.2.7 A 2.4 kb cassette carrying the *S*. 4,[5],12:i:- *sopE* allele was conserved between the three distinct phage clusters, P27-like, D6-like and P2-like phages

The presence of an identical *sopE* allele in phylogenetically distinct bacteriophages suggested recent transfer. The transfer of the *sopE* gene between unrelated bacteriophages was postulated to occur *via* homologous recombination as the sequences flanking *sopE* were conserved in distinct bacteriophages (Mirold et al., 2001b).

In order to investigate whether an analogous mechanism of homologous recombination may be responsible for the transfer of the *sopE* allele associated with *S*. 4,[5],12:i:- the genetic context of the *sopE* allele in diverse bacteriophages was studied. To this end, a multiple sequence alignment comprising the *sopE* gene and its flanking sequences was generated and inspected for conserved

regions. The multiple sequence alignment generated, moreover, was analysed at a nucleotide level to detect possible sequence variation that may alter the amino acid sequence of the genes in the cassette or their expression.

The coordinates of the *sopE* cassettes conserved in the unrelated phages mTmV, pS01569 and ΦBAA were first identified by BLAST, then the sequences were isolated and aligned. The prototypical *sop*E cassette of SopEΦ from *S*. Typhimurium SL1344 was included in the analysis. A 2.4 kb cassette comprising the *sopE* gene and upstream region, a tail fiber assembly protein and a hypothetical protein was conserved between mTmV, pS01569 and ΦBAA and between the SopEΦ prophage (Figure 6.13). This conserved sequence was longer than the 1.2 kb *sopE* cassette previously described based on the comparison of SopEΦ with the lambdoid prophage in *S*. *enterica* serovars Hadar, Gallinarum and Enteritidis and the *sopE* region from SARC16 (Mirold et al., 2001b).

In detail, a 2,446 bp cassette conserved in mTmV and pS01569 shared over 99% nucleotide sequence identity (Figure 6.13). mTmV shared a 2,440 bp cassette with ΦBAA at a nucleotide sequence identity of 98%. The different degree of identity indicated that the *sopE* cassette from mTmV and pS01569 likely shared a more recent common ancestor compared with the cassette from ΦBAA. Notably, a cassette of a similar size (2,446 bp) was conserved between mTmV and SopEΦ of *S.* Typhimurium SL1344 (98.6% nucleotide identity). The start and end nucleotide positions of the cassettes in the SopEΦ and pS01569 in reference to the mTmV cassette did not coincide but differed for only 145-150 bases*. sopE* in ΦBAA was located on a large 9.6 kb long fragment which shared 97% sequence identity to a region on the SopEΦ prophage (Figure 6.13). Overall, ΦBAA shared 75% of the genome with SopEΦ (not shown).

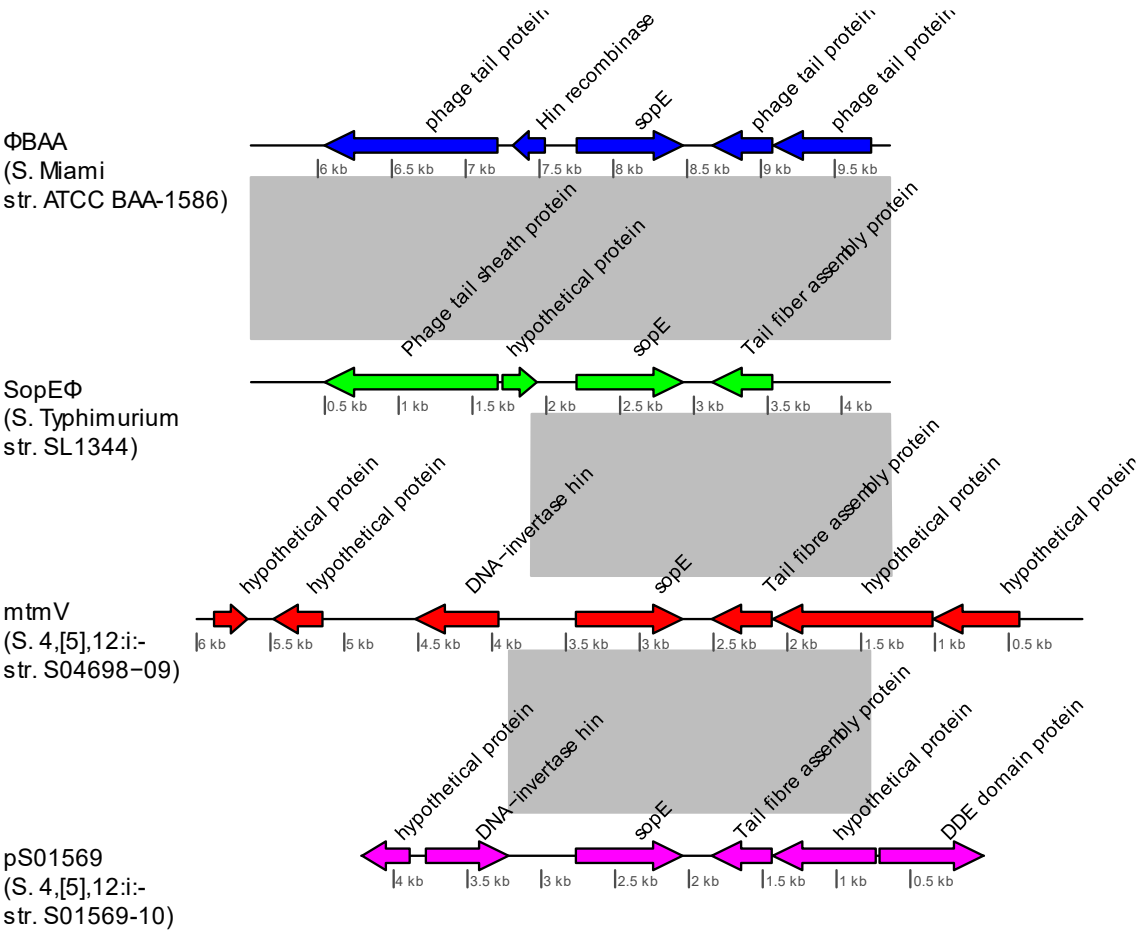

Figure 6.13 A 2.4 kb *sopE* cassette was conserved between ΦBAA, SopEΦ, mTmV and pS01569. Sequence comparison of sopE and flanking sequences in ΦBAA, SopEΦ, mTmV and pS01569 (grey shading indicates >90% nucleotide identity).

The *sopE* cassettes identified in mTmV, pS01569, ΦBAA and SopEΦ were inspected at a nucleotide level using Unipro UGENE (Okonechnikov et al., 2012). Additionally, sequences from the mTmV phage variants mTmV2, mTmV1569 and Ψm1569 were included to determine the degree to which the cassette is conserved in the mTmV-like prophages. A further 500 bp flanking the cassette was included to investigate whether partial homology still exists outside the conserved stretch.

A portion of the *hin* gene and upstream sequence were only partially conserved between the mTmV-like prophages, pS01569 and ΦBAA, whereas the region approximately 300 nucleotides upstream of *sopE* coding sequence was highly conserved across all the prophages (Figure 6.14). Within this region, four SNPs were common to all the prophages carrying *sopE* in *S.* 4,[5],12:i:- (mTmV, mTmV2, mTmV1569, Ψm1569 and pS01569) and ΦBAA from *S.* Miami ATCC

BAA-1586 with reference to SopEΦ. Prediction of the promoter region using BPROM (Solovyev and Salamov, 2010) revealed that one of these SNPs was 3 bases upstream of the site of transcription start and 4 bases downstream the -10 box, a second SNP was two bases from the -35 box, and the remaining two SNPs were 83 and 223 nucleotides upstream of the transcription start site, respectively. Furthermore, ΦBAA presented a single nucleotide insertion 84 nucleotides upstream the transcription start site (Figure 6.14). The proximity of the polymorphisms to regulatory regions may alter the expression of the *sopE* gene.

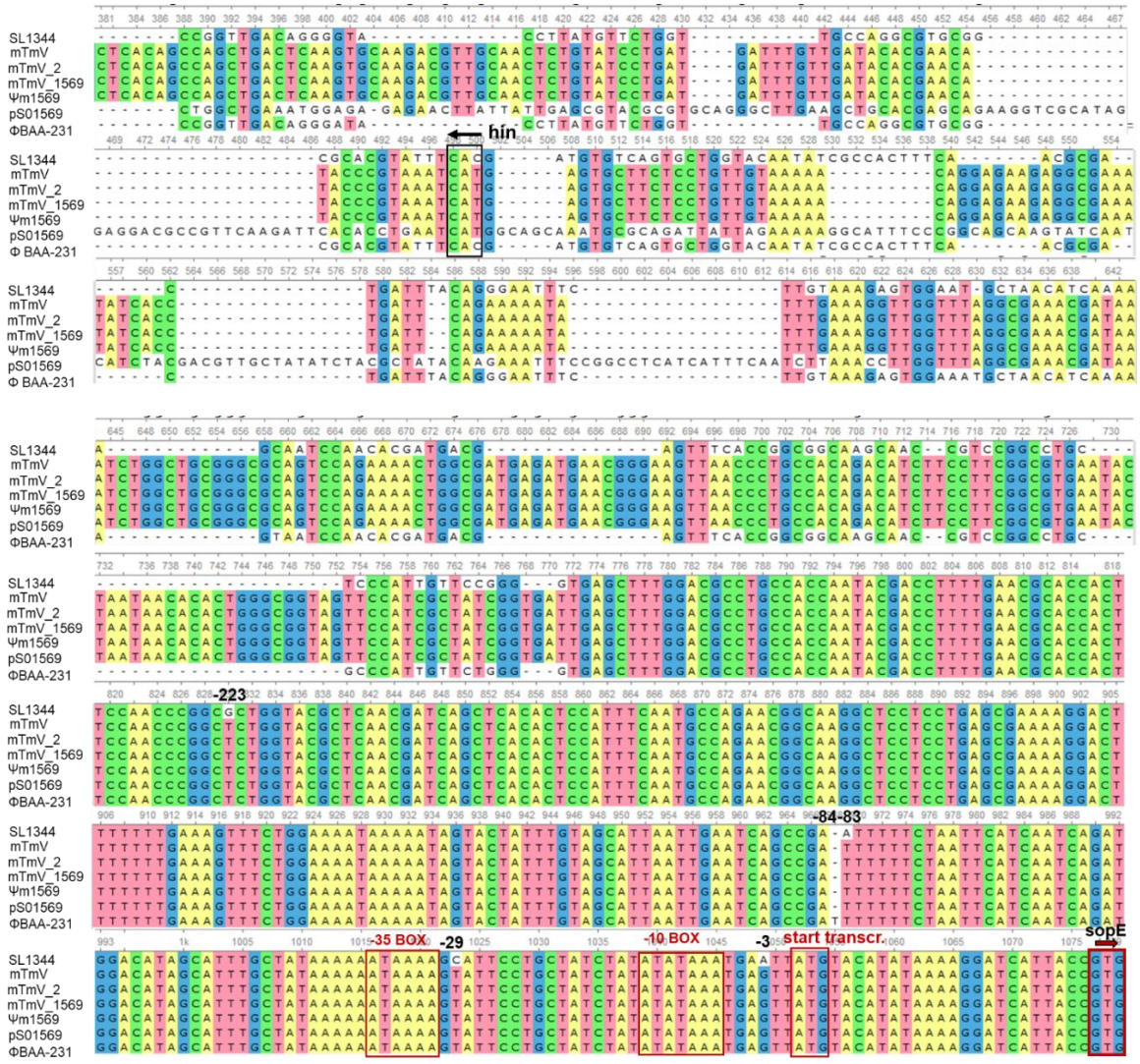

Figure 6.14 Multiple sequence alignment of the 5' flanking sequence of the *sopE* coding sequence in SopEΦ (SL1344), mTmV, mTmV2, mTmV1569, Ψm1569, pS01569 and ΦBAA. The start codon of the *hin* and *sopE* genes are indicated with an arrow. The predicted *sopE* promoter regions are reported. The coordinates of the SNPs and insertion refer to the start transcription site.

The *sopE* coding sequence (723 nt) was identical in mTmV, mTmV2, mTmV_1569, Ψm1569, pS01569 and ΦBAA, and differed by a single nucleotide from the sequence of SopEΦ (Figure 6.15). The polymorphism with respect to *sopE* in SopEΦ was a transversion (G/A) in nucleotide +27 from the start codon. The SNP was in the last nucleotide of the codon (CAG/CAA), and did not alter the subsequent amino acid translation as glutamine.

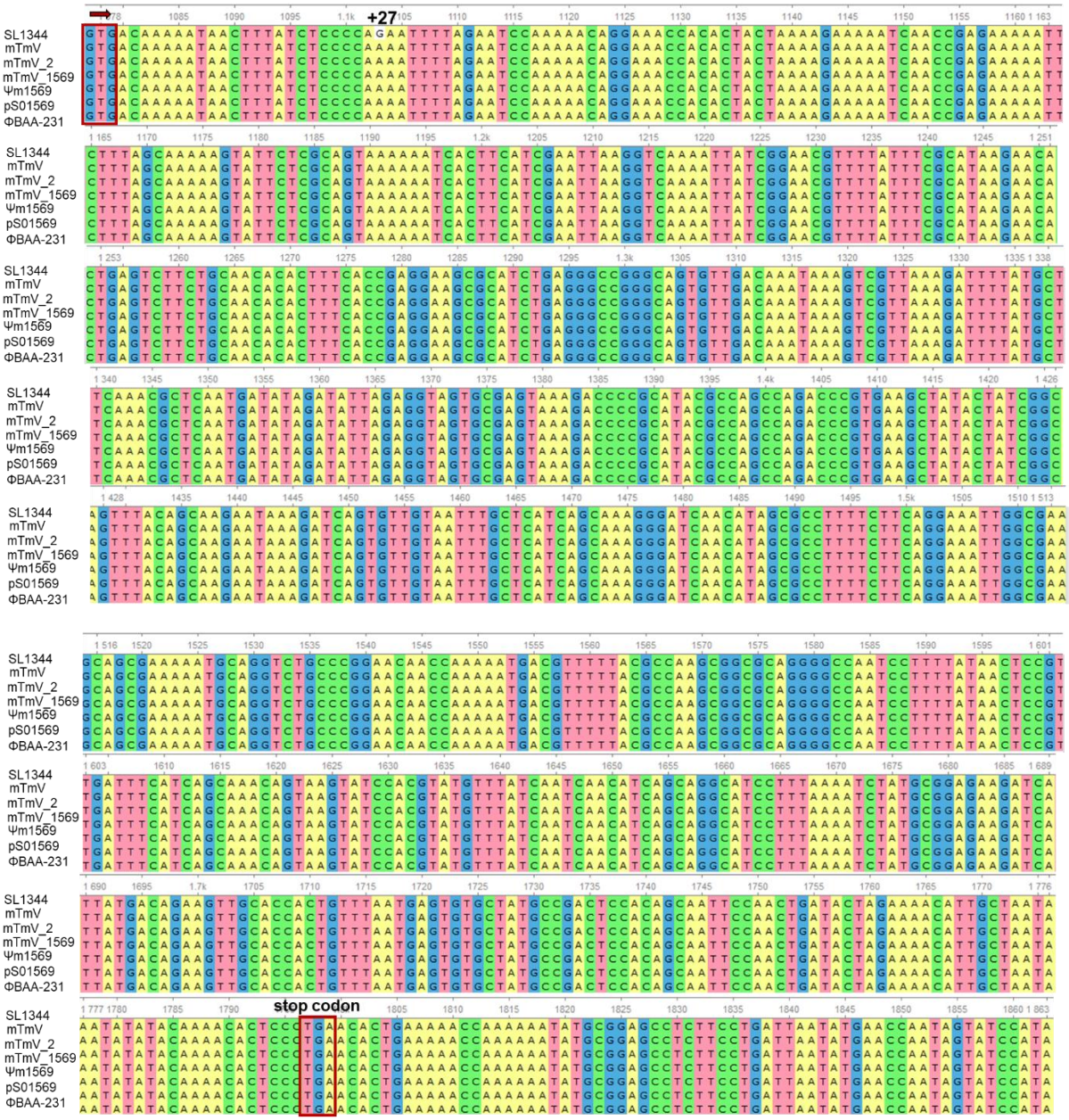

Figure 6.15 Multiple sequence alignment of *sopE* coding sequence in SopEΦ (SL1344), mTmV, mTmV2, mTmV1569, Ψm1569, pS01569 and ΦBAA. The start and stop codons are delimited by a red box; the position of the synonymous SNP (nucleotide 27) identified in reference with the SopEΦ sequence from *S.* Typhimurium SL1344 is indicated.

The intergenic region between *sopE* and the neighbouring prophage gene encoding for a tail fibre assembly protein was identical in all the bacteriophages with reference to SopEΦ prophage, with the exception of a single SNP. The tail fibre assembly protein (408 bp) was conserved among mTmV, mTmV2, mTmV_1569, pS01569 Ψm1569 (except for a single nucleotide deletion in the latter), which differed from the sequence of SopEΦ by only 4 SNPs. In contrast, ΦBAA exhibited 13 SNPs of difference when compared with the sequence of SopEΦ. Variation in tail fibre proteins may influence the host range of the phage as they are often involved in the interaction with the surface microbial receptors (Bertozzi Silva et al., 2016).

In the hypothetical protein flanking the tail fibre assembly gene (1,083 bp), a stretch of 666 nt was overall conserved in all the 7 sequences analysed. In the prophages from *S.* 4,[5],12:i:- it was almost identical, with the exception of a single SNP in mTmV2, and a SNP conserved within the three phages from S01569-10, in reference to the sequence of mTmV. The sequences from the prophages in *S.* 4,[5],12:i:- had 16 SNPs of difference compared with SopEΦ, whereas ΦBAA differed by only 7 SNPs from SopEΦ.

In conclusion, the presence of conserved sequences flanking the *sopE* allele associated with *S.* 4,[5],12:i:- in distinct bacteriophages suggested that *sopE* was likely mobilised by homologous recombination. The finding that in mTmV, pS01569, ΦBAA and SopEΦ bacteriophages the conserved *sopE* cassette was longer than that previously described (Mirold et al., 2001b) suggested that the *sopE* transfer between these bacteriophages occurred more recently. Additionally, variation in *sopE* promoter region was identified in *S.* 4,[5],12:i:- and *S*. Miami in comparison with S. Typhimurium SL1344 which could potentially alter the expression of the virulence gene.

6.3 Discussion

Although *sopE* was mostly associated with the mTmV prophage in clinical *S*. 4,[5],12:i:- isolates from the UK (Chapter 4), novel bacteriophages were identified in the current study associated with *sopE* copy number variation. These bacteriophages were either variants of the previously described mTmV which emerged following exchange of a large portion of the genome with other prophages, or unrelated phages which acquired the *sopE* cassette. In both cases homologous recombination was the likely mechanism of transfer.

Variants of the mTmV prophage comprised mTmV2, associated with the *S*. 4,[5],12:i:- ST34 epidemic in Italy (Palma et al., 2018), and the mTmV1569 prophage, associated with a sub-population of *S*. 4,[5],12:i:- ST34 circulating in the UK. The identification of a recombination breakpoint conserved in mTmV2 and mTmV1569 in reference to mTmV suggested that homologous recombination occurred.

It is known that the presence of homologous sequences may promote homologous recombination between prophages and give rise to novel, mosaic viruses (De Paepe et al., 2014). In the case of S01569-10, one of the mTmV variants (mTmV1569) was generated via recombination between homologous sequences of the ST64B prophage, resident in the tRNA-Ser locus, and mTmV, which caused a 2 Mb genomic inversion. Bacteriophages are known for promoting genomic rearrangements by acting as anchor points for recombination (Brussow et al., 2004), generating further variation within a microbial population.

In both prophage variants, the genes involved in immunity differed from those of mTmV and from each other. This likely expanded the host range of the prophages by circumventing immunity mechanisms mediated by resident prophages, as previously demonstrated (Mirold et al., 2001b), enhancing their chances to infect permissive host cells and promoting the dissemination of the phage-encoded toxin.

An additional mTmV variant identified in *S.* 4,12:i:- 17437230 isolated in Canada had a shorter 14 kb segment in common with mTmV suggesting the presence of multiple recombination breakpoints. Recombination generally occurs between intergenic regions and genes which encode for the same biological function are usually transferred together to minimise the possibility of genomic disruption (Lefeuvre et al., 2009, Hatfull and Hendrix, 2011). The block conserved in all the

mTmV variants contained genes encoding tail proteins that are commonly used for the adsorption and binding to the bacterial receptor (Rakhuba et al., 2010). Additionally, a short prophage (Ψm1569) resulting from duplication and recombination of mTmV sequences, including *sopE*, was identified in S01569-10, resulting in an increased chromosomal dose of *sopE*.

The acquisition of an unrelated, episomal bacteriophage, pS01569, was associated with a further increase of *sopE* dosage. The transfer of the *sopE* cassette between unrelated phages, suggested to occur *via* homologous recombination (Mirold et al., 2001a), further contributes to the dissemination of *sopE*. Although not conserved between the distinct bacteriophages, *hin* DNA invertase genes were always found flanking the *sopE* coding sequence indicating their possible role in the mobilisation of the cassette.

While in the literature phylogenetically distinct *sopE* alleles were shown to be in unrelated bacteriophages within a conserved 1.2 kb cassette (Mirold et al., 2001b), in this study an identical *sopE* allele was described in bacteriophages from multiple families and on a larger conserved region (2.4 kb), suggesting a recent transfer of the cassette. The boundaries of the 2.4 kb *sopE* cassette were not identical in all the bacteriophages indicating the lack of a mechanisms of site-specific recombination.

The bacteriophages carrying the *sopE* allele associated with *S*. 4,[5],12:i:- belonged to three distinct families. The mTmV prophage and its variants were related to P27-like myoviruses. *S*. Miami ATCC BAA-1586 (Boyd et al., 1993) harboured a P2-like prophage with the same *sopE* allele as *S*. 4,[5],12:i:-. The presence of the same polymorphisms of *S*. 4,[5],12:i:- in the coding sequence and in the promoter region of *sopE*, in comparison with the prototypical sequence from SopEΦ, suggested a recent common ancestor between the *sopE* allele in *S*. Miami and *S*. 4,[5],12:i:- and indicated that the allele was already circulating in *Salmonella* and did not emerge in *S*. 4,[5],12:i:-. Whether the SNPs in the coding sequence and in the promoter region of the *sopE* gene affect its expression and activity remains to be studied. Finally, pS01569 was related to P1-like phages; this is the first report of *sopE* in this class of bacteriophages. During the lysogenic cycle, the phage replicates in the form of an autonomous, circular plasmid, as in S01569-10. However, it can also integrate in the genome of the host cell, as demonstrated by the presence of a highly homologous phage into the genome of

S. 4,[5],12:i:- DA34867. pS01569 was highly homologous to circular elements published in GenBank as plasmids from historical *S*. *dysinteriae*. This information, together with the presence of several tRNA loci commonly found on mobile genetic elements in *Escherichia* and *Shigella* suggested that the element commonly circulates in these genera, to which it is not restricted.

This study reported for the first time the presence of three *sopE* copies in a single strain of *Salmonella*. Previously, up to two copies of the *sopE* gene were reported in *S*. Heidelberg associated with an outbreak (Hoffmann et al., 2014). Nevertheless, the two copies differed by 8 SNPs and were harboured by distinct prophages indicating they were acquired by distinct sources. Although the effect of an increased dosage of *sopE* remains un-investigated, enhanced virulence resulting from the increased production of SopE is possible. Analogously, multiple Stx bacteriophages can simultaneously integrate into a single cell resulting in multiple copies of the toxin. Although the presence of two *stx* copies associated with enhanced expression and production of the protein *in vitro*, no correlation with the clinical outcome was identified (Bielaszewska et al., 2006). However, while *stx* expression is activated by the induction of the phage lytic cycle, therefore depending on the stability/vulnerability of the phage, *sopE* expression relies on microbial factors encoded within SPI-1 (InvF/SicA) (Brussow et al., 2004). In addition, the translocation of SopE in the cytosol of the eukaryotic cells requires the chaperon protein InvB. Therefore, the availability of these microbial factors likely limit the amount of *sopE* expressed and translocated into its target location.

From a phage perspective the acquisition of the *sopE* cassette may be beneficial for multiple reasons. It may promote phage maintenance in the host cells as a result of the positive selection on *sopE*. Additionally, the gut inflammation sustained by SopE was shown to promote prophage induction and boost the dissemination of the virus to novel susceptible hosts (Diard et al., 2017).

6.4 Conclusion

The recent transfer of the *sopE* cassette to novel bacteriophages and the emergence of mTmV phage variants suggested that the dissemination of *sopE* was promoted, further corroborating the contribution of *sopE* to the success of *S*. 4,[5],12:i:- ST34. As a result, organisms with multiple *sopE* copies, up to three,

were identified potentially increasing the fitness and pathogenicity of the microbial cells.

7 General Discussion

The pork industry has a large impact on the economy of the European Union, including Ireland. Food safety is a main priority to ensure competitiveness and prosperity of the sector. *Salmonella*, and in particular *S*. Typhimurium, is one of the most common zoonoses associated with pigs and is responsible for food-borne infections in humans. Contamination with *Salmonella* can occur at each step of the farm-to-fork food production chain. However, the presence of *Salmonella* in pigs at slaughter represents a major risk factor for food safety; therefore, the reduction in the prevalence of *Salmonella* on pig farms is a major goal for the pig industry. The spread of resistance to antimicrobials is a global public health problem, which is associated with reduced effectiveness of antimicrobial drugs used for the treatment of clinical infections. The acquisition of antimicrobial resistance (AMR) by zoonotic agents threatens food safety due to the potential transmission of drug-resistant bacteria through the food chain. To this end, this study aimed to understand the population structure and the microevolutionary events that result in changes to phenotypes/characteristics of *Salmonella* important to persistence on the farm, transmission through the food chain to the human population, or AMR. This has the potential to inform the design of more effective strategies to control the zoonotic agent.

The contamination and the persistence of *S*. Typhimurium and the monophasic variant *S*. 4,[5],12;i:- ST34 on pig farms was investigated using whole-genome sequencing (WGS). Each farm was colonised by a specific *S*. Typhimurium or *S*. 4,[5],12;i:- population, often widespread across the farm site and multiple stages of pig production, which persisted between successive production cycles indicating that *S*. Typhimurium and *S*. 4,[5],12;i:- are well adapted to the pig farm environment. The presence of farm-specific clones suggests that the effectiveness of the measures adopted to control the pathogen may vary depending on the genotype and that the use of various interventions may be necessary to identify the most effective strategy for each farm. A number of studies using sub-typing methods reported the ability of *S*. Typhimurium and *S*. 4,[5],12;i:- to persist on pig farms. Nevertheless, the use of WGS in this study enabled me to shed light on the evolution of *S*. Typhimurium and *S*. 4,[5],12;i:- during persistence at a molecular level which affected the biofilm-forming ability and antimicrobial and heavy metal resistance.

This is the first study which investigated the ability of *S*. 4,[5],12:i:- ST34 to produce biofilms, likely an important strategy for environmental survival, transmission and persistence on farms between production cycles. The observation that enhanced biofilm formation in epidemic (*S*. 4,[5],12:i:- ST34 and *S*. Typhimurium DT104) compared to non-epidemic genotypes suggested selection towards this phenotype and the contribution it may play to the success of the isolates. Clones associated with high biofilm-formation may exhibit enhanced resilience to removal during cleaning and disinfection between production cycles. Enhanced cleaning and disinfection are associated with increased cost of production. Nevertheless, no unequivocal association between biofilm formation and persistence was observed which may be explained by the difficulty to reproduce real world conditions *in vitro* or differences in cleaning procedures between pig farms. Despite the high clonality of *S*. 4,[5],12:i:- ST34, enhanced biofilm formation was reported in closely related isolates over a short period of time indicating selection for biofilm-forming ability. *S*. 4,[5],12:i:- ST34 may also become capable of long-term persistence in drinking water systems as a consequence of its ability to produce strong pellicles in the air-liquid interface.

Plasmid-mediated gene flux was a major mechanism of variation over a short period of time. The acquisition of IncHI2 plasmids by a group of *S*. 4,[5],12:i:- ST34 isolated from a single farm and two feed mills conferred enhanced antimicrobial and heavy metal resistance and may contribute to the persistence of the clones in the pig farms and, therefore, in the food chain. The acquisition and spread of antimicrobial resistance (AMR) in zoonotic agents is concerning due to their potential transmission to humans through the food chain. Infections with antimicrobial-resistant *Salmonella* are associated with enhanced severity and frequency of drug treatment failure. The presence of AMR genes on mobile genetic elements facilitates the spread of resistance through horizontal gene transfer mechanisms. A *S*. Typhimurium isolate from the same farm of the IncHI2-positive *S*. 4,[5],12:i:- harboured a related plasmid suggesting the potential dissemination of drug-encoding plasmids on farms to other *Salmonella* genotypes and, conceivably, to other pathogens due to the promiscuous nature of InchI2 plasmids. Feed was the likely source of the plasmids further highlighting the important role of feed in the acquisition and spread of drug-resistant *Salmonella*. The IncHI2 plasmids also encoded resistance to heavy metals, in

particular to silver and copper. Nevertheless, all *S*. 4,[5],12:i:- already harboured an extended chromosomal resistance to heavy metals encoded on *Salmonella* genomic island 4 (SGI-4), whose acquisition drove the emergence of the *S*. 4,[5],12:i:- ST34 clones. In particular, resistance to copper played an important role in the success of the isolates as the heavy metal is extensively used as a growth promoter and antimicrobial in pig rearing. Contrary to other studies, none of the *S*. 4,[5],12:i:- isolates lost the SGI-4 further suggesting the importance of the fitness advantage conferred by the island for their persistence on pig farms. Overall, these data indicate that *S*. Typhimurium and *S*. 4,[5],12:i:- ST34 are well adapted to the farm environment and can rapidly evolve during persistence on pig farms.

The phage-mediated acquisition of the virulence factor SopE contributed to the evolution of *S*. 4,[5],12:i:- ST34. Although no variation in *sopE* presence occurred over a short period of time in pig farms, its acquisition represented a source of variation within *S*. 4,[5],12:i:- ST34 on a larger evolutionary scale. Several lines of evidence suggested that its acquisition was selected during expansion of the *S*. 4,[5],12:i:- ST34 epidemic and that *sopE* played an important role in the emergence and epidemiological success of *S*. 4,[5],12:i:- ST34. These included multiple events of acquisition followed by clonal expansion, acquisition concomitant with the emergence of *S*. 4,[5],12:i:- ST34 in pig farms and dissemination to various genetic elements. The acquisition of *sopE* may not only promote fitness of *S*. 4,[5],12:i:- ST34 in the pig host and their dissemination through enhanced excretion, but also enhance the rate of transmission of the zoonosis to humans with implications for food safety. The location of *sopE* on mobile genetic elements such as temperate bacteriophages facilitated the dissemination of the virulence gene. Cell stresses induced by antimicrobials and gut inflammation can trigger the switch from lysogenic to lytic cycle of temperate bacteriophages with the consequent dissemination of the phage-encoded virulence gene, although mTmV was capable of spontaneous induction. The mTmV phage was predominantly associated with the acquisition of *sopE* in *S*. 4,[5],12:i:- ST34. Although mTmV was never identified in *S*. Typhimurium, it was acquired by other *Salmonella* serovars circulating within the pork and poultry chain potentially leading to the emergence of novel pathogenic variants threatening food safety. The recent transfer of the *sopE* cassette between

unrelated bacteriophages further supported the relevance of *sopE* for *S*. 4,[5],12:i:- ST34. The recombination of the mTmV prophage and the emergence of novel, distinct phages carrying *sopE* was associated with enhanced copy number of the virulence gene. Although it was a rare event and the impact of the dosage remains to be studied, it could potentially increase the pathogenicity of the strains. Also, the emergence of mTmV prophage variants and the transfer of *sopE* between distinct bacteriophages likely contribute to the dissemination of the virulence factor to novel susceptible microbial hosts.

Overall, this study revealed the remarkable level of microevolution of the *S*. 4,[5],12:i:- ST34 epidemic clade over a short time scale, measured in months. Mobile genetic elements (MGEs) played a major role in its microevolution affecting antimicrobial resistance and pathogenicity. Strategies aimed at reducing the transfer of MGEs could reduce the evolution of *Salmonella* on pig farms. For instance, molecules aimed at preventing conjugation have been identified (Getino and De la Cruz, 2018), and vaccination has been suggested as a measure to reduce phage transfer promoted by conditions in the inflamed intestine (Diard et al., 2017).

8 Future Research

- Investigate the impact of the SNP in the *edd* gene in biofilms formation. To test the hypothesis that the SNP in the *edd* gene enhanced the biofilm formation of pig *S*. 4,[5],12;i;- isolates in subclade A1, the SNP will be transferred to a closely related *S*. 4,[5],12;i;- isolate in subclade A2 which displayed a lower biofilm-forming ability and the formation of biofilms assessed by microtiter plate assay. Similarly, the transfer of the wild-type allele of *edd* to an isolate representative of subclade A1 will be performed to test whether it results in lowered biofilm formation. Additionally, isolates with the mutation complemented will also be included to further check the involvement of the SNP in the phenotype observed.

  If the SNP in the *edd* gene associates with increased biofilm formation compared with the wild-type form of the gene in isogenic isolates, characterisation of the Edd protein will be performed aimed at understanding how the presence of the SNP alters the activity of the protein.
- Test the impact of increased gene dosage of *sopE* on pathogenicity of *S*. 4,[5],12:i:-. In order to test the hypothesis that the carriage of multiple copies of *sopE* enhances the pathogenicity of *S*. 4,[5],12:i:-, the invasive ability of isogenic *S*. 4,[5],12:i:- encoding either a single or multiple copies of *sopE* will be compared using epithelial cells *in vitro*. Additionally, i*n vivo* experiments using a mouse model would also be performed in order to compare the ability of the isolates to invade, replicate and cause damage in the host.
- Test whether the polymorphisms in the promoter region of *sopE* in *S*. 4,[5],12:i:- alter the expression of *sopE* in comparison with prototypical *sopE* from *S*. Typhimurium SL1344 by RT-qPCR.

References

ABE, T., IKEMURA, T., OHARA, Y., UEHARA, H., KINOUCHI, M., KANAYA, S., YAMADA, Y., MUTO, A. & INOKUCHI, H. 2009. tRNADB-CE: tRNA gene database curated manually by experts. *Nucleic Acids Res,* 37**,** D163-8.

ADZITEY, F., HUDA, N. & ALI, G. R. 2013. Molecular techniques for detecting and typing of bacteria, advantages and application to foodborne pathogens isolated from ducks. *3 Biotech,* 3**,** 97-107.

ALEKSIC, S., HEINZERLING, F. & BOCKEMÜHL, J. 1996. Human infection caused by salmonellae of subspecies II to VI in Germany, 1977-1992. *Zbl. Bakt.,* 283**,** 391-398.

ALIKHAN, N. F., PETTY, N. K., BEN ZAKOUR, N. L. & BEATSON, S. A. 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics,* 12**,** 402.

ALIKHAN, N. F., ZHOU, Z., SERGEANT, M. J. & ACHTMAN, M. 2018. A genomic overview of the population structure of Salmonella. *PLoS Genet,* 14**,** e1007261.

ALLISON, G. E., ANGELES, D., TRAN-DINH, N. & VERMA, N. K. 2002. Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of Shigella flexneri. *J Bacteriol,* 184**,** 1974-87.

ALLISON, G. E., ANGELES, D. C., THI HUAN, P. & VERMA, N. K. 2003. Morphology of temperate bacteriophage SfV and characterisation of the DNA packaging and capsid genes: the structural genes evolved from two different phage families. *Virology,* 308**,** 114-127.

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.

ALVAREZ, J., PORWOLLIK, S., LACONCHA, I., GISAKIS, V., VIVANCO, A. B., GONZALEZ, I., ECHENAGUSIA, S., ZABALA, N., BLACKMER, F., MCCLELLAND, M., REMENTERIA, A. & GARAIZAR, J. 2003. Detection of a Salmonella enterica Serovar California Strain Spreading in Spanish Feed Mills and Genetic Characterization with DNA Microarrays. *Applied & Environmental Microbiology*.

ANDERSON, E. S., WARD, L. R., SAXE, M. J. & DE SA, J. D. 1977. Bacteriophage-typing designations of Salmonela typhimurium. *J Hyg (Lond),* 78**,** 297-300.

ANDRES, V. M. & DAVIES, R. H. 2015. Biosecurity Measures to Control Salmonella and Other Infectious Agents in Pig Farms: A Review. *Comprehensive Reviews in Food Science and Food Safety,* 14**,** 317-335.

ANDRES, V. M. & DAVIES, R. H. 2105. Biosecurity Measures to Control Salmonella and Other Infectious Agents in Pig Farms: A Review. *Comprehensive Reviews in Food Science and Food Safety,* 14**,** 317-335.

ANDRES-BARRANCO, S., VICO, J. P., GARRIDO, V., SAMPER, S., HERRERA-LEON, S., DE FRUTOS, C. & MAINAR-JAIME, R. C. 2014. Role of wild bird and rodents in the epidemiology of subclinical salmonellosis in finishing pigs. *Foodborne Pathog Dis,* 11**,** 689-97.

ANDREWS-POLYMENIS, H. L., BAUMLER, A. J., MCCORMICK, B. A. & FANG, F. C. 2010. Taming the elephant: Salmonella biology, pathogenesis, and prevention. *Infect Immun,* 78**,** 2356-69.

ANONYMOUS 2010. Scientific Opinion on monitoring and assessment of the public health risk of "Salmonella Typhimurium-like" strains. *EFSA Journal,* 8**,** 1826.

ANRIANY, Y., SAHU, S. N., WESSELS, K. R., MCCANN, L. M. & JOSEPH, S. W. 2006. Alteration of the rugose phenotype in waaG and ddhC mutants of Salmonella enterica serovar Typhimurium DT104 is associated with inverse production of curli and cellulose. *Appl Environ Microbiol,* 72**,** 5002-12.

APHA 2015. OVERVIEW OF SALMONELLA IN LIVESTOCK AND PEOPLE.

APHA 2018a. OVERVIEW OF SALMONELLA IN LIVESTOCK AND PEOPLE.

APHA 2018b. REPORTS OF SALMONELLA IN PIGS.

ARABYAN, N., PARK, D., FOUTOUHI, S., WEIS, A. M., HUANG, B. C., WILLIAMS, C. C., DESAI, P., SHAH, J., JEANNOTTE, R., KONG, N., LEBRILLA, C. B. & WEIMER, B. C. 2016. Salmonella Degrades the Host Glycocalyx Leading to Altered Infection and Glycan Remodeling. *Sci Rep,* 6**,** 29525.

ARAI, N., SEKIZUKA, T., TAMAMURA, Y., TANAKA, K., BARCO, L., IZUMIYA, H., KUSUMOTO, M., HINENOYA, A., YAMASAKI, S., IWATA, T., WATANABE, A., KURODA, M., UCHIDA, I. & AKIBA, M. 2018. Phylogenetic characterization of Salmonella enterica serovar Typhimurium and its monophasic variant isolated from food animals in Japan revealed replacement of major epidemic clones in the last four decades. *J Clin Microbiol*.

ARNDT, D., GRANT, J. R., MARCU, A., SAJED, T., PON, A., LIANG, Y. & WISHART, D. S. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res,* 44**,** W16-21.

BAKER, S., HANAGE, W. P. & HOLT, K. E. 2010. Navigating the future of bacterial molecular epidemiology. *Curr Opin Microbiol,* 13**,** 640-5.

BAKER, S., HARDY, J., SANDERSON, K. E., QUAIL, M., GOODHEAD, I., KINGSLEY, R. A., PARKHILL, J., STOCKER, B. & DOUGAN, G. 2007a. A Novel Linear Plasmid Mediates Flagellar Variation in Salmonella Typhi. *PLoS Pathog.,* 3.

BAKER, S., HOLT, K., WHITEHEAD, S., GOODHEAD, I., PERKINS, T., STOCKER, B., HARDY, J. & DOUGAN, G. 2007b. A linear plasmid truncation induces unidirectional flagellar phase change in H:z66 positive Salmonella Typhi. *Mol Microbiol,* 66**,** 1207-18.

BARKER, J. & BLOOMFIELD, S. F. 2000. Survival of Salmonella in bathrooms and toilets in domestic homes following salmonellosis. *Journal of Applied Microbiology*.

BAUGH, S., EKANAYAKA, A. S., PIDDOCK, L. J. & WEBBER, M. A. 2012. Loss of or inhibition of all multidrug resistance efflux pumps of Salmonella enterica serovar Typhimurium results in impaired ability to form a biofilm. *J Antimicrob Chemother,* 67**,** 2409-17.

BAUMLER, A. & FANG, F. C. 2013. Host specificity of bacterial pathogens. *Cold Spring Harb Perspect Med,* 3**,** a010041.

BAWN, M., THILLIEZ, G., KIRKWOOD, M., WHEELER, N., PETROVSKA, L., DALLMAN, T. J., HALL, N. & KINGSLEY, R. A. 2019. Evolution of Salmonella enterica serotype Typhimurium driven by anthropogenic selection and niche adaptation. *bioRxiv***,** 804674.

BEBEACUA, C., LORENZO FAJARDO, J. C., BLANGY, S., SPINELLI, S., BOLLMANN, S., NEVE, H., CAMBILLAU, C. & HELLER, K. J. 2013. X-ray structure of a superinfection exclusion lipoprotein from phage TP-J34 and identification of the tape measure protein as its target. *Mol Microbiol,* 89**,** 152-65.

BERTOZZI SILVA, J., STORMS, Z. & SAUVAGEAU, D. 2016. Host receptors for bacteriophage adsorption. *FEMS Microbiol Lett,* 363.

BESSER, J., CARLETON, H. A., GERNER-SMIDT, P., LINDSEY, R. L. & TREES, E. 2018. Next-generation sequencing technologies and their application to the study and control of bacterial infections. *Clin Microbiol Infect,* 24**,** 335-341.

BHAN, M. K., BAHL, R. & BHATNAGAR, S. 2005. Typhoid and paratyphoid fever. *Lancet,* 366**,** 749-62.

BIELASZEWSKA, M., PRAGER, R., ZHANG, W., FRIEDRICH, A. W., MELLMANN, A., TSCHÄPE, H. & KARCH, H. 2006. Chromosomal dynamism in progeny of outbreak-related sorbitol-fermenting enterohemorrhagic Escherichia coli O157:NM. *Appl Environ Microbiol,* 72**,** 1900-9.

BILLMAN-JACOBE, H., LIU, Y., HAITES, R., WEAVER, T., ROBINSON, L., MARENDA, M. & DYALL-SMITH, M. 2018. pSTM6-275, a conjugative IncHI2 plasmid of Salmonella that confers antibiotic and heavy metal resistance under changing physiological conditions. *Antimicrob Agents Chemother*.

BISWAS, S., LI, Y., ELBEDIWI, M. & YUE, M. 2019. Emergence and Dissemination of mcr-Carrying Clinically Relevant Salmonella Typhimurium Monophasic Clone ST34. *Microorganisms,* 7.

BOBAY, L. M., ROCHA, E. P. & TOUCHON, M. 2013. The adaptation of temperate bacteriophages to their host genomes. *Mol Biol Evol,* 30**,** 737-51.

BONARDI, S. 2017. Salmonella in the pork production chain and its impact on human health in the European Union. *Epidemiol Infect,* 145**,** 1513-1526.

BONARDI, S., MORGANTI, M., PUPILLO, G. & BRINDANI, F. 2018. Salmonella Brandenburg in the pork chain in Italy: Genetic comparison with the human isolates. *Ital J Food Saf,* 7**,** 6938.

BOSSI, L., FUENTES, J. A., MORA, G. & FIGUEROA-BOSSI, N. 2003. Prophage Contribution to Bacterial Population Dynamics.

BOUCKAERT, R., VAUGHAN, T. G., BARIDO-SOTTANI, J., DUCHÊNE, S., FOURMENT, M., GAVRYUSHKINA, A., HELED, J., JONES, G., KÜHNERT, D., DE MAIO, N., MATSCHINER, M., MENDES, F. K., MÜLLER, N. F., OGILVIE, H. A., DU PLESSIS, L., POPINGA, A., RAMBAUT, A., RASMUSSEN, D., SIVERONI, I., SUCHARD, M. A., WU, C.-H., XIE, D., ZHANG, C., STADLER, T. & DRUMMOND, A. J. 2019. BEAST 2.5: An advanced software platform for Bayesian evolutionary analysis. *PLOS Computational Biology,* 15**,** e1006650.

BOYD, D., CLOECKAERT, A., CHASLUS-DANCLA, E. & MULVEY, M. R. 2002. Characterization of Variant Salmonella Genomic Island 1 Multidrug Resistance Regions from Serovars Typhimurium DT104 and Agona. *Antimicrobial Agents and Chemotherapy,* 46**,** 1714-1722.

BOYD, E. F., WANG, F. S., BELTRAN, P., PLOCK, S. A., NELSON, K. & SELANDER, R. K. 1993. Salmonella reference collection B (SARB): strains of 37 serovars of subspecies I. *Journal of general microbiology,* 139 Pt 6**,** 1125-1132.

BRANCHU, P., BAWN, M. & KINGSLEY, R. A. 2018. Genome Variation and Molecular Epidemiology of Salmonella enterica Serovar Typhimurium Pathovariants. *Infect Immun,* 86.

BRANCHU, P., CHARITY, O. J., BAWN, M., THILLIEZ, G., DALLMAN, T. J., PETROVSKA, L. & KINGSLEY, R. A. 2019. SGI-4 in monophasic Salmonella Typhimurium ST34 is a novel ICE that enhances resistance to copper. *Frontiers in Microbiology*.

BRIANI, F., DEHO, G., FORTI, F. & GHISOTTI, D. 2001. The plasmid status of satellite bacteriophage P4. *Plasmid,* 45**,** 1-17.

BROOKS, J. D. & FLINT, S. H. 2008. Biofilms in the food industry: problems and potential solutions. *International Journal of Food Science & Technology,* 43**,** 2163-2176.

BRUSSOW, H., CANCHAYA, C. & HARDT, W. D. 2004. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev,* 68**,** 560-602, table of contents.

BUCHWALD, G., FRIEBEL, A., GALAN, J. E., HARDT, W. D., WITTINGHOFER, A. & SCHEFFZEK, K. 2002. Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE. *EMBO J,* 21**,** 3286-95.

BURNS, A. M. 2015. *Assessing and managing the risk posed by Salmonella in pig feed.* PhD, Waterford Institute of Technology.

BURNS, A. M., LAWLOR, P. G., GARDINER, G. E., MCCABE, E. M., WALSH, D., MOHAMMED, M., GRANT, J. & DUFFY, G. 2015. Salmonella occurrence and Enterobacteriaceae counts in pig feed ingredients and compound feed from feed mills in Ireland. *Prev Vet Med,* 121**,** 231-9.

CALLOW, B. R. 1959. A new phage-typing scheme for *Salmonella typhimurium*. *J. Hyg.,* 57**,** 346-359.

CAMPOS, J., MOURAO, J., MARCAL, S., MACHADO, J., NOVAIS, C., PEIXE, L. & ANTUNES, P. 2016. Clinical Salmonella Typhimurium ST34 with metal tolerance genes and an IncHI2 plasmid carrying oqxAB-aac(6')-Ib-cr from Europe. *J Antimicrob Chemother,* 71**,** 843-5.

CAMPOS, J., MOURAO, J., PEIXE, L. & ANTUNES, P. 2019. Non-typhoidal Salmonella in the Pig Production Chain: A Comprehensive Analysis of Its Impact on Human Health. *Pathogens,* 8.

CARATTOLI, A., BERTINI, A., VILLA, L., FALBO, V., HOPKINS, K. L. & THRELFALL, E. J. 2005. Identification of plasmids by PCR-based replicon typing. *Journal of microbiological methods,* 63**,** 219-228.

CARATTOLI, A., ZANKARI, E., GARCIA-FERNANDEZ, A., VOLDBY LARSEN, M., LUND, O., VILLA, L., MOLLER AARESTRUP, F. & HASMAN, H. 2014. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother,* 58**,** 3895-903.

CASADEVALL, A. & PIROFSKI, L. A. 2003. The damage-response framework of microbial pathogenesis. *Nat Rev Microbiol,* 1**,** 17-24.

CERI, H., OLSON, M. E., STREMICK, C., READ, R. R., MORCK, D. & BURET, A. 1999. The Calgary Biofilm Device: new technology for rapid determination of antibiotic susceptibilities of bacterial biofilms. *Journal of clinical microbiology,* 37**,** 1771-1776.

CHAN, W., COSTANTINO, N., LI, R., LEE, S. C., SU, Q., MELVIN, D., COURT, D. L. & LIU, P. 2007. A recombineering based approach for high-throughput conditional knockout targeting vector construction. *Nucleic Acids Res,* 35**,** e64.

CHEN, L., ZHENG, D., LIU, B., YANG, J. & JIN, Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. *Nucleic Acids Res,* 44**,** D694-7.

CHIOU, C.-S. 2010. Multilocus variable-number tandem repeat analysis as a molecular tool for subtyping and phylogenetic analysis of bacterial pathogens. *Expert review of molecular diagnostics*.

CHIU, C. H., SU, L. H. & CHU, C. 2004. Salmonella enterica serotype Choleraesuis: epidemiology, pathogenesis, clinical disease, and treatment. *Clin Microbiol Rev,* 17**,** 311-22.

CLARK, R. G., FENWICK, S. G., NICOL, C. M., MARCHANT, R. M., SWANNEY, S., GILL, J. M., HOLMES, J. D., LEYLAND, M. & DAVIES, P. R. 2004. Salmonella Brandenburg - emergence of a new strain affecting stock and humans in the South Island of New Zealand. *N Z Vet J,* 52**,** 26-36.

CLSI 2008. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically; Approved Standard-Seventh Edition. 26.

COBURN, B., GRASSL, G. A. & FINLAY, B. B. 2007. Salmonella, the host and disease: a brief review. *Immunology and cell biology,* 85**,** 112-118.

COHEN, H. J., MECHANDA, S. M. & LIN, W. 1996. PCR amplification of the fimA gene sequence of Salmonella typhimurium, a specific method for detection of Salmonella spp. *Appl Environ Microbiol,* 62**,** 4303-8.

CONDELL, O., IVERSEN, C., COONEY, S., POWER, K. A., WALSH, C., BURGESS, C. & FANNING, S. 2012. Efficacy of biocides used in the modern food industry to control salmonella enterica, and links between biocide tolerance and resistance to clinically relevant antimicrobial compounds. *Appl Environ Microbiol,* 78**,** 3087-97.

CORCORAN, M., MORRIS, D., DE LAPPE, N., O'CONNOR, J., LALOR, P., DOCKERY, P. & CORMICAN, M. 2014. Commonly used disinfectants fail to eradicate Salmonella enterica biofilms from food contact surface materials. *Appl Environ Microbiol,* 80**,** 1507-14.

CROUCHER, N. J., PAGE, A. J., CONNOR, T. R., DELANEY, A. J., KEANE, J. A., BENTLEY, S. D., PARKHILL, J. & HARRIS, S. R. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res,* 43**,** e15.

CRUM-CIANFLONE, N. F. 2008. Salmonellosis and the gastrointestinal tract: more than just peanut butter. *Curr Gastroenterol Rep,* 10**,** 424-31.

DALLMAN, T., ASHTON, P., SCHAFER, U., JIRONKIN, A., PAINSET, A., SHAABAN, S., HARTMAN, H., MYERS, R., UNDERWOOD, A., JENKINS, C. & GRANT, K.

2018. SnapperDB: a database solution for routine sequencing analysis of bacterial isolates. *Bioinformatics,* 34**,** 3028-3029.

DANECEK, P., AUTON, A., ABECASIS, G., ALBERS, C. A., BANKS, E., DEPRISTO, M. A., HANDSAKER, R. E., LUNTER, G., MARTH, G. T., SHERRY, S. T., MCVEAN, G., DURBIN, R. & GENOMES PROJECT ANALYSIS, G. 2011. The variant call format and VCFtools. *Bioinformatics,* 27**,** 2156-8.

DATSENKO, K. A. & WANNER, B. L. 2000. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc Natl Acad Sci U S A,* 97**,** 6640-5.

DE LA TORRE, E., ZAPATA, D., TELLO, M., MEJIA, W., FRIAS, N., GARCIA PENA, F. J., MATEU, E. M. & TORRE, E. 2003. 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,* 41**,** 2395-400.

DE LUCIA, A., RABIE, A., SMITH, R. P., DAVIES, R., OSTANELLO, F., AJAYI, D., PETROVSKA, L. & MARTELLI, F. 2018. Role of wild birds and environmental contamination in the epidemiology of Salmonella infection in an outdoor pig farm. *Vet Microbiol,* 227**,** 148-154.

DE PAEPE, M., HUTINET, G., SON, O., AMARIR-BOUHRAM, J., SCHBATH, S. & PETIT, M. A. 2014. Temperate phages acquire DNA from defective prophages by relaxed homologous recombination: the role of Rad52-like recombinases. *PLoS Genet,* 10**,** e1004181.

DELESALLE, V. A., TANKE, N. T., VILL, A. C. & KRUKONIS, G. P. 2016. Testing hypotheses for the presence of tRNA genes in mycobacteriophage genomes. *Bacteriophage,* 6**,** e1219441.

DIARD, M., BAKKEREN, E., CORNUAULT, J. K., MOOR, K., HAUSMANN, A., SELLIN, M. E., LOVERDO, C., AERTSEN, A., ACKERMANN, M., DE PAEPE, M., SLACK, E. & HARDT, W. D. 2017. Inflammation boosts bacteriophage transfer between Salmonella spp. *Science,* 355**,** 1211-1215.

DIDELOT, X., CROUCHER, N. J., BENTLEY, S. D., HARRIS, S. R. & WILSON, D. J. 2018. Bayesian inference of ancestral dates on bacterial phylogenetic trees. *Nucleic Acids Res,* 46**,** e134.

DONLAN, R. M. 2001. Biofilm Formation: A Clinically Relevant Microbiological Process. *Clinical Infectious Diseases,* 33.

DONLAN, R. M. & COSTERTON, J. W. 2002. Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev,* 15**,** 167-93.

DYALL-SMITH, M. L., LIU, Y. & BILLMAN-JACOBE, H. 2017. Genome Sequence of an Australian Monophasic Salmonella enterica subsp. enterica Typhimurium Isolate

(TW-Stm6) Carrying a Large Plasmid with Multiple Antimicrobial Resistance Genes. *Genome Announc,* 5.

DYSON, Z. A., THANH, D. P., BODHIDATTA, L., MASON, C. J., SRIJAN, A., RABAA, M. A., VINH, P. V., THANH, T. H., THWAITES, G. E., BAKER, S. & HOLT, K. E. 2017. Whole Genome Sequence Analysis of Salmonella Typhi Isolated in Thailand before and after the Introduction of a National Immunization Program. *PLoS Negl Trop Dis,* 11**,** e0005274.

ECDC 2015. Expert Opinion on the introduction of next-generation typing methods for food- and waterborne diseases in the EU and EEA.

ECDC, E. 2017. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2016. *EFSA Journal,* 15.

ECDC, E. A. 2018. The European Union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2017. *EFSA Journal,* 16.

ECHEITA, M. A., ALADUENA, A., CRUCHAGA, S. & USERA, M. A. 1999. Emergence and Spread of an Atypical Salmonella enterica subsp. enterica Serotype 4,5,12:i:- Strain in Spain. *Journal of Clinical Microbiology*.

ECHEITA, M. A., HERRERA, S. & USERA, M. A. 2001. Atypical, fljB-negative Salmonella enterica subsp. enterica strain of serovar 4,5,12:i:- appears to be a monophasic variant of serovar Typhimurium. *J Clin Microbiol,* 39**,** 2981-3.

EFSA 2010a. Scientific Opinion on a Quantitative Microbiological Risk Assessment ofSalmonellain slaughter and breeder pigs. *EFSA Journal,* 8.

EFSA 2010b. Scientific Opinion on monitoring and assessment of the public health risk of “SalmonellaTyphimurium-like” strains. *EFSA Journal,* 8.

EFSA 2013. Scientific Opinion on the evaluation of molecular typing methods for major food-borne microbiological hazards and their use for attribution modelling, outbreak investigation and scanning surveillance: Part 1 (evaluation of methods and applications). *EFSA Journal,* 11.

EFSA 2014. Scientific Opinion on the evaluation of molecular typing methods for major food-borne microbiological hazards and their use for attribution modelling, outbreak investigation and scanning surveillance: Part 2 (surveillance and data management activities). *EFSA Journal,* 12.

EHRBAR, K. & HARDT, W. D. 2005. Bacteriophage-encoded type III effectors in Salmonella enterica subspecies 1 serovar Typhimurium. *Infect Genet Evol,* 5**,** 1-9.

ELLINGSON, J. L., ANDERSON, J. L., CARLSON, S. A. & SHARMA, V. K. 2004. Twelve hour real-time PCR technique for the sensitive and specific detection of Salmonella in raw and ready-to-eat meat products. *Mol Cell Probes,* 18**,** 51-7.

ELNEKAVE, E., HONG, S., MATHER, A. E., BOXRUD, D., TAYLOR, A. J., LAPPI, V., JOHNSON, T. J., VANNUCCI, F., DAVIES, P., HEDBERG, C., PEREZ, A. & ALVAREZ, J. 2018. Salmonella enterica Serotype 4,[5],12:i:- in Swine in the United States Midwest: An Emerging Multidrug-Resistant Clade. *Clin Infect Dis,* 66**,** 877-885.

ENG, S.-K., PUSPARAJAH, P., AB MUTALIB, N.-S., SER, H.-L., CHAN, K.-G. & LEE, L.-H. 2015. Salmonella: A review on pathogenesis, epidemiology and antibiotic resistance. *Frontiers in Life Science,* 8**,** 284-293.

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

FEASEY, N. A., DOUGAN, G., KINGSLEY, R. A., HEYDERMAN, R. S. & GORDON, M. A. 2012. Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. *Lancet,* 379**,** 2489-99.

FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. *J Mol Evol,* 17**,** 368-76.

FIGUEROA-BOSSI, N., UZZAU, S., MALORIOL, D. & BOSSI, L. 2001. Variable assortment of prophages provides a transferable repertoire of pathogenic determinants in Salmonella. *Mol Microbiol,* 39**,** 260-272.

FOLEY, S. L., ZHAO, S. & WALKER, R. D. 2007. Comparison of molecular typing methods for the differentiation of Salmonella foodborne pathogens. *Foodborne Pathog Dis,* 4**,** 253-76.

FRIEBEL, A. & HARDT, W. D. 2000. Purification and biochemical activity of Salmonella exchange factor SopE. *Methods Enzymol,* 325**,** 82-91.

FRIEBEL, A., ILCHMANN, H., AEPFELBACHER, M., EHRBAR, K., MACHLEIDT, W. & HARDT, W. D. 2001. SopE and SopE2 from Salmonella typhimurium activate different sets of RhoGTPases of the host cell. *J Biol Chem,* 276**,** 34035-40.

FU, L., NIU, B., ZHU, Z., WU, S. & LI, W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics,* 28**,** 3150-2.

FUX, C. A., COSTERTON, J. W., STEWART, P. S. & STOODLEY, P. 2005. Survival strategies of infectious biofilms. *Trends Microbiol,* 13**,** 34-40.

GAL-MOR, O., BOYLE, E. C. & GRASSL, G. A. 2014. Same species, different diseases: how and why typhoidal and non-typhoidal Salmonella enterica serovars differ. *Front Microbiol,* 5**,** 391.

GARAIZAR, J., PORWOLLIK, S., ECHEITA, A., REMENTERIA, A., HERRERA, S., WONG, R. M. Y., FRYE, J., USERA, M. A. & MCCLELLAND, M. 2002. DNA Microarray-Based Typing of an Atypical Monophasic Salmonella enterica Serovar. *Journal of Clinical Microbiology,* 40**,** 2074-2078.

GARCIA, P., GUERRA, B., BANCES, M., MENDOZA, M. C. & RODICIO, M. R. 2011. IncA/C plasmids mediate antimicrobial resistance linked to virulence genes in the Spanish clone of the emerging Salmonella enterica serotype 4,[5],12:i. *J Antimicrob Chemother,* 66**,** 543-9.

GARCIA, P., MALORNY, B., HAUSER, E., MENDOZA, M. C. & RODICIO, M. R. 2013. Genetic types, gene repertoire, and evolution of isolates of the Salmonella enterica serovar 4,5,12:i:- Spanish clone assigned to different phage types. *J Clin Microbiol,* 51**,** 973-8.

GARCIA-DOVAL, C. & VAN RAAIJ, M. J. 2013. Bacteriophage Receptor Recognition and Nucleic Acid Transfer. *In:* MATEU, M. G. (ed.) *Structure and Physics of Viruses: An Integrated Textbook.* Dordrecht: Springer Netherlands.

GARRISON, E. & MARTH, G. 2012. Haplotype-based variant detection from short-read sequencing.

GEBREYES, W. A., ALTIER, C. & THAKUR, S. 2006. Molecular epidemiology and diversity of Salmonella serovar Typhimurium in pigs using phenotypic and genotypic approaches. *Epidemiol Infect,* 134**,** 187-98.

GERSTEL, U. & ROMLING, U. 2001. Oxygen tension and nutrient starvation are major signals that regulate agfD promoter activity and expression of the multicellular morphotype in Salmonella typhimurium. *Environmental Microbiology*.

GETINO, M. & DE LA CRUZ, F. 2018. Natural and Artificial Strategies To Control the Conjugative Transmission of Plasmids. *Microbiol Spectrum,* 6.

GIANNELLA, R. A., BROITMAN, S. A. & ZAMCHECK, N. 1971. Salmonella enteritis. II. Fulminant diarrhea in and effects on the small intestine. *Am J Dig Dis,* 16**,** 1007-13.

GILCREASE, E. B. & CASJENS, S. R. 2018. The genome sequence of Escherichia coli tailed phage D6 and the diversity of Enterobacteriales circular plasmid prophages. *Virology,* 515**,** 203-214.

GILLEN, K. L. & HUGHES, K. T. 1991. Negative regulatory loci coupling flagellin synthesis to flagellar assembly in Salmonella typhimurium. *J Bacteriol,* 173**,** 2301-10.

GOKER, M., GARCIA-BLAZQUEZ, G., VOGLMAYR, H., TELLERIA, M. T. & MARTIN, M. P. 2009. Molecular taxonomy of phytopathogenic fungi: a case study in Peronospora. *PLoS One,* 4**,** e6319.

GRIMONT, P. A. & WEILL, F. X. 2007. Antigenic Formulae of the Salmonella serovars. *WHO Collaborating Centre for Reference and Research on Salmonella,* 9th edition.

GUERRA, B., SOTO, S. M., ARGUELLES, J. M. & MENDOZA, M. C. 2001. Multidrug resistance is mediated by large plasmids carrying a class 1 integron in the emergent Salmonella enterica serotype [4,5,12:i:-]. *Antimicrob Agents Chemother,* 45**,** 1305-8.

GUY, L., KULTIMA, J. R. & ANDERSSON, S. G. 2010. genoPlotR: comparative gene and genome visualization in R. *Bioinformatics,* 26**,** 2334-5.

HAMILTON, D., HOLDS, G., HOGG, G., VALCANIS, M. & KIERMEIER, A. 2015. Longitudinal study of an Australian pig farm infected with monophasic Salmonella Typhimurium-like PT193 (1,4,[5],12:i:- PT193) using MLVA. *Safe Pork: Epidemiology and Control of Hazards in Pork Production Chain*.

HANAHAN, D. 1983. Studies on transformation of Escherichia coli with plasmids. *J Mol Biol,* 166**,** 557-80.

HARDT, W. D., CHEN, L. M., SCHUEBEL, K. E., BUSTELO, X. R. & GALAN, J. E. 1998a. S. typhimurium encodes an activator of Rho GTPases that induces membrane ruffling and nuclear responses in host cells. *Cell,* 93**,** 815-26.

HARDT, W. D., URLAUB, H. & GALAN, J. E. 1998b. A substrate of the centisome 63 type III protein secretion system of Salmonella typhimurium is encoded by a cryptic bacteriophage. *Proc Natl Acad Sci U S A,* 95**,** 2574-9.

HARRISON, E. & BROCKHURST, M. A. 2017. Ecological and Evolutionary Benefits of Temperate Phage: What Does or Doesn't Kill You Makes You Stronger. *Bioessays,* 39.

HASEGAWA, M., KISHINO, H. & YANO, T. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. *J Mol Evol,* 22**,** 160-74.

HATFULL, G. F. & HENDRIX, R. W. 2011. Bacteriophages and their genomes. *Curr Opin Virol,* 1**,** 298-303.

HAUSER, E., TIETZE, E., HELMUTH, R., JUNKER, E., BLANK, K., PRAGER, R., RABSCH, W., APPEL, B., FRUTH, A. & MALORNY, B. 2010. Pork contaminated with Salmonella enterica serovar 4,[5],12:i:-, an emerging health risk for humans. *Appl Environ Microbiol,* 76**,** 4601-10.

HENDRIKSEN, R. S., PEDERSEN, S. K., LEEKITCHAROENPHON, P., MALORNY, B., BOROWIAK, M., BATTISTI, A., FRANCO, A., ALBA, P., CARFORA, V., RICCI, A., MASTRORILLI, E., LOSASSO, C., LONGO, A., PETRIN, S., BARCO, L., WOŁKOWICZ, T., GIERCZYŃSKI, R., ZACHARCZUK, K., WOLANIUK, N., WASYL, D., ZAJAC, M., WIECZOREK, K., PŹOŁTORAK, K., PETROVSKA-HOLMES, L., DAVIES, R., TANG, Y., GRANT, K., UNDERWOOD, A., DALLMAN, T., PAINSET, A., HARTMAN, H., AL-SHABIB, A. & COWLEY, L.

2018. Final report of ENGAGE - Establishing Next Generation sequencing Ability for Genomic analysis in Europe. *EFSA Supporting Publications,* 15.

HENDRIX, R. W., LAWRENCE, J. G., HATFULL, G. F. & CASJENS, S. R. 2000. The origins and ongoing evolution of viruses. *Trends in Microbiol.,* 8.

HO, T. D., FIGUEROA-BOSSI, N., WANG, M., UZZAU, S., BOSSI, L. & SLAUCH, J. M. 2002. Identification of GtgE, a novel virulence factor encoded on the Gifsy-2 bacteriophage of Salmonella enterica serovar Typhimurium. *J Bacteriol,* 184**,** 5234-9.

HOELZER, K., MORENO SWITT, A. I. & WIEDMANN, M. 2011. Animal contact as a source of human non-typhoidal salmonellosis. *Vet Res,* 42**,** 34.

HOELZER, K., SOYER, Y., RODRIGUEZ-RIVERA, L. D., CUMMINGS, K. J., MCDONOUGH, P. L., SCHOONMAKER-BOPP, D. J., ROOT, T. P., DUMAS, N. B., WARNICK, L. D., GROHN, Y. T., WIEDMANN, M., BAKER, K. N., BESSER, T. E., HANCOCK, D. D. & DAVIS, M. A. 2010. The prevalence of multidrug resistance is higher among bovine than human Salmonella enterica serotype Newport, Typhimurium, and 4,5,12:i:- isolates in the United States but differs by serotype and geographic region. *Appl Environ Microbiol,* 76**,** 5947-59.

HOFER, B., RUGE, M. & DREISEIKELMANN, B. 1995. The superinfection exclusion gene (sieA) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. *J Bacteriol,* 177**,** 3080-6.

HOFFMAN, S., MACULLOCH, B. & BATZ, M. 2015. Economic Burden of Major Foodborne Illnesses Acquired in the United States. *United States Department of Agriculture*.

HOFFMANN, M., ZHAO, S., PETTENGILL, J., LUO, Y., MONDAY, S. R., ABBOTT, J., AYERS, S. L., CINAR, H. N., MURUVANDA, T., LI, C., ALLARD, M. W., WHICHARD, J., MENG, J., BROWN, E. W. & MCDERMOTT, P. F. 2014. Comparative genomic analysis and virulence differences in closely related salmonella enterica serotype heidelberg isolates from humans, retail meats, and animals. *Genome Biol Evol,* 6**,** 1046-68.

HOIBY, N., BJARNSHOLT, T., GIVSKOV, M., MOLIN, S. & CIOFU, O. 2010. Antibiotic resistance of bacterial biofilms. *Int J Antimicrob Agents,* 35**,** 322-32.

HOOTON, S. P., ATTERBURY, R. J. & CONNERTON, I. F. 2011. Application of a bacteriophage cocktail to reduce Salmonella Typhimurium U288 contamination on pig skin. *Int J Food Microbiol,* 151**,** 157-63.

HOOTON, S. P., TIMMS, A. R., CUMMINGS, N. J., MORETON, J., WILSON, R. & CONNERTON, I. F. 2014. The complete plasmid sequences of Salmonella enterica serovar Typhimurium U288. *Plasmid,* 76**,** 32-9.

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.

HOPKINS, K. L., KIRCHNER, M., GUERRA, B., GRANIER, S. A., LUCARELLI, C., PORRERO, M. C., JACKUBCZAK, A., THRELFALL, E. J. & MEVIUS, D. 2010. Multiresistant Salmonella enterica serovar 4,[5],12:i:- in Europe: a new pandemic strain? *Euro Surveill.*

HOPKINS, K. L. & THRELFALL, E. J. 2004. Frequency and polymorphism of sopE in isolates of Salmonella enterica belonging to the ten most prevalent serotypes in England and Wales. *J Med Microbiol,* 53**,** 539-43.

HUMPHREYS, D., DAVIDSON, A., HUME, P. J. & KORONAKIS, V. 2012. Salmonella virulence effector SopE and Host GEF ARNO cooperate to recruit and activate WAVE to trigger bacterial invasion. *Cell Host Microbe,* 11**,** 129-39.

HUNT, M., MATHER, A. E., SÁNCHEZ-BUSÓ, L., PAGE, A. J., PARKHILL, J., KEANE, J. A. & HARRIS, S. R. 2017a.

HUNT, M., MATHER, A. E., SÁNCHEZ-BUSÓ, L., PAGE, A. J., PARKHILL, J., KEANE, J. A. & HARRIS, S. R. 2017b. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads.

HUNT, M., SILVA, N. D., OTTO, T. D., PARKHILL, J., KEANE, J. A. & HARRIS, S. R. 2015. Circlator: automated circularization of genome assemblies using long sequencing reads. *Genome Biol,* 16**,** 294.

IDE, N. & KUTSUKAKE, K. 1997. Identification of a novel Escherichia coli gene whose expression is dependent on the flagellum-specific sigma factor, FliA, but dispensable for motility development. *Gene***,** 19–23.

INOUYE, M., DASHNOW, H., RAVEN, L. A., SCHULTZ, M. B., POPE, B. J., TOMITA, T., ZOBEL, J. & HOLT, K. E. 2014. SRST2: Rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med,* 6**,** 90.

ISHIKAWA, S. A., ZHUKOVA, A., IWASAKI, W. & GASCUEL, O. 2019. A Fast Likelihood Method to Reconstruct and Visualize Ancestral Scenarios. *Mol Biol Evol*.

JASLOW, S. L., GIBBS, K. D., FRICKE, W. F., WANG, L., PITTMAN, K. J., MAMMEL, M. K., THADEN, J. T., FOWLER, V. G., JR., HAMMER, G. E., ELFENBEIN, J. R. & KO, D. C. 2018. Salmonella Activation of STAT3 Signaling by SarA Effector Promotes Intracellular Replication and Production of IL-10. *Cell Rep,* 23**,** 3525-3536.

JONAS, K., TOMENIUS, H., KADER, A., NORMARK, S., ROMLING, U., BELOVA, L. M. & MELEFORS, O. 2007. Roles of curli, cellulose and BapA in Salmonella biofilm morphology studied by atomic force microscopy. *BMC Microbiol,* 7**,** 70.

JUKES, T. H. & CANTOR, C. R. 1969. *Evolution of protein molecules*.

KADO, C. I. & LIU, S. T. 1981. Rapid procedure for detection and isolation of large and small plasmids. *J Bacteriol,* 145**,** 1365-73.

KIDGELL, C., REICHARD, U., WAIN, J., LINZ, B., TORPDAHL, M., DOUGAN, G. & ACHTMAN, M. 2002. Salmonella typhi, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect Genet Evol,* 2**,** 39-45.

KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution,* 16**,** 111-120.

KINGSLEY, R. A., KAY, S., CONNOR, T., BARQUIST, L., SAIT, L., HOLT, K. E., SIVARAMAN, K., WILEMAN, T., GOULDING, D., CLARE, S., HALE, C., SESHASAYEE, A., HARRIS, S., THOMSON, N. R., GARDNER, P., RABSCH, W., WIGLEY, P., HUMPHREY, T., PARKHILL, J. & DOUGAN, G. 2013. Genome and transcriptome adaptation accompanying emergence of the definitive type 2 host-restricted Salmonella enterica serovar Typhimurium pathovar. *MBio,* 4**,** e00565-13.

KINGSLEY, R. A., MSEFULA, C. L., THOMSON, N. R., KARIUKI, S., HOLT, K. E., GORDON, M. A., HARRIS, D., CLARKE, L., WHITEHEAD, S., SANGAL, V., MARSH, K., ACHTMAN, M., MOLYNEUX, M. E., CORMICAN, M., PARKHILL, J., MACLENNAN, C. A., HEYDERMAN, R. S. & DOUGAN, G. 2009. Epidemic multiple drug resistant Salmonella Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res,* 19**,** 2279-87.

KLIEM, M. & DREISEIKELMANN, B. 1989. The superimmunity gene sim of bacteriophage P1 causes superinfection exclusion. *Virology,* 171**,** 350-5.

KNODLER, L. A. & STEELE-MORTIMER, O. 2003. Taking possession: biogenesis of the Salmonella-containing vacuole. *Traffic,* 4**,** 587-99.

KORBER, B. 2000. HIV Signature and Sequence Variation Analysis. *Computational Analysis of HIV Molecular Sequences, Chapter 4***,** 55-72.

KROPINSKI, A. M., SULAKVELIDZE, A., KONCZY, P. & POPPE, C. 2007. Salmonella Phages and Prophages—Genomics and Practical Aspects. *In:* SCHATTEN, H. & EISENSTARK, A. (eds.) *Salmonella: Methods and Protocols.* Totowa, NJ: Humana Press.

KURTZ, S., PHILLIPPY, A., DELCHER, A. L., SMOOT, M., SHUMWAY, M., ANTONESCU, C. & SALZBERG, S. L. 2004. Versatile and open software for comparing large genomes. *Genome Biol,* 5**,** R12.

LANGMEAD, B. & SALZBERG, S. L. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods,* 9**,** 357-9.

LAORDEN, L., HERRERA-LEON, S., MARTINEZ, I., SANCHEZ, A., KROMIDAS, L., BIKANDI, J., REMENTERIA, A., ECHEITA, A. & GARAIZAR, J. 2010. Genetic evolution of the Spanish multidrug-resistant Salmonella enterica 4,5,12:i:- monophasic variant. *J Clin Microbiol,* 48**,** 4563-6.

LATASA, C., ROUX, A., TOLEDO-ARANA, A., GHIGO, J. M., GAMAZO, C., PENADES, J. R. & LASA, I. 2005. BapA, a large secreted protein required for biofilm formation and host colonization of Salmonella enterica serovar Enteritidis. *Mol Microbiol,* 58**,** 1322-39.

LEFEUVRE, P., LETT, J. M., VARSANI, A. & MARTIN, D. P. 2009. Widely conserved recombination patterns among single-stranded DNA viruses. *J Virol,* 83**,** 2697-707.

LEFORT, V., DESPER, R. & GASCUEL, O. 2015. FastME 2.0: A Comprehensive, Accurate, and Fast Distance-Based Phylogeny Inference Program. *Mol Biol Evol,* 32**,** 2798-800.

LETUNIC, I. & BORK, P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Research,* 47**,** W256-W259.

LEWANDOWSKI, Z. & EVANS, L. V. 2000. Structure and function of biofilms,. *Biofilms: recent advances in their study and control, Amsterdam Harwood Academic Publishers***,** 1-17.

LI, H. 2015. FermiKit: assembly-based variant calling for Illumina resequencing data. *Bioinformatics,* 31**,** 3694-6.

LI, H. 2016. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics,* 32**,** 2103-2110.

LI, H. & DURBIN, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England),* 25**,** 1754-1760.

LI, H., HANDSAKER, B., WYSOKER, A., FENNELL, T., RUAN, J., HOMER, N., MARTH, G., ABECASIS, G., DURBIN, R. & GENOME PROJECT DATA PROCESSING, S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics,* 25**,** 2078-9.

LI, Y., PULFORD, C. V., DÍAZ, P., PEREZ-SEPULVEDA, B. M., DUARTE, C., PREDEUS, A. V., WIESNER, M., HEAVENS, D., LOW, R., SCHUDOMA, C., LIPSCOMBE, J., MONTAÑO, A., HALL, N., MORENO, J. & HINTON, J. C. D. 2019. Genomic and phylogenetic analysis of Salmonella Typhimurium and its monophasic variants responsible for invasive endemic infections in Colombia. *bioRxiv*.

LIM, J. S., SHIN, M., KIM, H. J., KIM, K. S., CHOY, H. E. & CHO, K. A. 2014. Caveolin-1 mediates Salmonella invasion via the regulation of SopE-dependent Rac1 activation and actin reorganization. *J Infect Dis,* 210**,** 793-802.

LIM, S. K., BYUN, J. R., NAM, H. M., LEE, H. S. & JUNG, S. C. 2011. Phenotypic and genotypic characterization of Salmonella spp. Isolated from pigs and their farm environment in Korea. *J Microbiol Biotechnol,* 21**,** 50-4.

LINDSTEDT, B. A., HEIR, E., GJERNES, E. & KAPPERUD, G. 2003. DNA Fingerprinting of Salmonella enterica subsp. enterica Serovar Typhimurium with Emphasis on Phage Type DT104 Based on Variable Number of Tandem Repeat Loci. *Journal of Clinical Microbiology,* 41**,** 1469-1479.

LINDSTEDT, B. A., TORPDAHL, M., VERGNAUD, G., LE HELLO, S., WEILL, F. X., TIETZE, E., MALORNY, B., PRENDERGAST, D. M., NI GHALLCHOIR, E., LISTA, R. F., SCHOULS, L. M., SODERLUND, R., BORJESSON, S. & AKERSTROM, S. 2013. Use of multilocus variable-number tandem repeat analysis (MLVA) in eight European countries, 2012. *Euro Surveill,* 18**,** 20385.

LO FO WONG, D. M., HALD, T., VAN DEL WOLF, P. J. & SWANENBURG, M. 2002. Epidemiology and control measures for Salmonella in pigs and pork. *Livestock Production Science*.

LOPEZ, C. A., WINTER, S. E., RIVERA-CHAVEZ, F., XAVIER, M. N., POON, V., NUCCIO, S. P., TSOLIS, R. M. & BAUMLER, A. J. 2012. Phage-mediated acquisition of a type III secreted effector protein boosts growth of salmonella by nitrate respiration. *MBio,* 3.

LUCARELLI, C., DIONISI, A. M., FILETICI, E., OWCZAREK, S., LUZZI, I. & VILLA, L. 2012. Nucleotide sequence of the chromosomal region conferring multidrug resistance (R-type ASSuT) in Salmonella Typhimurium and monophasic Salmonella Typhimurium strains. *J Antimicrob Chemother,* 67**,** 111-4.

LUCARELLI, C., DIONISI, A. M., TORPDAHL, M., VILLA, L., GRAZIANI, C., HOPKINS, K., THRELFALL, J., CAPRIOLI, A. & LUZZI, I. 2010. Evidence for a second genomic island conferring multidrug resistance in a clonal group of strains of Salmonella enterica serovar Typhimurium and its monophasic variant circulating in Italy, Denmark, and the United Kingdom. *J Clin Microbiol,* 48**,** 2103-9.

LYMAN, M. B., STEWARD, J. P. & ROANTREE, R. J. 1976. Characterization of the virulence and antigenic structure of Salmonella typhimurium strains with lipopolysaccharide core defects. *Infect Immun,* 13**,** 1539-42.

LYNCH, H., WALIA, K., LEONARD, F. C., LAWLOR, P. G., MANZANILLA, E. G., GRANT, J., DUFFY, G., GARDINER, G. E., CORMICAN, M., KING, J., MARKEY, B. K., FANNING, S. & ARGUELLO, H. 2018. Salmonella in breeding pigs:

Shedding pattern, transmission of infection and the role of environmental contamination in Irish commercial farrow-to-finish herds. *Zoonoses Public Health,* 65**,** e196-e206.

MACKENZIE, K. D., PALMER, M. B., KOSTER, W. L. & WHITE, A. P. 2017. Examining the Link between Biofilm Formation and the Ability of Pathogenic Salmonella Strains to Colonize Multiple Host Species. *Front Vet Sci,* 4**,** 138.

MAGIORAKOS, A. P., SRINIVASAN, A., CAREY, R. B., CARMELI, Y., FALAGAS, M. E., GISKE, C. G., HARBARTH, S., HINDLER, J. F., KAHLMETER, G., OLSSON-LILJEQUIST, B., PATERSON, D. L., RICE, L. B., STELLING, J., STRUELENS, M. J., VATOPOULOS, A., WEBER, J. T. & MONNET, D. L. 2012. Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin Microbiol Infect,* 18**,** 268-81.

MAINAR-JAIME, R. C., ATASHPARVAR, N., CHIRINO-TREJO, M. & RAHN, K. 2008. Survey on Salmonella prevalence in slaughter pigs from Saskatchewan. *Can Vet J*.

MAJOWICZ, S. E., MUSTO, J., SCALLAN, E., ANGULO, F. J., KIRK, M., O'BRIEN, S. J., JONES, T. F., FAZIL, A., HOEKSTRA, R. M. & INTERNATIONAL COLLABORATION ON ENTERIC DISEASE 'BURDEN OF ILLNESS, S. 2010. The global burden of nontyphoidal Salmonella gastroenteritis. *Clin Infect Dis,* 50**,** 882-9.

MALCOVA, M., HRADECKA, H., KARPISKOVA, R. & RYCHLIK, I. 2008. Biofilm formation in field strains of Salmonella enterica serovar Typhimurium: identification of a new colony morphology type and the role of SGI1 in biofilm formation. *Vet Microbiol,* 129**,** 360-6.

MALORNY, B., PACCASSONI, E., FACH, P., BUNGE, C., MARTIN, A. & HELMUTH, R. 2004. Diagnostic Real-Time PCR for Detection of Salmonella in Food. *Applied and Environmental Microbiology,* 70**,** 7046-7052.

MARTELLI, F., LAMBERT, M., BUTT, P., CHENEY, T., TATONE, F. A., CALLABY, R., RABIE, A., GOSLING, R. J., FORDON, S., CROCKER, G., DAVIES, R. H. & SMITH, R. P. 2017. Evaluation of an enhanced cleaning and disinfection protocol in Salmonella contaminated pig holdings in the United Kingdom. *PLoS One,* 12**,** e0178897.

MATHER, A. E., PHUONG, T. L. T., GAO, Y., CLARE, S., MUKHOPADHYAY, S., GOULDING, D. A., HOANG, N. T. D., TUYEN, H. T., LAN, N. P. H., THOMPSON, C. N., TRANG, N. H. T., CARRIQUE-MAS, J., TUE, N. T., CAMPBELL, J. I., RABAA, M. A., THANH, D. P., HARCOURT, K., HOA, N. T., TRUNG, N. V.,

SCHULTSZ, C., PERRON, G. G., COIA, J. E., BROWN, D. J., OKORO, C., PARKHILL, J., THOMSON, N. R., CHAU, N. V. V., THWAITES, G. E., MASKELL, D. J., DOUGAN, G., KENNEY, L. J. & BAKER, S. 2018. New Variant of Multidrug-Resistant Salmonella enterica Serovar Typhimurium Associated with Invasive Disease in Immunocompromised Patients in Vietnam. *MBio,* 9.

MATHER, A. E., REID, S. W. J., MASKELL, D. J., PARKHILL, J., FOOKES, M. C., HARRIS, S. R., BROWN, D. J., COIA, J. E., MULVEY, M. R., GILMOUR, M. W., PETROVSKA, L., DE PINNA, E., KURODA, M., AKIBA, M., IZUMIYA, H., CONNOR, T. R., SUCHARD, M. A., LEMEY, P., MELLOR, D. J., HAYDON, D. T. & THOMSON, N. R. 2013. Distinguishable Epidemics of Multidrug-Resistant Salmonella Typhimurium DT104 in Different Hosts. *Science,* 341**,** 1514-1517.

MCGOVERN, V. J. & SLAVUTIN, L. J. 1979. Pathology of salmonella colitis. *Am J Surg Pathol,* 3**,** 483-90.

MCQUISTON, J. R., FIELDS, P. I., TAUXE, R. V. & LOGSDON, J. M., JR. 2008. Do Salmonella carry spare tyres? *Trends Microbiol,* 16**,** 142-8.

MEIER-KOLTHOFF, J. P. & GOKER, M. 2017. VICTOR: genome-based phylogeny and classification of prokaryotic viruses. *Bioinformatics,* 33**,** 3396-3404.

MERRITT, J. H., KADOURI, D. E. & O'TOOLE, G. A. 2005. Growing and analyzing static biofilms. *Current protocols in microbiology,* Chapter 1**,** Unit-1B.1.

MILLER, A. J., TWOMEY, D. F., DAVIES, R. H., TEALE, C. J., WILLIAMSON, S. M., REICHEL, R., FEATHERSTONE, C. A., COOK, A. J., SNOW, L. C. & ARMSTRONG, J. D. 2011. Salmonella serovars and antimicrobial resistance patterns on a sample of high seroprevalence pig farms in England and Wales (2003-2008). *Zoonoses Public Health,* 58**,** 549-59.

MINTZ, E. D., CARTTER, M. L., HADLER, J. L., WASSELL, J. T., ZINGESER, J. A. & TAUXE, R. V. 1994. Dose-response effects in an outbreak of Salmonella enteritidis. *Epidemiol Infect,* 112**,** 13-23.

MIROLD, S., EHRBAR, K., WEISSMULLER, A., PRAGER, R., TSCHAPE, H., RUSSMANN, H. & HARDT, W. D. 2001a. Salmonella Host Cell Invasion Emerged by Acquisition of a Mosaic of Separate Genetic Elements, Including Salmonella Pathogenicity Island 1 (SPI1), SPI5, and sopE2. *J Bacteriol,* 183**,** 2348-2358.

MIROLD, S., RABSCH, W., ROHDE, M., STENDER, S., TSCHAPE, H., RUSSMANN, H., IGWE, E. & HARDT, W. D. 1999. Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic Salmonella typhimurium strain. *Proc Natl Acad Sci U S A,* 96**,** 9845-50.

MIROLD, S., RABSCH, W., TSCHAPE, H. & HARDT, W. D. 2001b. Transfer of the Salmonella type III effector sopE between unrelated phage families. *J Mol Biol,* 312**,** 7-16.

MOURA, A., SOARES, M., PEREIRA, C., LEITAO, N., HENRIQUES, I. & CORREIA, A. 2009. INTEGRALL: a database and search engine for integrons, integrases and gene cassettes. *Bioinformatics,* 25**,** 1096-8.

MOURAO, J., MACHADO, J., NOVAIS, C., ANTUNES, P. & PEIXE, L. 2014. Characterization of the emerging clinically-relevant multidrug-resistant Salmonella enterica serotype 4,[5],12:i:- (monophasic variant of S. Typhimurium) clones. *Eur J Clin Microbiol Infect Dis,* 33**,** 2249-57.

MUELLER-DOBLIES, D., SPEED, K. C. R., KIDD, S. & DAVIES, R. H. 2018. Salmonella Typhimurium in livestock in Great Britain - trends observed over a 32-year period. *Epidemiol Infect,* 146**,** 409-422.

MULLER, A. J., HOFFMANN, C., GALLE, M., VAN DEN BROEKE, A., HEIKENWALDER, M., FALTER, L., MISSELWITZ, B., KREMER, M., BEYAERT, R. & HARDT, W. D. 2009. The S. Typhimurium effector SopE induces caspase-1 activation in stromal cells to initiate gut inflammation. *Cell Host Microbe,* 6**,** 125-36.

MULVEY, M. R., BOYD, D. A., OLSON, A. B., DOUBLET, B. & CLOECKAERT, A. 2006. The genetics of Salmonella genomic island 1. *Microbes and Infection,* 8**,** 1915-1922.

NICHOLSON, F. A., CHAMBERS, B. J., WILLIAMS, J. R. & UNWIN, R. J. 1999. Heavy metal contents of livestock feeds and animal manures in England and Wales. *Bioresource Technology,* 70**,** 23-31.

NOBLE, D. J., LANE, C., LITTLE, C. L., DAVIES, R., DE PINNA, E., LARKIN, L. & MORGAN, D. 2012. Revival of an old problem: an increase in Salmonella enterica serovar Typhimurium definitive phage type 8 infections in 2010 in England and Northern Ireland linked to duck eggs. *Epidemiol Infect,* 140**,** 146-9.

O'LEARY, D., CABE, E. M., MCCUSKER, M. P., MARTINS, M., FANNING, S. & DUFFY, G. 2013. Microbiological study of biofilm formation in isolates of Salmonella enterica Typhimurium DT104 and DT104b cultured from the modern pork chain. *Int J Food Microbiol,* 161**,** 36-43.

OBENG, N., PRATAMA, A. A. & ELSAS, J. D. V. 2016. The Significance of Mutualistic Phages for Bacterial Ecology and Evolution. *Trends Microbiol,* 24**,** 440-449.

OCHMAN, H., SONCINI, F. C., SOLOMON, F. & GROISMAN, E. A. 1996. Identification of a pathogenicity island required for Salmonella survival in host cells. *Proc Natl Acad Sci U S A,* 93**,** 7800-4.

OKONECHNIKOV, K., GOLOSOVA, O., FURSOV, M. & TEAM, U. 2012. Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics,* 28**,** 1166-7.

OKORO, C. K., BARQUIST, L., CONNOR, T. R., HARRIS, S. R., CLARE, S., STEVENS, M. P., ARENDS, M. J., HALE, C., KANE, L., PICKARD, D. J., HILL, J., HARCOURT, K., PARKHILL, J., DOUGAN, G. & KINGSLEY, R. A. 2015. Signatures of adaptation in human invasive Salmonella Typhimurium ST313 populations from sub-Saharan Africa. *PLoS Negl Trop Dis,* 9**,** e0003611.

OKORO, C. K., KINGSLEY, R. A., QUAIL, M. A., KANKWATIRA, A. M., FEASEY, N. A., PARKHILL, J., DOUGAN, G. & GORDON, M. A. 2012. High-resolution single nucleotide polymorphism analysis distinguishes recrudescence and reinfection in recurrent invasive nontyphoidal Salmonella typhimurium disease. *Clin Infect Dis,* 54**,** 955-63.

PALMA, F., MANFREDA, G., SILVA, M., PARISI, A., BARKER, D. O. R., TABOADA, E. N., PASQUALI, F. & ROSSI, M. 2018. Genome-wide identification of geographical segregated genetic markers in Salmonella enterica serovar Typhimurium variant 4,[5],12:i. *Sci Rep,* 8**,** 15251.

PARADIS, E. & SCHLIEP, K. 2018. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics,* 35**,** 526-528.

PARRY, C. M., HIEN, T. T., DOUGAN, G., WHITE, N. J. & FARRAR, J. J. 2002. Typhoid fever. *N Engl J Med,* 347**,** 1770-82.

PATHMANATHAN, S. G., CARDONA-CASTRO, N., SANCHEZ-JIMENEZ, M. M., CORREA-OCHOA, M. M., PUTHUCHEARY, S. D. & THONG, K. L. 2003. Simple and rapid detection of Salmonella strains by direct PCR amplification of the hilA gene. *J Med Microbiol,* 52**,** 773-6.

PELLUDAT, C., MIROLD, S. & HARDT, W. D. 2003. The SopEPhi phage integrates into the ssrA gene of Salmonella enterica serovar Typhimurium A36 and is closely related to the Fels-2 prophage. *Journal of bacteriology,* 185**,** 5182-5191.

PENNYCOTT, T. W., PARK, A. & MATHER, H. A. 2006. Isolation of different serovars of Salmonella enterica from wild birds in Great Britain between 1995 and 2003. *Vet Rec,* 158**,** 817-20.

PETROVSKA, L., MATHER, A. E., ABUOUN, M., BRANCHU, P., HARRIS, S. R., CONNOR, T., HOPKINS, K. L., UNDERWOOD, A., LETTINI, A. A., PAGE, A., BAGNALL, M., WAIN, J., PARKHILL, J., DOUGAN, G., DAVIES, R. & KINGSLEY, R. A. 2016. Microevolution of Monophasic Salmonella Typhimurium during Epidemic, United Kingdom, 2005-2010. *Emerg Infect Dis,* 22**,** 617-24.

PEYRU, G. & FRAENKEL, D. G. 1968. Genetic mapping of loci for glucose-6-phosphate dehydrogenase, gluconate-6-phosphate dehydrogenase, and gluconate-6-phosphate dehydrase in Escherichia coli. *J Bacteriol,* 95**,** 1272-8.

PHAN, M. D. & WAIN, J. 2008. IncHI plasmids, a dynamic link between resistance and pathogenicity. *J Infect Dev Ctries,* 2**,** 272-8.

PHE 2018. Zoonoses report UK 2017.

PILAR, A. V., REID-YU, S. A., COOPER, C. A., MULDER, D. T. & COOMBES, B. K. 2012. GogB is an anti-inflammatory effector that limits tissue damage during Salmonella infection through interaction with human FBXO22 and Skp1. *PLoS Pathog,* 8**,** e1002773.

POON, A. F. 2016. Impacts and shortcomings of genetic clustering methods for infectious disease outbreaks. *Virus Evol,* 2**,** vew031.

PORWOLLIK, S., WONG, R. & MCCLELLAND, M. 2002. Evolutionary genomics of Salmonella: Gene acquisitions revealed by microarray analysis. *Proceedings of the National Academy of Sciences,* 99**,** 8956-8961.

R CORE TEAM 2019. R: A language and environment for statistical computing. *In:* COMPUTING, R. F. F. S. (ed.).

RABSCH, W. 2007. Salmonella Typhimurium Phage Typing for Pathogens. *In:* SCHATTEN, H. & EISENSTARK, A. (eds.) *Salmonella: Methods and Protocols.* Totowa, NJ: Humana Press.

RABSCH, W., SIMON, S. & HUMPHREY, T. 2013. *Public Health Aspects of Salmonella Infections*.

RABSCH, W., TRUEPSCHUCH, S., WINDHORST, D. & GERLACH, R. G. 2011. *Typing phages and prophages of Salmonella,* Norfolk, UK, Caister Academic Press.

RABSCH, W., TSCHAPE, H. & BAUMLER, A. J. 2001. Non-typhoidal salmonellosis: emerging problems. *Microbes Infect,* 3**,** 237-47.

RAFFATELLU, M., WILSON, R. P., WINTER, S. E. & BAUMLER, A. J. 2008. Clinical pathogenesis of typhoid fever. *J Infect Dev Ctries,* 2**,** 260-6.

RAKHUBA, D. V., KOLOMIETS, E. I., SZWAICER DERY, E. & NOVIK, G. I. 2010. Bacteriophage Receptors, Mechanisms of Phage Adsorption and Penetration into Host Cell. *Polish Journal of Microbiology*.

RAYA, R. L. R. & H'BERT, E. M. 2009. Isolation of Phage via Induction of Lysogens. *In:* CLOKIE, M. R. J. & KROPINSKI, A. M. (eds.) *Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions.* Totowa, NJ: Humana Press.

ROBERTS, F., ALLISON, G. E. & VERMA, N. K. 2007. Transcription-termination-mediated immunity and its prevention in bacteriophage SfV of Shigella flexneri. *J Gen Virol,* 88**,** 3187-97.

ROBERTSON, J., YOSHIDA, C., GURNIK, S. & NASH, J. H. E. 2018. Completed Genome Sequences of Strains from 36 Serotypes of Salmonella. *Genome Announc,* 6.

ROMLING, U. 2005. Characterization of the rdar morphotype, a multicellular behaviour in Enterobacteriaceae. *Cell Mol Life Sci,* 62**,** 1234-46.

ROMLING, U., BOKRANZ, W., RABSCH, W., ZOGAJ, X., NIMTZ, M. & TSCHAPE, H. 2003. Occurrence and regulation of the multicellular morphotype in Salmonella serovars important in human disease. *Int J Med Microbiol,* 293**,** 273-85.

ROMLING, U. & ROHDE, M. 1999. Flagella modulate the multicellular behavior of Salmonella typhimurium on the community level. *Fems Microbiology Letters*.

ROMLING, U., SIERRALTA, W. D., ERIKSSON, K. & NORMARK, S. 1998. Multicellular and aggregative behaviour of Salmonella typhimurium strains is controlled by mutations in the agfDpromoter. *Molecular Microbiology*.

RONQUIST, F., TESLENKO, M., VAN DER MARK, P., AYRES, D. L., DARLING, A., HÖHNA, S., LARGET, B., LIU, L., SUCHARD, M. A. & HUELSENBECK, J. P. 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol,* 61**,** 539-42.

ROSTAGNO, M. H. & CALLAWAY, T. R. 2012. Pre-harvest risk factors for Salmonella enterica in pork production. *Food Research International,* 45**,** 634-640.

ROWE, T. A., LEONARD, F. C., KELLY, G., LYNCH, P. B., EGAN, J., QUIRKE, A. M. & QUINN, P. J. 2003. Salmonella serotypes present on a sample of Irish pig farms. *Vet Rec,* 153**,** 453-6.

RUTHERFORD, K., PARKHILL, J., CROOK, J., HORSNELL, T., RICE, P., RAJANDREAM, M. A. & BARRELL, B. 2000. Artemis: sequence visualization and annotation. *Bioinformatics (Oxford, England),* 16**,** 944-945.

SAN ROMAN, B., GARRIDO, V., SANCHEZ, S., MARTINEZ-BALLESTEROS, I., GARAIZAR, J., MAINAR-JAIME, R. C., MIGURA-GARCIA, L. & GRILLO, M. J. 2018. Relationship between Salmonella infection, shedding and serology in fattening pigs in low-moderate prevalence areas. *Zoonoses Public Health,* 65**,** 481-489.

SCHER, K., ROMLING, U. & YARON, S. 2005. Effect of heat, acidification, and chlorination on Salmonella enterica serovar typhimurium cells in a biofilm formed at the air-liquid interface. *Appl Environ Microbiol,* 71**,** 1163-8.

SCHLUMBERGER, M. C., FRIEBEL, A., BUCHWALD, G., SCHEFFZEK, K., WITTINGHOFER, A. & HARDT, W. D. 2003. Amino acids of the bacterial toxin SopE involved in G nucleotide exchange on Cdc42. *J Biol Chem,* 278**,** 27149-59.

SCHREIBER, F., KAY, S., FRANKEL, G., CLARE, S., GOULDING, D., VAN DE VOSSE, E., VAN DISSEL, J. T., STRUGNELL, R., THWAITES, G., KINGSLEY, R. A., DOUGAN, G. & BAKER, S. 2015. The Hd, Hj, and Hz66 flagella variants of Salmonella enterica serovar Typhi modify host responses and cellular interactions. *Sci Rep,* 5**,** 7947.

SCHULTE, M. & HENSEL, M. 2016. Models of intestinal infection by Salmonella enterica: introduction of a new neonate mouse model. *F1000Research,* 5**,** F1000 Faculty Rev-1498.

SEEMANN, T. 2015. snippy: fast bacterial variant calling from NGS reads. *https://github.com/tseemann/snippy*.

SEEMANN, T. 2017. snp-dists. *https://github.com/tseemann/snp-dists*.

SEIXAS, R., MACHADO, J., BERNARDO, F., VILELA, C. & OLIVEIRA, M. 2014. Biofilm formation by Salmonella enterica serovar 1,4,[5],12:i:- Portuguese isolates: a phenotypic, genotypic, and socio-geographic analysis. *Curr Microbiol,* 68**,** 670-7.

SIGUIER, P., PEROCHON, J., LESTRADE, L., MAHILLON, J. & CHANDLER, M. 2006. ISfinder: the reference centre for bacterial insertion sequences. *Nucleic acids research,* 34**,** D32-6.

SIIRA, L., MACDONALD, E., HOLMBAKKEN, G. M., SUNDAR, T., MEYER-MYKLESTAD, L., LANGE, H., BRANDAL, L. T., NASEER, U., JOHANNESSEN, G. S., BERGSJØ, B., ESPENHAIN, L., VOLD, L. & NYGÅRD, K. 2019. Increasing incubation periods during a prolonged monophasic Salmonella Typhimurium outbreak with environmental contamination of a commercial kitchen at Oslo Airport, Norway, 2017. *Eurosurveillance*.

SIMM, R., AHMAD, I., RHEN, M., LE GUYON, S. & ROMLING, U. 2014. Regulation of biofilm formation in Salmonella enterica serovar Typhimurium. *Future Microbiol*.

SIRINAVIN, S. & GARNER, P. 2000. Antibiotics for treating salmonella gut infections. *Cochrane Database Syst Rev***,** Cd001167.

SOLOVYEV, V. & SALAMOV, A. 2010. Automatic Annotation of Microbial Genomes and Metagenomic Sequences. In Metagenomics and its Applications in Agriculture. *Biomedicine and Environmental Studies, Nova Science Publishers (Ed. R.W. Li)***,** p. 61-78.

SOYER, Y., MORENO SWITT, A., DAVIS, M. A., MAURER, J., MCDONOUGH, P. L., SCHOONMAKER-BOPP, D. J., DUMAS, N. B., ROOT, T., WARNICK, L. D., GROHN, Y. T. & WIEDMANN, M. 2009. Salmonella enterica serotype 4,5,12:i:-,

an emerging Salmonella serotype that represents multiple distinct clones. *J Clin Microbiol,* 47**,** 3546-56.

STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22**,** 2688-2690.

STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics,* 30**,** 1312-3.

STEENACKERS, H., HERMANS, K., VANDERLEYDEN, J. & DE KEERSMAECKER, S. C. J. 2012. Salmonella biofilms: An overview on occurrence, structure, regulation and eradication. *Food Research International,* 45**,** 502-531.

STOCKER, B. A. D. 1949. Meassurement of the rate of mutation of flagellar antigenic phase in *Salmonella typhimurium*. *J. Hyg.,* 47**,** 398-413.

STOCKI, S. L., ANNETT, C. B., SIBLEY, C. D., MCLAWS, M., CHECKLEY, S. L., SINGH, N., SURETTE, M. G. & WHITE, A. P. 2007. Persistence of Salmonella on egg conveyor belts is dependent on the belt type but not on the rdar morphotype. *Poult Sci,* 86**,** 2375-83.

STOJILJKOVIC, I., BÄUMLER, A. J. & HEFFRON, F. 1995. Ethanolamine utilization in *Salmonella typhimurium:* nucleotide sequence, protein expression and mutational analysis of the *cchA cchB eutE eutJ eutG eutH* gene cluster. *J. Bacteriol.,* 177**,** 1357-1366.

SUN, J., KE, B., HUANG, Y., HE, D., LI, X., LIANG, Z. & KE, C. 2014. The Molecular Epidemiological Characteristics and Genetic Diversity of Salmonella Typhimurium in Guangdong, China, 2007–2011. *PLoS ONE,* 9.

SWITT, A. I. M., SULAKVELIDZE, A., WIEDMANN, M., KROPINSKI, A. M., WISHART, D. S., POPPE, C. & LIANG, Y. 2015. Salmonella Phages and Prophages: Genomics, Taxonomy, and Applied Aspects. *In:* SCHATTEN, H. & EISENSTARK, A. (eds.) *Salmonella: Methods and Protocols.* New York, NY: Springer New York.

TAVARÉ, S. & MIURA, R. M. 1986. Some probabilistic and statistical problems in the analysis of DNA. *Lectures on mathematics in the life sciences,* 17.

TAYLOR, D. E. & SUMMERS, A. O. 1979. Association of Tellurium Resistance and Bacteriophage Inhibition Conferred by R Plasmids. *Journal of Bacteriology,* Vol. 137**,** p. 1430-1433.

THRELFALL, E. J. 2000. Epidemic Salmonella typhimurium DT 104 - a truly international multiresistant clone. *Journal of Antimicrobial Chemotherapy,* 46**,** 7-10.

TRUPSCHUCH, S., LAVERDE GOMEZ, J. A., EDIBERIDZE, I., FLIEGER, A. & RABSCH, W. 2010. Characterisation of multidrug-resistant Salmonella

Typhimurium 4,[5],12:i:- DT193 strains carrying a novel genomic island adjacent to the thrW tRNA locus. *Int J Med Microbiol,* 300**,** 279-88.

UZZAU, S., BOSSI, L. & FIGUEROA-BOSSI, N. 2002. Differential accumulation of Salmonella[Cu, Zn] superoxide dismutases SodCI and SodCII in intracellular bacteria: correlation with their relative contribution to pathogenicity. *Mol Microbiol,* 46**,** 147-56.

VEGGE, C. S., JANSEN VAN RENSBURG, M. J., RASMUSSEN, J. J., MAIDEN, M. C., JOHNSEN, L. G., DANIELSEN, M., MACINTYRE, S., INGMER, H. & KELLY, D. J. 2016. Glucose Metabolism via the Entner-Doudoroff Pathway in Campylobacter: A Rare Trait that Enhances Survival and Promotes Biofilm Formation in Some Isolates. *Front Microbiol,* 7**,** 1877.

VESTBY, L. K., MORETRO, T., LANGSRUD, S., HEIR, E. & NESSE, L. L. 2009. Biofilm forming abilities of Salmonella are correlated with persistence in fish meal- and feed factories. *BMC Vet Res,* 5**,** 20.

VISSCHER, C. F., KLEIN, G., VERSPOHL, J., BEYERBACH, M., STRATMANN-SELKE, J. & KAMPHUES, J. 2011. Serodiversity and serological as well as cultural distribution of Salmonella on farms and in abattoirs in Lower Saxony, Germany. *Int J Food Microbiol,* 146**,** 44-51.

VON BUBNOFF, A. 2008. Next-Generation Sequencing: The Race Is On. *Cell,* 132**,** 721-723.

VONAESCH, P., SELLIN, M. E., CARDINI, S., SINGH, V., BARTHEL, M. & HARDT, W. D. 2014. The Salmonella Typhimurium effector protein SopE transiently localizes to the early SCV and contributes to intracellular replication. *Cell Microbiol,* 16**,** 1723-35.

WAHL, A., BATTESTI, A. & ANSALDI, M. 2019. Prophages in Salmonella enterica: a driving force in reshaping the genome and physiology of their bacterial host? *Mol Microbiol,* 111**,** 303-316.

WALIA, K., ARGUELLO, H., LYNCH, H., GRANT, J., LEONARD, F. C., LAWLOR, P. G., GARDINER, G. E. & DUFFY, G. 2017. The efficacy of different cleaning and disinfection procedures to reduce Salmonella and Enterobacteriaceae in the lairage environment of a pig abattoir. *Int J Food Microbiol,* 246**,** 64-71.

WANG, Y., ZHANG, A., YANG, Y., LEI, C., JIANG, W., LIU, B., SHI, H., KONG, L., CHENG, G., ZHANG, X., YANG, X. & WANG, H. 2017. Emergence of Salmonella enterica serovar Indiana and California isolates with concurrent resistance to cefotaxime, amikacin and ciprofloxacin from chickens in China. *Int J Food Microbiol,* 262**,** 23-30.

WEAVER, T., VALCANIS, M., MERCOULIA, K., SAIT, M., TUKE, J., KIERMEIER, A., HOGG, G., POINTON, A., HAMILTON, D. & BILLMAN-JACOBE, H. 2017. Longitudinal study of Salmonella 1,4,[5],12:i:- shedding in five Australian pig herds. *Prev Vet Med,* 136**,** 19-28.

WHITE, A. P., GIBSON, D. L., KIM, W., KAY, W. W. & SURETTE, M. G. 2006. Thin aggregative fimbriae and cellulose enhance long-term survival and persistence of Salmonella. *J Bacteriol,* 188**,** 3219-27.

WHITE, A. P. & SURETTE, M. G. 2006. Comparative genetics of the rdar morphotype in Salmonella. *J Bacteriol,* 188**,** 8395-406.

WHITE, A. P., WELJIE, A. M., APEL, D., ZHANG, P., SHAYKHUTDINOV, R., VOGEL, H. J. & SURETTE, M. G. 2010. A global metabolic shift is linked to Salmonella multicellular development. *PLoS One,* 5**,** e11814.

WHO 2018. Whole genome sequencing for foodborne disease surveillance.

WHYTE, P., MC GILL, K., COLLINS, J. D. & GORMLEY, E. 2002. The prevalence and PCR detection of Salmonella contamination in raw poultry. *Vet Microbiol,* 89**,** 53-60.

WINFIELD, M. & GROISMAN, E. 2003a. Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. *Applied and Environmental Microbiology,* 69**,** 3687-3694.

WINFIELD, M. D. & GROISMAN, E. A. 2003b. Role of Nonhost Environments in the Lifestyles of Salmonella and Escherichia coli. *Applied and Environmental Microbiology,* 69**,** 3687-3694.

WINTER, S. E., THIENNIMITR, P., WINTER, M. G., BUTLER, B. P., HUSEBY, D. L., CRAWFORD, R. W., RUSSELL, J. M., BEVINS, C. L., ADAMS, L. G., TSOLIS, R. M., ROTH, J. R. & BAUMLER, A. J. 2010a. Gut inflammation provides a respiratory electron acceptor for Salmonella. *Nature,* 467**,** 426-9.

WINTER, S. E., WINTER, M. G., GODINEZ, I., YANG, H. J., RUSSMANN, H., ANDREWS-POLYMENIS, H. L. & BAUMLER, A. J. 2010b. A rapid change in virulence gene expression during the transition from the intestinal lumen into tissue promotes systemic dissemination of Salmonella. *PLoS Pathog,* 6.

WONG, M. H. Y., YAN, M., CHAN, E. W. C., LIU, L. Z., KAN, B. & CHEN, S. 2013. Expansion of Salmonella enterica Serovar Typhimurium ST34 Clone Carrying Multiple Resistance Determinants in China. *Antimicrobial Agents and Chemotherapy,* 57**,** 4599-4601.

XIE, X., WANG, Z., ZHANG, K., LI, Y., HU, Y., PAN, Z., CHEN, X., LI, Q. & JIAO, X. 2020. Pig as a reservoir of CRISPR type TST4 Salmonella enterica serovar

Typhimurium monophasic variant during 2009-2017 in China. *Emerg Microbes Infect,* 9**,** 1-4.

YAMAMOTO, S. & KUTSUKAKE, K. 2006. FljA-Mediated Posttranscriptional Control of Phase 1 Flagellin Expression in Flagellar Phase Variation of Salmonella enterica Serovar Typhimurium. *Journal of Bacteriology,* 188**,** 958.

YANG, Z. 1993. Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol Biol Evol,* 10**,** 1396-401.

YANG, Z. & RANNALA, B. 2012. Molecular phylogenetics: principles and practice. *Nat Rev Genet,* 13**,** 303-14.

YOON, H., ANSONG, C., ADKINS, J. N. & HEFFRON, F. 2011. Discovery of Salmonella virulence factors translocated via outer membrane vesicles to murine macrophages. *Infect Immun,* 79**,** 2182-92.

YU, G., SMITH, D. K., ZHU, H., GUAN, Y., LAM, T. T.-Y. & MCINERNY, G. 2017. ggtree: anrpackage for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution,* 8**,** 28-36.

ZANKARI, E., HASMAN, H., COSENTINO, S., VESTERGAARD, M., RASMUSSEN, S., LUND, O., AARESTRUP, F. M. & LARSEN, M. V. 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother,* 67**,** 2640-4.

ZHANG, S., KINGSLEY, R. A., SANTOS, R. L., ANDREWS-POLYMENIS, H., RAFFATELLU, M., FIGUEIREDO, J., NUNES, J., TSOLIS, R. M., ADAMS, L. G. & BAUMLER, A. J. 2003. Molecular pathogenesis of Salmonella enterica serotype typhimurium-induced diarrhea. *Infect Immun,* 71**,** 1-12.

ZHANG, S., YIN, Y., JONES, M. B., ZHANG, Z., DEATHERAGE KAISER, B. L., DINSMORE, B. A., FITZGERALD, C., FIELDS, P. I. & DENG, X. 2015. Salmonella serotype determination utilizing high-throughput genome sequencing data. *J Clin Microbiol,* 53**,** 1685-92.

ZHANG, X. L., TSUI, I. S., YIP, C. M., FUNG, A. W., WONG, D. K., DAI, X., YANG, Y., HACKETT, J. & MORRIS, C. 2000. Salmonella enterica serovar typhi uses type IVB pili to enter human intestinal epithelial cells. *Infect Immun,* 68**,** 3067-73.

ZHOU, X., SHEN, X.-X., HITTINGER, C. T. & ROKAS, A. 2017. Evaluating Fast Maximum Likelihood-Based Phylogenetic Programs Using Empirical Phylogenomic Data Sets. *Molecular Biology and Evolution,* 35**,** 486-503.

Appendices

Appendix 1 *S*. Typhimurium and *S*. 4,[5],12:i:- isolates from Irish pig farms used for the study in Chapter 3. The strain name, the serotype, the farm of isolation, the stage of production, the sample type, the date of isolation and accession number are reported. S.Tm: S. Typhimurium

Strain	Serotype	Farm	Stage of production	Sample type	Date	Accession number
0153D	4,12:i:-	A	1st Stage Weaners	faecal	26/03/2012	SAMN08734271
0160A	4,12:i:-	A	1st Stage Weaners	water	26/03/2012	SAMN08734326
0167D	4,12:i:-	A	1st Stage Weaners	water drinker	26/03/2012	SAMN08734331
0171B	4,12:i:-	A	1st Stage Weaners	faecal	26/03/2012	SAMN08734338
0179B	4,12:i:-	A	2nd Stage Weaners	faecal	26/03/2012	SAMN08734350
0180B	4,12:i:-	A	2nd Stage Weaners	water	26/03/2012	SAMN08734360
0181A	4,12:i:-	A	2nd Stage Weaners	water drinker	26/03/2012	SAMN08734302
0182B	4,12:i:-	A	2nd Stage Weaners	feed	26/03/2012	SAMN08734366
0185A	4,12:i:-	A	2nd Stage Weaners	faecal	26/03/2012	SAMN08734380
0187A	4,12:i:-	A	2nd Stage Weaners	water drinker	26/03/2012	SAMN08734392
0191A	4,12:i:-	A	2nd Stage Weaners	faecal	26/03/2012	SAMN08734274
0197B	4,12:i:-	A	2nd Stage Weaners	faecal	26/03/2012	SAMN08734282
0211A	4,12:i:-	A	Finishers	faecal	26/03/2012	SAMN08734293
0216A	4,12:i:-	A	Finishers	pen	26/03/2012	SAMN08734275
0217A	4,5,12:i:-	A	Finishers	pen	26/03/2012	SAMN08734295
0229A	4,12:i:-	A	Finishers	faecal	26/03/2012	SAMN08734297
0234B	4,5,12:i:-	A	Finishers	pen	26/03/2012	SAMN08734306

0248A	4,12:i:-	A	Gilts	pen	26/03/2012	SAMN08734276
0286B	4,12:i:-	A	Farrowing	pen	26/03/2012	SAMN08734277
0287A	4,5,12:i:-	A	Farrowing feed bin		26/03/2012	SAMN08734307
0309C	4,12:i:-	B	1st Stage Weaners	faecal	16/04/2012	SAMN08734286
0314A	4,12:i:-	B	1st Stage Weaners	pen	16/04/2012	SAMN08734322
0315D	4,12:i:-	B	1st Stage Weaners	faecal	16/04/2012	SAMN08734308
0316E	4,12:i:-	B	1st Stage Weaners	water	16/04/2012	SAMN08734323
0322D	4,12:i:-	B	1st Stage Weaners	water drinker	16/04/2012	SAMN08734327
0323A	4,12:i:-	B	1st Stage Weaners	water drinker	16/04/2012	SAMN08734310
0327B	4,12:i:-	B	1st Stage Weaners	faecal	16/04/2012	SAMN08734311
0329B	4,12:i:-	B	1st Stage Weaners	water drinker	16/04/2012	SAMN08734328
0335B	4,12:i:-	B	2nd Stage Weaners	faecal	16/04/2012	SAMN08734312
0337D	4,12:i:-	B	2nd Stage Weaners	water drinker	16/04/2012	SAMN08734278
0342E	4,12:i:-	B	2nd Stage Weaners	water	16/04/2012	SAMN08734314
0343A	4,12:i:-	B	2nd Stage Weaners	water drinker	16/04/2012	SAMN08734292
0364A	4,12:i:-	B	Finishers	feed	16/04/2012	SAMN08734329
0384C	4,5,12:i:-	B	Finishers	pen	16/04/2012	SAMN08734317
0591C	4,5,12:i:-	D	Dry sow	water drinker	21/05/2012	SAMN08734318
0598B	*S.* Tm	D	1st Stage Weaners	water	21/05/2012	SAMN08734324
0660A	*S.* Tm	D	Finishers	pen	21/05/2012	SAMN08734384
0678A	*S.* Tm	D	Gilts	feed	21/05/2012	SAMN08734313
0694E	*S.* Tm	D	Farrowing	trough	21/05/2012	SAMN08734325

0911A	*S.* Tm	F	2nd Stage Weaners	pen	02/07/2012	SAMN08734386
1048A	*S.* Tm	G	2nd Stage Weaners	pen	16/07/2012	SAMN08734387
1054D	*S.* Tm	G	2nd Stage Weaners	pen	16/07/2012	SAMN08734388
1060C	*S.* Tm	G	2nd Stage Weaners	pen	16/07/2012	SAMN08734372
1067B	*S.* Tm	G	Finishers	trough	16/07/2012	SAMN08734404
1089D	*S.* Tm	G	Gilts	faecal	16/07/2012	SAMN08734391
1098A	*S.* Tm	G	Gilts	feed	16/07/2012	SAMN08734374
1106A	*S.* Tm	G	Gilts	pen	16/07/2012	SAMN08734395
1344A	*S.* Tm	I	Finishers	pen	14/08/2012	SAMN08734330
1446A	4,12:i:-	J	1st Stage Weaners	faecal	03/09/2012	SAMN08734333
1448A	4,12:i:-	J	1st Stage Weaners	water drinker	03/09/2012	SAMN08734334
1465A	4,12:i:-	J	1st Stage Weaners	water	03/09/2012	SAMN08734335
1468A	4,12:i:-	J	1st Stage Weaners	trough	03/09/2012	SAMN08734279
1469C	4,12:i:-	J	1st Stage Weaners	pen	03/09/2012	SAMN08734336
1495C	4,12:i:-	J	2nd Stage Weaners	pen	03/09/2012	SAMN08734337
1728A	4,12:i:-	A	1st Stage Weaners	faecal	19/11/2012	SAMN09435861
1733C	4,12:i:-	A	1st Stage Weaners	pen	19/11/2012	SAMN08734339
1740A	4,12:i:-	A	1st Stage Weaners	faecal	19/11/2012	SAMN08734340
1746A	4,12:i:-	A	1st Stage Weaners	faecal	19/11/2012	SAMN08734341
1754C	4,12:i:-	A	2nd Stage Weaners	faecal	19/11/2012	SAMN08734342
1757A	4,12:i:-	A	2nd Stage Weaners	feed	19/11/2012	SAMN08734343

1758A	4,5,12:i:-	A	2nd Stage Weaners	Trough	19/11/2012	SAMN09435862
1759B	4,5,12:i:-	A	2nd Stage Weaners	pen	19/11/2012	SAMN08734344
1760A	4,12:i:-	A	2nd Stage Weaners	faecal	19/11/2012	SAMN08734345
1764C	4,12:i:-	A	2nd Stage Weaners	trough	19/11/2012	SAMN08734283
1765A	4,5,12:i:-	A	2nd Stage Weaners	pen	19/11/2012	SAMN08734285
1769B	4,12:i:-	A	2nd Stage Weaners	feed	19/11/2012	SAMN08734347
1770A	4,12:i:-	A	2nd Stage Weaners	trough	19/11/2012	SAMN08734348
1772B	4,12:i:-	A	2nd Stage Weaners	faecal	19/11/2012	SAMN08734349
1774A	4,5,12:i:-	A	2nd Stage Weaners	water drinker	19/11/2012	SAMN08734287
1777A	4,5,12:i:-	A	2nd Stage Weaners	pen	19/11/2012	SAMN08734303
1792A	4,12:i:-	A	Finishers	pen	19/11/2012	SAMN08734288
1870C	4,5,12:i:-	A	Finishers pen 6	water drinker	19/11/2012	SAMN08734352
1873A	4,5,12:i:-	A	Finishers pen 6	pen	19/11/2012	SAMN08734289
1882A	4,12:i:-	A	Finishers pen 8	water drinker	19/11/2012	SAMN08734353
1885A	4,12:i:-	A	Finishers pen 8	pen	19/11/2012	SAMN08734290
1939A	4,12:i:-	B	1st Stage Weaners	faecal	03/12/2012	SAMN08734346
1943A	4,12:i:-	B	1st Stage Weaners	trough	03/12/2012	SAMN08734304
1948B	4,12:i:-	B	1st Stage Weaners	feed	03/12/2012	SAMN08734351
1950A	4,5,12:i:-	B	1st Stage Weaners	pen	03/12/2012	SAMN08734354

1951A	4,12:i:-	B	1st Stage Weaners	faecal	03/12/2012	SAMN08734356
1956A	4,12:i:-	B	1st Stage Weaners	pen	03/12/2012	SAMN08734291
1958A	4,12:i:-	B	1st Stage Weaners	water	03/12/2012	SAMN08734357
1959A	4,12:i:-	B	1st Stage Weaners	water drinker	03/12/2012	SAMN08734305
1962A	4,12:i:-	B	1st Stage Weaners	pen	03/12/2012	SAMN08734294
1969A	4,12:i:-	B	2nd Stage Weaners	faecal	03/12/2012	SAMN08734355
1971A	4,12:i:-	B	2nd Stage Weaners	water drinker	03/12/2012	SAMN08734358
1973A	4,12:i:-	B	2nd Stage Weaners	trough	03/12/2012	SAMN08734298
1976A	4,12:i:-	B	2nd Stage Weaners	water	03/12/2012	SAMN08734359
1979A	4,12:i:-	B	2nd Stage Weaners	trough	03/12/2012	SAMN08734319
1980A	4,12:i:-	B	2nd Stage Weaners	pen	03/12/2012	SAMN08734299
1981A	4,12:i:-	B	2nd Stage Weaners	faecal	03/12/2012	SAMN08734320
1989B	4,12:i:-	B	2nd Stage Weaners	water drinker	03/12/2012	SAMN08734361
2002A	4,12:i:-	B	Finishers	water	03/12/2012	SAMN08734368
2005A	4,12:i:-	B	Finishers	trough	03/12/2012	SAMN08734296
2024A	4,12:i:-	B	Finishers	pen	03/12/2012	SAMN08734370
2207A	4,5,12:i:-	C	Finishers	pen	07/01/2013	SAMN08734371
2213A	4,5,12:i:-	C	Finishers	pen	07/01/2013	SAMN08734362
2278A	4,5,12:i:-	Mill C		Finisher meal	17/01/2013	SAMN09435863
2455C	*S.* Tm	D	Farrowing	feed	21/01/2013	SAMN08734396
2888B	4,12:i:-	Mill E		Soybean meal	10/04/2013	SAMN08734363
2906A	*S.* Tm	G	Dry sow	water	18/04/2013	SAMN08734398

2911A	*S.* Tm	G	Dry sow	faecal	18/04/2013	SAMN08734399
2930A	*S.* Tm	G	1st Stage Weaners	pen	18/04/2013	SAMN08734315
2954A	*S.* Tm	G	2nd Stage Weaners	water drinker	18/04/2013	SAMN08734400
2956A	*S.* Tm	G	2nd Stage Weaners	trough	18/04/2013	SAMN08734401
2957A	*S.* Tm	G	2nd Stage Weaners	pen	18/04/2013	SAMN08734377
2963A	*S.* Tm	G	2nd Stage Weaners	pen	18/04/2013	SAMN08734316
2971A	*S.* Tm	G	Finishers	pen	18/04/2013	SAMN08734378
2977A	*S.* Tm	G	Finishers	pen	18/04/2013	SAMN08734379
2978A	*S.* Tm	G	Finishers	faecal	18/04/2013	SAMN08734272
2985A	4,5,12:i:-	G	Finishers	water	18/04/2013	SAMN08734373
2994A	4,5,12:i:-	G	Gilts	water drinker	18/04/2013	SAMN08734364
2999A	4,5,12:i:-	G	Gilts	water	18/04/2013	SAMN08734365
3005A	*S.* Tm	G	Gilts	water	18/04/2013	SAMN08734385
3083C	4,5,12:i:-	H	1st Stage Weaners	water drinker	08/05/2013	SAMN08734375
3338A	*S.* Tm	I	Finishers	water	20/05/2013	SAMN08734273
3339A	*S.* Tm	I	Finishers	water drinker	20/05/2013	SAMN08734280
3341A	*S.* Tm	I	Finishers	trough	20/05/2013	SAMN08734389
3356A	*S.* Tm	I	Finishers	water	20/05/2013	SAMN08734390
3379A	*S.* Tm	I	Gilts	trough	20/05/2013	SAMN08734402
3391A	*S.* Tm	I	Gilts	trough	20/05/2013	SAMN08734403
3476A	4,5,12:i:-	J	1st Stage Weaners	faecal	11/06/2013	SAMN08734367
3496A	4,5,12:i:-	J	2nd Stage Weaners	faecal	11/06/2013	SAMN08734393
3500B	4,5,12:i:-	J	2nd Stage Weaners	trough	11/06/2013	SAMN08734300
3508A	4,5,12:i:-	J	2nd Stage Weaners	faecal	11/06/2013	SAMN08734394

3519A	4,5,12:i:-	J	2nd Stage Weaners	pen	11/06/2013	SAMN08734397
3520A	4,5,12:i:-	J	2nd Stage Weaners	faecal	11/06/2013	SAMN08734405
3521A	4,5,12:i:-	J	2nd Stage Weaners	water	11/06/2013	SAMN08734301
3524A	4,5,12:i:-	J	2nd Stage Weaners	trough	11/06/2013	SAMN08734321
3525A	4,5,12:i:-	J	2nd Stage Weaners	pen	11/06/2013	SAMN08734406
3528A	4,5,12:i:-	J	Gilts	faecal	11/06/2013	SAMN08734376
3529A	4,5,12:i:-	J	Gilts	water	11/06/2013	SAMN08734407
3541A	*S.* Tm	J	Gilts	water	11/06/2013	SAMN08734281
3542B	*S.* Tm	J	Gilts	water drinker	11/06/2013	SAMN08734309
3593A	*S.* Tm	J	Finishers	water drinker	11/06/2013	SAMN08734284
3836A	4,5,12:i:-	Mill B		Wheat	01/08/2013	SAMN08734381
3844A	4,5,12:i:-	Mill D		Dry Sow Meal	01/08/2013	SAMN08734382
3845A	4,5,12:i:-	Mill D		Dry Sow Pellets	01/08/2013	SAMN08734383

Appendix 2 List of representative *S*. Typhimurium and *S*. 4,[5],12;i:- used in Chapter 4 to show the distribution of *sopE* in *S*. Typhimurium. The isolate name, date of isolation, sample type, phage type and accession number are reported together with *sopE* presence which was detected with ARIBA (Martin Hunt et al., 2017b).

Isolate	**Date**	**Sample type**	**Phage type**	**Accession n.**	***sopE***
105841997	1997		ND	ERR024629	absent
01_2888		pigeon	DT2		absent
03-715		pigeon	DT2		absent
100419-1995	1995	cattle	DT10	ERR028279	absent
1008-1995	1995	cattle	DT170	ERR028312	absent
10084-1995	1995	cattle	DT10	ERR028278	absent
1013-1997	1995	pig	193	ERR038760	absent
10177-1993	1993	pig	193	ERR038757	absent
10246-1993	1993	cattle	193	ERR038758	absent
10258-1997	1997	turkey	DT85	ERR029225	absent
10382-1995	1995	cattle	193	ERR038759	absent
10902-1996	1996	environment	DT110	ERR028284	absent
10984-1996	1996	cattle	DT108	ERR028292	absent
11020-1996	1996	pig	DT193	ERR028308	absent
1164-1995	1995	cattle	DT141	ERR028295	absent
11671-1996	1996	cattle	193	ERR038761	absent
12005-1995	1995	pig	U323	ERR028290	absent
12342-1996	1996	duck	DT9	ERR029229	absent
1402-2000	1995	cattle	DT170	ERR028313	absent
1713-1998	1998	cattle	DT110	ERR028294	absent
1731-1999	1999	pig	ND	ERR038771	absent
2036-2000	2000	feed	DT85	ERR029228	absent
2087-1997	1997	bird	DT56	ERR029221	absent
2610-1998	1998	horse	DT160	ERR028309	absent
2798-2001	2001	pig	ND	ERR038774	absent
3193-1995			DT102	ERR028272	absent
3203-1997	1997	pig	DT106	ERR028291	absent
3299-1997	1997	bird	DT160	ERR028306	absent
3543-2002	2002	pig	ND	ERR038775	absent
388-1998	1998	dog	DT204	ERR028634	absent
4061-1997	1997	pig	193	ERR038762	absent

4179-2001	2001	bird	DT121	ERR028301	absent
4284-1995	1998	pig	DT120	ERR028300	absent
4300-2001	2001	duck	DT9	ERR029230	absent
4582-1995	1995	cattle	DT104C	ERR028289	absent
5102-1999	1999	pig	ND	ERR039367	absent
5465-1998	1998	duck	DT8	ERR029215	absent
547-2001	2001	pig	DT108	ERR028283	absent
5544-1997	1997	environment	DT179	ERR028307	absent
6164-1997	1997	chicken	DT49	ERR029219	present
6353-1997	1997	pig	DT167	ERR028310	absent
6887-2000	2000	pig	ND	ERR038773	absent
6940-1998	1998	dog	DT126	ERR028303	absent
7302-199	1999	pig	DT104A	ERR028287	absent
7396-1998	1998	bird	DT36	ERR028633	absent
7828-1995	1995	pig	DT103	ERR028285	absent
7830-1995	1995	pig	DT103	ERR028286	absent
818-1998	1998	chicken	DT141	ERR028296	absent
8380-1996	1996	pig	DT179	ERR028316	absent
8721-1997	1997	environment	DT167	ERR028311	absent
8767-1998	1998	pig	DT35	ERR028641	absent
8935-1995	1997	sheep	DT126	ERR028302	absent
9115-1996	1996	cattle	DT204	ERR028631	present
A130	2001	human	ST313	ERS000023	absent
A53_C21	2018	sausage			present
ATCC_14028		ref			absent
B45_C27	2018	block			present
B50_C42	2018	tripe			present
B54_C6	2018	chicken meat			present
B58_C24	2018	chicken meat			present
B7_C4	2018	tripe			present
B8_C7	2018	tripe			absent
B82	2018	chicken meat			present
B9_C8	2018	tripe			present
D23580	2004	human	ST313	FN424405.1	absent
DT1	2009	human	DT1	ERR024408	absent
DT104			DT104	GCA_000493675.1	absent
DT104b			DT104b	ERR028656	absent

DT12	2009	human	DT12	ERR024387	absent
DT120	2009	human	DT120	ERR024391	absent
DT135	2009	human	DT135	ERR024392	absent
DT177	2008	human	DT177	ERR024394	absent
DT191A	2009	human	DT191A	ERR024395	absent
DT193	2009	human	DT193	ERR024396	absent
DT195	2009	human	DT195	ERR024397	absent
DT2		pigeon	DT2	GCA_000493535.1	absent
DT24	2009	human	DT24	ERR024388	absent
DT2B	2009	human	DT2B	ERR024398	absent
DT41	2009	human	DT41	ERR024389	absent
DT41b			DT41b	ERR028658	absent
DT56	2009	human	DT56	ERR024400	absent
DT7	2009	human	DT7	ERR024404	absent
DT8	2009	human	DT8	ERR024405	absent
DT97	2008	human	DT97	ERR024406	absent
DT99	2009	human	DT99	ERR024407	absent
L0004109	2009	dog	DT193	ERR024954	absent
L00961-04	2004	pig	ND	ERR039366	absent
L01157		duck	DT8		absent
L01730-06	2006	pig	DT193	ERR039368	absent
S00060-07	2007	horse	DT40	ERR029217	absent
S00250-07	2007	pig	ND	ERR039364	absent
S00454-09	2009	horse	DT101	ERR028280	absent
S00705-09			DT101	ERR028271	absent
S00914-05	2005	cattle	DT104B	ERR028288	absent
S01491-06	2006	pigeon	DT2A	ERR028638	absent
S01569-10	2010	pig		ERS023498	present
S01960-05	2005	pig	U288	ERR029231	absent
S0272405	2005	pig	rough	ERR024634	absent
S02909-08	2008	pig	DT193	ERR039369	absent
S03188-03	2003	pig	DT170B	ERR028314	absent
S03433-05	2005	chicken	DT36	ERR029213	absent
S03512-08	2008	cat	120	ERR038763	absent
S03645-11		duck	DT8		absent
S04178-09	2009	duck	DT30	ERR028640	absent
S04199-08	2008	pig	U302	ERR029232	absent

S04454-08	2008	pig	DT208	ERR028636	absent
S04527-10		duck	DT8		absent
S04696-09	2009	bird	DT1	ERR028277	absent
S04698-09	2009	cattle	DT193	GCF_001540845.1	present
S04744-08	2008	chicken	DT35	ERR028632	absent
S04782-03	2003	pig	U310	ERR029235	absent
S05081-04	2004	horse	DT40	ERR029216	absent
S05416-06	2006	chicken	DT7	ERR029214	absent
S05451-08	2008	dog	DT56var	ERR029223	absent
S05738-08			U323	ERR028282	absent
S05894-09	2009	pig	DT193	ERR039371	present
S05968-02	2002	pig	U308	ERR029233	absent
S06221-07	2007	cattle	DT135	ERR028304	absent
S06281-04	2004	duck	DT30	ERR028639	absent
S06356-04	2004	feed	DT2A	ERR028637	absent
S07292-07	2007	pig	U308	ERR029234	absent
S07676-03	2003	bird	DT56var	ERR029222	absent
S08300-02	2002	pig	DT12	ERR028299	absent
S08313-02	2002	pig	DT208	ERR028635	absent
S09207-07	2007	pig	DT170B	ERR028315	absent
S09304-02	2002	cattle	DT41	ERR029218	absent
S09313-03	2002	cattle	DT49	ERR029220	present
S33_C6	2018	swab			present
S35_C13	2018	swab			absent
S38_C16	2018	swab			absent
S4489-10	2010	other	ND	ERR038792	absent
S5828-08	2008	chicken	ND	ERR038789	absent
SL1344	1974	cattle	SL1344	GCA_000210855.2	present
SR11-WT			U310	ERR029226	present
U276	2009	human	U276	ERR024409	absent
U277	2009	human	U277	ERR024410	absent
U288	2009	human	U288	ERR024411	absent
U302	2009	human	U302	ERR024401	absent
U310	2009	human	U310	ERR024402	absent
U311			U311	ERR028643	absent
U313	2008	human	U313	ERR024413	absent
U319	2008	human	U319	ERR024417	absent

Appendix 3 List of representative *S.* 4,[5],12:i:- isolates from UK pig farms (2006-2015). The sequences were used in Chapter 4 for ancestral state reconstruction analysis for *sopE* acquisition/loss and inference of ancestral dates.

The isolate name, phage type, year of isolation and the reference is reported together with presence/absence of *sopE* determined with SRST2 (Inouye et al., 2014) using the sequence of S04698-09 as reference (accession number NZ_LN999997.1). The isolates obtained from APHA were in-house sequenced.

Strain name	**Phage type**	**Year of isolation**	**Reference**	***sopE***
L01730-06	ND	2006	(Petrovska et al., 2016)	absent
S0806007	193	2007	(Petrovska et al., 2016)	absent
S0337107	DT193	2007	(Petrovska et al., 2016)	absent
L0064707	DT193	2007	(Petrovska et al., 2016)	absent
S03407-08	DT193	2008	APHA	absent
L01074-08	DT193	2008	APHA	absent
S04619-08	DT193	2008	APHA	absent
S05574-08	DT193	2008	APHA	absent
L01176-08	ND	2008	(Petrovska et al., 2016)	absent
L01189-08	ND	2008	(Petrovska et al., 2016)	absent
S02909-08	DT193	2008	(Petrovska et al., 2016)	absent
L00028-09	DT193	2009	APHA	absent
S03797-09	DT193	2009	APHA	absent
S04409-09	DT193	2009	APHA	present
S05893-09	DT193	2009	(Petrovska et al., 2016)	present
S0433209	DT193	2009	(Petrovska et al., 2016)	present
SO5894-09	DT193	2009	APHA	present
S01523-10	DT193	2010	APHA	absent
S01299-10	DT193	2010	APHA	absent
4824-10	ND	2010	(Petrovska et al., 2016)	present
S03113-10	DT193	2010	(Petrovska et al., 2016)	present
S01569-10	DT120	2010	(Petrovska et al., 2016)	present
S00823-11	DT193	2011	APHA	absent
S01402-11	DT193	2011	APHA	absent
S00801-11	DT193	2011	APHA	absent

S01024-11	DT193	2011	APHA	absent
S02728-11	DT193	2011	APHA	absent
L00197-11	DT193	2011	APHA	present
S00655-11	DT193	2011	APHA	present
L01204-11	DT193	2011	APHA	present
S03515-12	DT193	2012	APHA	present
L00153-12	DT193	2012	APHA	present
S00479-12	DT193	2012	APHA	present
S01271-12	DT193	2012	APHA	present
S01401-12	DT193	2012	APHA	present
S02335-12	DT193	2012	APHA	present
S02692-12	DT193	2012	APHA	present
S03101-12	DT193	2012	APHA	present
S01051-13	DT193	2013	APHA	absent
S01879-13	DT193	2013	APHA	absent
S01052-13	DT193	2013	APHA	present
S01599-13	DT193	2013	APHA	present
S02698-13	DT193	2013	APHA	present
S00968-13	DT193	2013	APHA	present
S01539-13	DT193	2013	APHA	present
S02597-13	DT193	2013	APHA	present
S02933-13	DT193	2013	APHA	present
S01892-14	DT193	2014	APHA	present
S02680-14	DT193	2014	APHA	present
S03407-14	DT193	2014	APHA	present
L00907-14	DT193	2014	APHA	present
L01338-14	DT193	2014	APHA	present
L01783-14	DT193	2014	APHA	present
S00888-14	DT193	2014	APHA	present
S01036-14	DT193	2014	APHA	present
S02412-14	DT193	2014	APHA	present
S02503-14	DT193	2014	APHA	present
L00644-15	DT193	2015	APHA	absent
S00313-15	DT193	2015	APHA	present

S01209-15	DT193	2015	APHA	present
S01410-15	DT193	2015	APHA	present
L00036-15	DT193	2015	APHA	present

Appendix 4 List of *S*. Typhimurium and *S*. 4,[5],12:i:- from isolated from human infections in the UK. The dataset was used in Chapter 4 and 6. The isolate name, sequence type (ST), type of sample, date of isolation phage type and accession number are reported.

Isolate	**ST**	**Sample**	**Date**	**Phage Type**	**Accession number**
H124521012	34	Chicken	06/11/2012	DT 120	SRR1959239
H124521014	34	Chicken	06/11/2012	DT 120	SRR1962284
H124521018	34	Poultry	06/11/2012	DT 120	SRR1959495
H124521082	19	Human	06/11/2012	DT 135	SRR1963221
H124521083	19	Human	06/11/2012	DT 8	SRR1963502
H124521119	19	Human	06/11/2012	DT 135	SRR1960376
H124521124	19	Human	06/11/2012	DT 9	SRR1960184
H124521125	19	Human	06/11/2012	DT 208	SRR1960179
H124521137	19	Human	06/11/2012	DT 64	SRR1959429
H124521169	19	Human	06/11/2012	DT 135	SRR1960041
H124521176	19	Human	06/11/2012	RDNC	SRR1963370
H124521199	19	Human	06/11/2012	DT 64	SRR1963484
H124521212	19	Human	06/11/2012	DT 208	SRR1959219
H124521302	19	Chicken	06/11/2012	DT 35	SRR1960135
H124521308	19	Chicken	06/11/2012	DT 135	SRR1963374
H124521316	19	Chicken	06/11/2012	DT 30	SRR1960204
H124521322	19	Poultry	06/11/2012	DT 135	SRR1963343
H124521330	19	Animal	06/11/2012	DT 8	SRR1960230
H124521333	19	Poultry	06/11/2012	DT 35	SRR1963522
H124521334	19	Animal	06/11/2012	DT 64	SRR1963460
H132260914	34	Chicken	30/05/2013	DT 120	SRR1963298
H132260959	34	Human	30/05/2013	DT 120	SRR1960058
H132260971	19	Human	30/05/2013	DT 64	SRR1960381
H133020731	19	Human	23/07/2013	DT 8	SRR1967848
H141420383	19	Human	01/04/2014	None	SRR3049370
H141420385	19	Human	01/04/2014	DT 193	SRR3049711

H141420391	19	Human	01/04/2014	DT 193	SRR3049860
H141420393	34	Human	01/04/2014	RDNC	SRR3049071
H141420398	19	Human	01/04/2014	DT 104	SRR3048655
H141420400	19	Human	01/04/2014	DT 193	SRR3049214
H141420402	34	Human	01/04/2014	DT 193	SRR1970222
H141460669	34	Human	03/04/2014	DT 193	SRR1968124
H141480357	19	Human	04/04/2014	DT 15a	SRR1969055
H141480359	19	Human	04/04/2014	DT 177	SRR1967933
H141480439	313	Human	04/04/2014	DT 56	SRR1967883
H141500474	34	Human	07/04/2014	DT 193	SRR1968522
H141500475	34	Human	07/04/2014	None	SRR1969016
H141500476	34	Human	07/04/2014	DT 193	SRR3049541
H141500477	34	Human	07/04/2014	DT 193	SRR1968097
H141500478	19	Human	07/04/2014	RDNC	SRR1970169
H141500484	19	Human	07/04/2014	DT 193	SRR1966721
H141500485	34	Human	07/04/2014	DT 193	SRR1968694
H141520436	19	Human	08/04/2014	RDNC	SRR3049211
H141520468	213	Human	08/04/2014	None	SRR1970253
H141520470	313	Human	08/04/2014	DT 56	SRR1967248
H141540514	19	Human	09/04/2014	DT 104	SRR1967639
H141540529	34	Human	09/04/2014	DT 120	SRR1966181
H141560271	34	Human	10/04/2014	DT 193	SRR1965093
H141580495	34	Raw meat	11/04/2014	DT 193	SRR1968666
H141580499	19	Human	11/04/2014	DT 104b	SRR1967710
H141580500	19	Human	11/04/2014	DT 195	SRR1968142
H141580501	313	Human	11/04/2014	DT 143	SRR1966408
H141580502	19	Human	11/04/2014	RDNC	SRR1970078
H141580503	34	Human	11/04/2014	DT 193	SRR1965395
H141580513	313	Human	11/04/2014	DT 56	SRR1966068
H141600085	34	Human	14/04/2014	DT 193	SRR1968340
H141600086	34	Human	14/04/2014	DT 193	SRR1965475
H141600088	34	Human	14/04/2014	DT 193	SRR1965672
H141600092	34	Human	14/04/2014	None	SRR1968741
H141600104	19	Human	14/04/2014	None	SRR1959465
H141600104	19	Human	14/04/2014	None	SRR1967309
H141600104	19	Human	14/04/2014	None	SRR1968116
H141600105	34	Human	14/04/2014	None	SRR1967578

H141620518	313	Human	15/04/2014	DT 56	SRR1965662
H141620518	313	Human	15/04/2014	DT 56	SRR1969775
H141620520	19	Human	15/04/2014	Untypable	SRR1967439
H141620520	19	Human	15/04/2014	Untypable	SRR1969989
H141620522	19	Human	15/04/2014	DT 8	SRR1968816
H141620522	19	Human	15/04/2014	DT 8	SRR1970008
H141620526	34	Human	15/04/2014	DT 193	SRR3049520
H141620528	19	Human	15/04/2014	DT 193	SRR1966271
H141620528	19	Human	15/04/2014	DT 193	SRR3049244
H141620529	313	Human	15/04/2014	DT 56	SRR1970264
H141620529	313	Human	15/04/2014	DT 56	SRR3048536
H141620533	34	Human	15/04/2014	DT 120	SRR1965137
H141640734	19	Human	16/04/2014	DT 104b	SRR1967148
H141640739	19	Human	16/04/2014	None	SRR1969201
H141640739	19	Human	16/04/2014	None	SRR1970027
H141640740	19	Human	16/04/2014	None	SRR1967480
H141640740	19	Human	16/04/2014	None	SRR1969510
H141660473	34	Human	17/04/2014	DT 193	SRR1965994
H141660474	19	Human	17/04/2014	DT 104b	SRR1969233
H141660475	19	Raw pork	17/04/2014	PT U311	SRR1967514
H141660476	34	Human	17/04/2014	DT 104	SRR1965389
H141660477	19	Human	17/04/2014	DT 120	SRR1965710
H141660478	34	Human	17/04/2014	None	SRR1965476
H141660479	19	Human	17/04/2014	DT 104b	SRR1969130
H141660481	34	Pork	17/04/2014	DT 193	SRR1967813
H141660485	34	Raw meat	17/04/2014	DT 193	SRR1969121
H141741067	19	Human	23/04/2014	DT 104	SRR1967013
H141741091	34	Human	23/04/2014	DT 120	SRR1966667
H141741092	34	Human	23/04/2014	PT U323	SRR1968532
H141741094	34	Human	23/04/2014	None	SRR3049401
H141741095	34	Human	23/04/2014	None	SRR1970033
H141741114	34	Unknown	23/04/2014	DT 193	SRR1969145
H141741117	34	Unknown	23/04/2014	DT 193	SRR1970037
H141741118	19	Unknown	23/04/2014	DT 104b	SRR1965698
H141741121	19	NCTC 5713B3	23/04/2014	DT 86	SRR1966220
H141760554	34	Human	24/04/2014	DT 193	SRR1968894
H141760555	19	Human	24/04/2014	PT U313	SRR1967882

H141760556	34	Human	24/04/2014	DT 193	SRR1968317
H141760557	34	Human	24/04/2014	None	SRR1969044
H141760561	19	Human	24/04/2014	None	SRR1967195
H141780732	34	Raw meat	25/04/2014	DT 21	SRR1966564
H141780739	19	Human	25/04/2014	None	SRR1969481
H141780740	34	Human	25/04/2014	DT 195	SRR1967777
H141780741	34	Human	25/04/2014	DT 193	SRR3049286
H141800491	34	Human	28/04/2014	DT 193	SRR1968308
H141800493	34	Human	28/04/2014	None	SRR1969290
H141800494	34	Human	28/04/2014	None	SRR1967214
H141820386	34	Human	29/04/2014	DT 193	SRR1965854
H141820387	34	Human	29/04/2014	DT 193	SRR1967415
H141820388	19	Human	29/04/2014	DT 193	SRR1969355
H141820389	19	Human	29/04/2014	DT 193	SRR1965524
H141820390	19	Human	29/04/2014	DT 104	SRR1969825
H141820391	19	Human	29/04/2014	PT U314	SRR1965880
H141820392	19	Human	29/04/2014	DT 193	SRR1966615
H141820393	19	Human	29/04/2014	PT U288	SRR1967525
H141820410	34	Human	29/04/2014	DT 193	SRR1968485
H141840455	19	Human	30/04/2014	PT U312	SRR1966007
H141840456	34	Human	30/04/2014	None	SRR1969254
H141840457	19	Human	30/04/2014	PT U312	SRR3048544
H141840458	19	Human	30/04/2014	DT 104	SRR1970109
H141840462	19	Human	30/04/2014	DT 193	SRR1968690
H141840463	213	EQA	30/04/2014	DT 1	SRR1967397
H141840464	34	Raw meat	30/04/2014	DT 21	SRR1968954
H141840475	313	Human	30/04/2014	None	SRR1966656
H141840480	19	Human	30/04/2014	DT 120	SRR1967406
H141840482	34	Human	30/04/2014	DT 120	SRR3049815
H141860444	19	Human	01/05/2014	RDNC	SRR1965324
H141860462	34	Human	01/05/2014	DT 120	SRR1967318
H141880251	19	Human	02/05/2014	PT U312	SRR1965534
H141880252	19	Human	02/05/2014	PT 193a	SRR1969647
H141880253	19	Human	02/05/2014	PT U317	SRR1967104
H141880254	34	Unknown	02/05/2014	DT 120	SRR1968669
H141880268	34	Human	02/05/2014	DT 193	SRR3049413
H141920581	19	Human	06/05/2014	DT 135	SRR1966888

H141920582	34	Human	06/05/2014	DT 193	SRR1967524
H141920583	19	Human	06/05/2014	DT 104	SRR1966639
H141920584	19	Human	06/05/2014	DT 22	SRR1966152
H141920596	19	Human	06/05/2014	None	SRR1966808
H141940684	34	Human	07/05/2014	None	SRR1966334
H141940696	34	Human	07/05/2014	None	SRR1969662
H141940697	19	Human	07/05/2014	DT 104	SRR1966348
H141940698	19	Human	07/05/2014	RDNC	SRR1967995
H141940699	34	Human	07/05/2014	DT 193	SRR1968379
H141940700	19	Human	07/05/2014	DT 104b	SRR1967031
H141940702	313	Human	07/05/2014	DT 132	SRR1968363
H141940704	19	Human	07/05/2014	DT 170b	SRR1966814
H141940705	34	Human	07/05/2014	DT 193	SRR1969473
H141960658	19	Human	08/05/2014	None	SRR1966126
H141960659	34	Human	08/05/2014	DT 120	SRR1969394
H141960666	34	Human	08/05/2014	PT U323	SRR1968427
H141960674	34	Human	08/05/2014	DT 7a	SRR1968613
H141960675	34	Human	08/05/2014	DT 193	SRR1967471
H141960676	34	Human	08/05/2014	DT 193	SRR1967241
H141980453	34	Unknown	09/05/2014	DT 193	SRR1969202
H141980458	19	NCTC	09/05/2014	DT 104	SRR1965423
H141980465	19	Human	09/05/2014	DT 74	SRR3049368
H141980466	19	Human	09/05/2014	DT 1	SRR1965331
H141980469	34	Human	09/05/2014	DT 193	SRR1967782
H142000598	34	Human	12/05/2014	None	SRR1967409
H142000599	313	Human	12/05/2014	DT 132	SRR1965263
H142000600	19	Human	12/05/2014	DT 1	SRR1969252
H142020343	34	Human	13/05/2014	DT 193	SRR1967759
H142020344	19	Human	13/05/2014	DT 41	SRR1968523
H142020349	34	Human	13/05/2014	None	SRR1967257
H142020351	19	Human	13/05/2014	None	SRR1968793
H142020418	34	Human	13/05/2014	DT 193	SRR1966666
H142020419	34	Human	13/05/2014	DT 193	SRR1966513
H142020420	19	Human	13/05/2014	RDNC	SRR1968022
H142020441	34	Human	13/05/2014	DT 193	SRR1967279
H142040424	34	Human	14/05/2014	DT 104b	SRR1966739
H142040426	19	Human	14/05/2014	DT 101	SRR1965467

H142040436	34	Human	14/05/2014	DT 104b	SRR1965297
H142040442	34	Human	14/05/2014	DT 120	SRR1966431
H142040445	19	Human	14/05/2014	RDNC	SRR1968950
H142040455	19	Human	14/05/2014	None	SRR1967783
H142060410	34	Human	15/05/2014	DT 193	SRR1966862
H142060412	568	Human	15/05/2014	PT U277	SRR1969941
H142060415	34	Human	15/05/2014	PT 193a	SRR1966299
H142080647	34	Human	16/05/2014	DT 193	SRR1969659
H142080648	19	Human	16/05/2014	None	SRR1966890
H142100254	34	Human	19/05/2014	DT 193	SRR1966572
H142100255	19	Human	19/05/2014	DT 170b	SRR1967022
H142100256	19	Human	19/05/2014	PT U302	SRR1966720
H142100257	34	Human	19/05/2014	DT 193	SRR1970287
H142100258	19	Human	19/05/2014	DT 41	SRR1969072
H142100259	19	Human	19/05/2014	DT 29	SRR1966645
H142120282	19	Human	20/05/2014	DT 64	SRR1967889
H142120283	34	Human	20/05/2014	DT 193	SRR1967841
H142140689	313	Human	21/05/2014	DT 103	SRR1965387
H142140691	19	Human	21/05/2014	DT 1	SRR1965859
H142140696	34	Human	21/05/2014	None	SRR1967753
H142140702	19	Human	21/05/2014	DT 193	SRR1967547
H142160283	34	Human	22/05/2014	DT 120	SRR1966328
H142160284	19	Human	22/05/2014	DT 104b	SRR1970092
H142160286	19	Human	22/05/2014	None	SRR1969680
H142160292	34	Raw meat	22/05/2014	DT 193	SRR1967276
H142160293	34	Human	22/05/2014	DT 193	SRR1966719
H142160295	34	Human	22/05/2014	DT 120	SRR1965971
H142180411	34	Human	23/05/2014	DT 193	SRR1966322
H142180412	34	Human	23/05/2014	DT 193	SRR1968149
H142180413	34	Human	23/05/2014	DT 193	SRR1968380
H142240388	19	Human	28/05/2014	PT U310	SRR3048970
H142240391	34	Human	28/05/2014	DT 193	SRR3048932
H142240393	19	Human	28/05/2014	DT 101	SRR3049018
H142240394	19	Human	28/05/2014	DT 161	SRR1965556
H142240396	34	Human	28/05/2014	DT 193	SRR3049084
H142240397	19	Human	28/05/2014	DT 22	SRR3048528
H142240472	19	Human	28/05/2014	DT 161	SRR1965311

H142260380	313	Human	29/05/2014	DT 15a	SRR1967419
H142260381	34	Human	29/05/2014	DT 193	SRR1968330
H142260381	34	Human	29/05/2014	DT 193	SRR3285426
H142260451	34	Human	29/05/2014	DT 193	SRR3049140
H142280446	34	Human	30/05/2014	DT 193	SRR3049246
H142300580	19	Human	02/06/2014	DT 104	SRR3048779
H142300582	34	Human	02/06/2014	DT 193	SRR3049151
H142320339	34	Human	03/06/2014	DT 120	SRR3049032
H142320344	19	Human	03/06/2014	RDNC	SRR3048906
H142320354	34	Human	03/06/2014	DT 193	SRR3049077
H142320355	34	Human	03/06/2014	DT 193	SRR3049233
H142320356	34	Human	03/06/2014	DT 120	SRR3049309
H142320357	34	Human	03/06/2014	DT 193	SRR3049492
H142320358	19	Human	03/06/2014	DT 104	SRR3048726
H142320411	19	Human	03/06/2014	PT U310	SRR3049236
H142340624	19	Human	04/06/2014	DT 2	SRR3048751
H142340625	34	Human	04/06/2014	DT 193	SRR3048554
H142340626	34	Human	04/06/2014	DT 193	SRR3049946
H142340627	34	Human	04/06/2014	DT 193	SRR3048936
H142340629	19	Human	04/06/2014	DT 73	SRR3049405
H142360379	19	Human	05/06/2014	DT 8	SRR3048728
H142360387	34	Human	05/06/2014	DT 120	SRR1969847
H142360389	34	Human	05/06/2014	DT 120	SRR1966276
H142380312	34	Human	06/06/2014	PT U311	SRR3049115
H142380313	34	Human	06/06/2014	PT U323	SRR3049268
H142380314	19	Human	06/06/2014	DT 2	SRR3049637
H142380315	34	Human	06/06/2014	DT 193	SRR3049935
H142380316	34	Human	06/06/2014	DT 193	SRR3048539
H142380317	19	Human	06/06/2014	DT 104	SRR3048556
H142380318	19	Human	06/06/2014	DT 2	SRR3049693
H142380320	34	Human	06/06/2014	DT 193	SRR3049419
H142400424	19	Human	09/06/2014	DT 191a	SRR3049141
H142420764	34	Human	10/06/2014	DT 193	SRR3049472
H142420765	19	Human	10/06/2014	DT 191a	SRR3049776
H142420766	19	Human	10/06/2014	RDNC	SRR1969512
H142420768	34	Human	10/06/2014	DT 193	SRR3048934
H142420769	34	Human	10/06/2014	DT 193	SRR3049572

H142420807	19	Human	10/06/2014	RDNC	SRR3048706
H142420808	34	Human	10/06/2014	None	SRR3049086
H142420821	19	Human	10/06/2014	None	SRR3049876
H142420822	34	Human	10/06/2014	DT 193	SRR3049197
H142440682	19	Human	11/06/2014	PT U302	SRR3049952
H142440683	34	Human	11/06/2014	DT 193	SRR3049186
H142440684	19	Human	11/06/2014	DT 177	SRR3048571
H142440685	34	Human	11/06/2014	DT 193	SRR3049696
H142440703	19	Human	11/06/2014	DT 8	SRR1966440
H142440705	34	Human	11/06/2014	DT 193	SRR1966049
H142440707	34	Human	11/06/2014	None	SRR3049373
H142460293	19	Human	12/06/2014	DT 94	SRR1966243
H142500094	19	Human	16/06/2014	DT 104b	SRR3049902
H142500101	34	Human	16/06/2014	DT 193	SRR3049259
H142520441	19	Human	17/06/2014	DT 203	SRR3049530
H142520442	19	Human	17/06/2014	DT 104	SRR3049263
H142520443	19	Human	17/06/2014	PT U302	SRR3049861
H142520444	34	Human	17/06/2014	PT U323	SRR1966811
H142520827	19	Human	17/06/2014	DT 193	SRR3048904
H142540301	34	Human	18/06/2014	DT 193	SRR1969506
H142540304	19	Human	18/06/2014	DT 49	SRR3048886
H142540308	19	Human	18/06/2014	DT 2	SRR3049632
H142540309	34	Human	18/06/2014	DT 193	SRR3049283
H142540310	19	Human	18/06/2014	DT 104	SRR3048685
H142540390	34	Human	18/06/2014	DT 193	SRR1957769
H142540395	19	Human	18/06/2014	DT 8	SRR1969521
H142540400	568	Human	18/06/2014	DT 178	SRR1968619
H142560489	19	Human	19/06/2014	DT 193	SRR1965047
H142560491	34	Human	19/06/2014	DT 193	SRR1970013
H142560538	19	Human	19/06/2014	DT 1	SRR1966291
H142580387	34	Human	20/06/2014	DT 120	SRR1968635
H142600609	34	Human	23/06/2014	DT 120	SRR1969733
H142600655	19	Human	23/06/2014	DT 104	SRR1969141
H142600656	19	Human	23/06/2014	DT 46	SRR1967292
H142600659	19	Human	23/06/2014	DT 193	SRR1970326
H142600660	34	Human	23/06/2014	DT 193	SRR1967420
H142600661	34	Human	23/06/2014	DT 193	SRR1969149

H142600665	34	Human	23/06/2014	DT 195	SRR1965964
H142620391	19	Human	24/06/2014	DT 2	SRR1969751
H142640381	34	Human	25/06/2014	DT 193	SRR1968528
H142640382	34	Human	25/06/2014	DT 193	SRR1967912
H142640383	19	Human	25/06/2014	DT 191a	SRR1970215
H142640384	19	Human	25/06/2014	DT 193	SRR1967048
H142640385	19	Human	25/06/2014	PT U302	SRR1965108
H142640386	19	Human	25/06/2014	None	SRR1957839
H142640387	19	Human	25/06/2014	None	SRR1958073
H142640394	34	Human	25/06/2014	DT 193	SRR1968735
H142660316	34	Human	26/06/2014	DT 193	SRR3048847
H142680476	34	Human	27/06/2014	DT 193	SRR1970057
H142680477	34	Human	27/06/2014	PT U311	SRR1969479
H142680479	19	Unknown	27/06/2014	RDNC	SRR1968072
H142680481	34	Human	27/06/2014	DT 193	SRR1966581
H142700048	19	Human	30/06/2014	DT 104	SRR1969171
H142720293	34	Human	01/07/2014	PT U311	SRR1968087
H142720294	34	Human	01/07/2014	RDNC	SRR1968412
H142720295	19	Human	01/07/2014	PT U314	SRR1967208
H142720297	19	Human	01/07/2014	DT 193	SRR1968437
H142720298	34	Human	01/07/2014	DT 193	SRR1968823
H142720299	34	Human	01/07/2014	DT 193	SRR1966731
H142740617	34	Human	02/07/2014	DT 193	SRR1966506
H142740622	34	Human	02/07/2014	DT 193	SRR1969703
H142740624	34	Human	02/07/2014	DT 193	SRR1965188
H142740625	19	Human	02/07/2014	DT 193	SRR1967093
H142740626	19	Human	02/07/2014	PT U302	SRR1966549
H142740627	19	Human	02/07/2014	DT 104b	SRR1965663
H142740628	19	Human	02/07/2014	DT 99	SRR1966494
H142740641	313	Human	02/07/2014	DT 56	SRR1966509
H142740643	19	Human	02/07/2014	DT 177	SRR1969251
H142760493	19	Human	03/07/2014	DT 132	SRR1970327
H142760495	34	Human	03/07/2014	PT U323	SRR1966501
H142760496	19	Human	03/07/2014	None	SRR1969658
H142760497	34	Human	03/07/2014	DT 193	SRR1968311
H142760499	34	Human	03/07/2014	DT 193	SRR1968253
H142760550	34	Human	03/07/2014	DT 193	SRR1969066

H142780347	34	Human	04/07/2014	None	SRR1966368
H142780372	19	Human	04/07/2014	DT 40	SRR1965151
H142780373	19	Human	04/07/2014	DT 35	SRR1970031
H142780374	19	Human	04/07/2014	DT 104	SRR1966980
H142780376	34	Human	04/07/2014	DT 193	SRR1969606
H142800188	34	Human	07/07/2014	DT 193	SRR3048565
H142800189	34	Human	07/07/2014	DT 193	SRR3049727
H142800190	19	Human	07/07/2014	DT 17	SRR1967332
H142800191	19	Human	07/07/2014	None	SRR3048981
H142800209	19	Human	07/07/2014	Untypable	SRR3049317
H142820396	34	Human	08/07/2014	None	SRR1966936
H142820412	19	Human	08/07/2014	DT 8	SRR1968247
H142820418	34	Human	08/07/2014	DT 193	SRR1965718
H142820420	19	Human	08/07/2014	DT 135	SRR1969904
H142820422	34	Human	08/07/2014	DT 193	SRR1966234
H142820424	19	Human	08/07/2014	DT 135	SRR1969646
H142820426	34	Human	08/07/2014	DT 193	SRR1969350
H142820427	34	Human	08/07/2014	DT 120	SRR3049754
H142820428	19	Human	08/07/2014	DT 101	SRR1966766
H142820429	19	Human	08/07/2014	DT 101	SRR1966443
H142820431	19	Human	08/07/2014	None	SRR1968706
H142820468	19	Human	08/07/2014	DT 193	SRR1969241
H142840476	34	Human	09/07/2014	None	SRR1968835
H142840477	19	Human	09/07/2014	DT 69	SRR1970219
H142840479	19	Human	09/07/2014	DT 101	SRR1965466
H142840480	34	Human	09/07/2014	DT 193	SRR1969999
H142840488	34	Human	09/07/2014	DT 193	SRR1966941
H142860385	19	Human	10/07/2014	DT 193	SRR1958551
H142860396	19	Human	10/07/2014	DT 135	SRR1958533
H142860397	34	Human	10/07/2014	DT 193	SRR1958354
H142860403	19	Human	10/07/2014	DT 193	SRR1958492
H142880242	34	Human	11/07/2014	DT 193	SRR1965696
H142880244	19	Human	11/07/2014	PT U310	SRR1968426
H142880245	34	Human	11/07/2014	None	SRR1968008
H142880246	34	Human	11/07/2014	None	SRR1968510
H142880253	19	Human	11/07/2014	DT 104	SRR1965981
H142900379	19	Human	14/07/2014	DT 104b	SRR1966835

H142900475	19	Human	14/07/2014	PT U311	SRR1966226
H142900476	34	Human	14/07/2014	DT 193	SRR1966576
H142920569	19	Human	15/07/2014	DT 135	SRR1966160
H142920572	19	Human	15/07/2014	PT U310	SRR1969046
H142920573	19	Human	15/07/2014	DT 49	SRR1968846
H142920575	19	Human	15/07/2014	DT 1	SRR1969701
H142920576	34	Human	15/07/2014	DT 193	SRR1969441
H142920583	19	Human	15/07/2014	PT U288	SRR1968818
H142920585	34	Human	15/07/2014	DT 193	SRR1966413
H142920602	19	Human	15/07/2014	PT U310	SRR3049571
H142920606	19	Human	15/07/2014	DT 104	SRR1970081
H142920607	19	Human	15/07/2014	DT 8	SRR3048768
H142920608	213	Human	15/07/2014	DT 1	SRR1965680
H142920612	19	Human	15/07/2014	DT 195	SRR3049501
H142940532	34	Human	16/07/2014	DT 193	SRR3049813
H142940534	34	Human	16/07/2014	DT 193	SRR3048914
H142940536	19	Human	16/07/2014	None	SRR3049928
H142940545	213	Human	16/07/2014	DT 1	SRR3049563
H142940549	213	Human	16/07/2014	DT 1	SRR3049232
H142940551	34	Human	16/07/2014	DT 193	SRR1965358
H142940552	19	Human	16/07/2014	PT U288	SRR3048526
H142940554	34	Human	16/07/2014	PT U311	SRR3048543
H142940555	34	Human	16/07/2014	PT U311	SRR3049138
H142940556	34	Human	16/07/2014	PT U311	SRR3049132
H142940557	P3080	Human	16/07/2014	PT U311	SRR3049417
H142940559	34	Human	16/07/2014	PT U311	SRR3049500
H142940577	19	Human	16/07/2014	DT 104	SRR3049082
H142940579	34	Human	16/07/2014	DT 193	SRR3048555
H142960507	19	Human	17/07/2014	DT 74	SRR1965261
H142960508	34	Human	17/07/2014	DT 193	SRR1966211
H142960510	19	Human	17/07/2014	PT U312	SRR1967045
H142960521	34	Human	17/07/2014	DT 193	SRR1968909
H142960527	19	Human	17/07/2014	DT 2	SRR1968883
H142980299	19	Human	18/07/2014	None	SRR1967836
H142980300	34	Human	18/07/2014	DT 193	SRR1967386
H142980300	34	Human	18/07/2014	DT 193	SRR1968598
H142980302	34	Human	18/07/2014	DT 193	SRR1967488

H143000283	19	Human	21/07/2014	PT U288	SRR3049753
H143000285	34	Human	21/07/2014	DT 193	SRR3049833
H143000286	34	Human	21/07/2014	None	SRR3049442
H143000286	34	Human	21/07/2014	None	SRR3286715
H143000295	34	Human	21/07/2014	DT 193	SRR1970024
H143000297	34	Human	21/07/2014	DT 193	SRR3048800
H143020367	19	Human	22/07/2014	PT U302	SRR1966551
H143020368	34	Human	22/07/2014	DT 193	SRR1967249
H143020369	34	Human	22/07/2014	DT 22	SRR1968578
H143020372	34	Human	22/07/2014	DT 193	SRR1967950
H143020373	34	Human	22/07/2014	DT 193	SRR1966514
H143020374	19	Human	22/07/2014	PT U288	SRR1966188
H143020408	34	Human	22/07/2014	DT 193	SRR3048518
H143040502	19	Human	23/07/2014	PT U288	SRR3049335
H143040504	34	Human	23/07/2014	DT 193	SRR3049016
H143040505	34	Human	23/07/2014	DT 193	SRR3049591
H143040507	34	Human	23/07/2014	PT U311	SRR3049156
H143040509	19	Human	23/07/2014	DT 195	SRR3049633
H143040511	34	Human	23/07/2014	DT 193	SRR3048857
H143040550	19	Human	23/07/2014	DT 101	SRR1970083
H143060181	34	Human	24/07/2014	PT U311	SRR1969000
H143060187	19	Human	24/07/2014	DT 205	SRR1965637
H143060188	34	Human	24/07/2014	DT 193	SRR1966614
H143060189	19	Human	24/07/2014	DT 193	SRR1968039
H143060191	19	Human	24/07/2014	DT 193	SRR1969511
H143060192	19	Human	24/07/2014	PT U302	SRR1969554
H143060271	34	Human	24/07/2014	PT U311	SRR1969014
H143080507	19	Human	25/07/2014	DT 193	SRR1969930
H143080514	34	Human	25/07/2014	DT 193	SRR1967651
H143080520	34	Human	25/07/2014	DT 193	SRR1965173
H143080521	19	Human	25/07/2014	DT 104	SRR1968560
H143080522	19	Human	25/07/2014	RDNC	SRR1970020
H143100440	19	Human	28/07/2014	RDNC	SRR1967256
H143100441	19	Human	28/07/2014	RDNC	SRR1969827
H143100443	34	Human	28/07/2014	DT 193	SRR1969308
H143100447	34	Human	28/07/2014	DT 193	SRR1969998
H143100449	19	Human	28/07/2014	PT U276	SRR3049573

H143100450	34	Human	28/07/2014	DT 193	SRR1969065
H143120411	34	Human	29/07/2014	PT U311	SRR1969420
H143120415	34	Human	29/07/2014	DT 193	SRR1970229
H143120421	19	Human	29/07/2014	DT 2	SRR1967843
H143120429	568	Human	29/07/2014	DT 40	SRR1969317
H143120432	34	Human	29/07/2014	DT 193	SRR1969480
H143120433	34	Human	29/07/2014	DT 193	SRR1969071
H143120493	19	Human	29/07/2014	RDNC	SRR1966038
H143120493	19	Human	29/07/2014	RDNC	SRR1968854
H143140475	19	Human	30/07/2014	DT 193	SRR1969019
H143140481	34	Human	30/07/2014	DT 193	SRR1968040
H143140483	34	Human	30/07/2014	DT 120	SRR1966371
H143140486	34	Human	30/07/2014	DT 193	SRR1969518
H143140487	34	Human	30/07/2014	DT 120	SRR1965941
H143140488	19	Human	30/07/2014	DT 9	SRR1968012
H143140489	19	Human	30/07/2014	DT 49	SRR1966380
H143140492	34	Human	30/07/2014	DT 193	SRR1966331
H143140492	34	Human	30/07/2014	DT 193	SRR1968825
H143140493	19	Human	30/07/2014	DT 193	SRR1966433
H143140578	568	Human	30/07/2014	None	SRR1969610
H143140582	19	Human	30/07/2014	DT 42	SRR1966232
H143140589	19	Human	30/07/2014	DT 193	SRR1967701
H143140591	19	Human	30/07/2014	DT 43	SRR1965088
H143160363	19	Human	31/07/2014	DT 191a	SRR1967365
H143160439	213	Human	31/07/2014	DT 193	SRR1967622
H143160444	19	Human	31/07/2014	DT 104	SRR1957845
H143160444	19	Human	31/07/2014	DT 104	SRR1967795
H143180347	34	Human	01/08/2014	PT U311	SRR1968837
H143180352	19	Human	01/08/2014	PT U288	SRR1967395
H143180353	19	Human	01/08/2014	DT 41b	SRR1966986
H143180355	34	Human	01/08/2014	DT 193	SRR1969888
H143180357	34	Human	01/08/2014	DT 193	SRR1958247
H143180357	34	Human	01/08/2014	DT 193	SRR1969164
H143200210	19	Human	04/08/2014	DT 104b	SRR1965074
H143200215	34	Human	04/08/2014	DT 193	SRR1969602
H143200218	34	Human	04/08/2014	DT 193	SRR1966516
H143200580	34	Human	04/08/2014	PT U311	SRR1967798

H143200611	19	Human	04/08/2014	DT 104	SRR1965446
H143200618	19	Human	04/08/2014	DT 8	SRR1965081
H143220671	34	Human	05/08/2014	DT 193	SRR1967312
H143220672	19	Human	05/08/2014	DT 8	SRR1968481
H143220674	19	Human	05/08/2014	None	SRR1965648
H143220678	34	Human	05/08/2014	DT 193	SRR1969177
H143220686	19	Human	05/08/2014	RDNC	SRR1967594
H143220688	34	Human	05/08/2014	DT 120	SRR1966342
H143220689	34	Human	05/08/2014	DT 193	SRR1969507
H143220691	34	Human	05/08/2014	DT 193	SRR1969207
H143220692	34	Human	05/08/2014	DT 104b	SRR1969299
H143220695	19	Human	05/08/2014	DT 1	SRR1966283
H143220698	34	Human	05/08/2014	DT 193	SRR1965965
H143220699	34	Human	05/08/2014	PT U311	SRR1968561
H143220700	34	Human	05/08/2014	PT U311	SRR1967917
H143220701	34	Cattle	05/08/2014	DT 7a	SRR1966580
H143220702	19	Cattle	05/08/2014	RDNC	SRR1968237
H143220707	34	Human	05/08/2014	DT 193	SRR1968695
H143220792	34	Human	05/08/2014	DT 193	SRR1969253
H143220793	19	Human	05/08/2014	DT 104b	SRR1967149
H143240626	34	Human	06/08/2014	DT 193	SRR1967096
H143240631	19	Human	06/08/2014	DT 8	SRR1969879
H143240635	19	Human	06/08/2014	PT U312	SRR1967461
H143240637	19	Human	06/08/2014	PT U312	SRR1965508
H143240640	34	Frozen raw chicken	06/08/2014	DT 193	SRR1965402
H143240642	19	Human	06/08/2014	PT U302	SRR1970132
H143240643	34	Human	06/08/2014	DT 193	SRR1969637
H143240644	34	Human	06/08/2014	DT 193	SRR1965343
H143240645	34	Human	06/08/2014	None	SRR1968160
H143240647	34	Human	06/08/2014	DT 120	SRR1965461
H143240650	19	Human	06/08/2014	RDNC	SRR1966147
H143240651	19	Human	06/08/2014	DT 101	SRR1966638
H143240683	34	Human	06/08/2014	DT 193	SRR1965061
H143240685	19	Human	06/08/2014	DT 2	SRR1969986
H143260616	34	Human	07/08/2014	DT 193	SRR1965099
H143260617	34	Human	07/08/2014	None	SRR1969405

H143260618	34	Human	07/08/2014	None	SRR1967715
H143260626	19	Human	07/08/2014	DT 193	SRR1966982
H143260627	19	Human	07/08/2014	DT 104	SRR1970255
H143260629	34	Human	07/08/2014	DT 193	SRR1965388
H143260631	19	Human	07/08/2014	DT 10	SRR1970252
H143260632	19	Human	07/08/2014	PT U276	SRR1966471
H143260633	34	Human	07/08/2014	DT 120	SRR1970136
H143280585	313	Human	08/08/2014	DT 2	SRR1966418
H143280585	313	Human	08/08/2014	DT 2	SRR1968052
H143280587	19	Human	08/08/2014	DT 101	SRR3049578
H143280603	34	Human	08/08/2014	DT 193	SRR1965953
H143280604	19	Human	08/08/2014	PT U302	SRR1966203
H143280605	34	Human	08/08/2014	DT 193	SRR1965944
H143280606	19	Human	08/08/2014	DT 104	SRR1969823
H143280607	34	Human	08/08/2014	DT 193	SRR1969766
H143280610	34	Human	08/08/2014	DT 193	SRR1966066
H143300400	34	Human	11/08/2014	DT 193	SRR3048695
H143300401	34	Human	11/08/2014	DT 193	SRR3049584
H143300402	34	Human	11/08/2014	PT U323	SRR1968378
H143300403	19	Human	11/08/2014	DT 12	SRR1967704
H143300404	34	Human	11/08/2014	DT 193	SRR1969214
H143300405	34	Human	11/08/2014	DT 120	SRR1965312
H143320328	19	Human	12/08/2014	None	SRR1970108
H143320447	568	Human	12/08/2014	DT 56	SRR1969075
H143320448	34	Human	12/08/2014	DT 193	SRR1967112
H143320465	34	Human	12/08/2014	DT 193	SRR3049201
H143320471	19	Human	12/08/2014	DT 104	SRR1968638
H143320472	19	Human	12/08/2014	RDNC	SRR1968536
H143320490	34	Human	12/08/2014	PT 4a	SRR1969389
H143340483	19	Human	13/08/2014	None	SRR1968290
H143340495	34	Human	13/08/2014	DT 193	SRR1958233
H143340496	19	Human	13/08/2014	PT U331	SRR1968373
H143340500	34	Human	13/08/2014	None	SRR1965940
H143340502	34	Human	13/08/2014	DT 104b	SRR1965816
H143340503	34	Human	13/08/2014	DT 193	SRR1969166
H143340505	34	Human	13/08/2014	DT 193	SRR1968291
H143340506	34	Human	13/08/2014	PT U329	SRR1966456

H143340515	34	Human	13/08/2014	None	SRR1965096
H143340517	34	Human	13/08/2014	DT 193	SRR1967577
H143340518	19	Human	13/08/2014	DT 104b	SRR1966047
H143340519	34	Human	13/08/2014	DT 193	SRR1967679
H143340521	19	Human	13/08/2014	DT 8	SRR1970049
H143340522	19	Human	13/08/2014	DT 8	SRR1968600
H143360556	34	Human	14/08/2014	DT 193	SRR1968092
H143380360	34	Human	15/08/2014	DT 193	SRR1968071
H143380438	34	Human	15/08/2014	PT U323	SRR1968204
H143380440	34	Human	15/08/2014	DT 193	SRR1966734
H143380441	34	Human	15/08/2014	DT 7a	SRR1965644
H143380442	34	Human	15/08/2014	None	SRR1966466
H143400432	34	Human	18/08/2014	PT U323	SRR1969866
H143400433	34	Human	18/08/2014	DT 193	SRR1969119
H143400434	34	Human	18/08/2014	PT U323	SRR1966414
H143400436	34	Human	18/08/2014	DT 195	SRR1968805
H143400440	19	Human	18/08/2014	DT 8	SRR1966178
H143400441	19	Human	18/08/2014	RDNC	SRR1969848
H143400442	34	Human	18/08/2014	None	SRR1966308
H143400481	34	Human	18/08/2014	DT 193	SRR1967173
H143420455	19	Human	19/08/2014	DT 135	SRR3049537
H143420586	34	Human	19/08/2014	None	SRR1968776
H143420588	34	Human	19/08/2014	None	SRR1966859
H143420589	19	Human	19/08/2014	None	SRR1966642
H143420590	34	Human	19/08/2014	DT 193	SRR1968187
H143420594	19	Human	19/08/2014	DT 193	SRR1967731
H143420595	34	Human	19/08/2014	DT 193	SRR1967281
H143420596	19	Human	19/08/2014	DT 101	SRR1967982
H143420602	19	Human	19/08/2014	DT 101	SRR1966673
H143420603	34	Human	19/08/2014	DT 193	SRR1966359
H143420605	34	Human	19/08/2014	DT 193	SRR1965582
H143440341	19	Human	20/08/2014	DT 193	SRR1965754
H143440347	34	Human	20/08/2014	None	SRR1969966
H143440350	19	Human	20/08/2014	DT 1	SRR1968852
H143440352	19	Human	20/08/2014	DT 208	SRR1966932
H143480449	34	Human	22/08/2014	PT U311	SRR1965472
H143520409	34	Human	26/08/2014	DT 193	SRR1958557

H143520413	34	Human	26/08/2014	DT 193	SRR1957881
H143520415	34	Human	26/08/2014	DT 193	SRR1969218
H143520418	19	Human	26/08/2014	DT 193	SRR1966499
H143520444	19	Human	26/08/2014	DT 49	SRR1967536
H143520446	19	Human	26/08/2014	RDNC	SRR1966257
H143520449	34	Human	26/08/2014	DT 120	SRR3048545
H143540763	19	Human	27/08/2014	DT 8	SRR1958434
H143540847	34	Human	27/08/2014	DT 193	SRR1966826
H143540860	19	Human	27/08/2014	DT 8	SRR1970276
H143540866	34	Human	27/08/2014	DT 120	SRR1969185
H143540867	19	Human	27/08/2014	DT 135	SRR1967606
H143540869	34	Human	27/08/2014	DT 193	SRR1958545
H143540869	34	Human	27/08/2014	DT 193	SRR1965597
H143540870	34	Human	27/08/2014	DT 193	SRR1968398
H143540872	19	Human	27/08/2014	DT 116	SRR1967087
H143540873	19	Human	27/08/2014	DT 135	SRR3049318
H143540874	34	Human	27/08/2014	DT 193	SRR1966840
H143540875	19	Human	27/08/2014	DT 177	SRR1967295
H143540876	568	Cat	27/08/2014	DT 56	SRR1968278
H143540877	19	Canary	27/08/2014	PT U308	SRR1967460
H143540878	19	Pig	27/08/2014	PT U308	SRR1969730
H143540879	34	Human	27/08/2014	PT U323	SRR1970307
H143540881	34	Human	27/08/2014	DT 193	SRR1965832
H143540882	34	Human	27/08/2014	DT 193	SRR1965589
H143540883	34	Human	27/08/2014	DT 193	SRR1968458
H143540884	19	Human	27/08/2014	PT U302	SRR1968343
H143540885	19	Human	27/08/2014	DT 135	SRR1968668
H143540899	19	Human	27/08/2014	PT U312	SRR1963379
H143540899	19	Human	27/08/2014	PT U312	SRR1967051
H143560431	34	Human	28/08/2014	DT 193	SRR1970076
H143560436	34	Human	28/08/2014	DT 193	SRR1968981
H143560439	34	Human	28/08/2014	DT 193	SRR1968113
H143560440	34	Human	28/08/2014	None	SRR1965780
H143560441	19	Human	28/08/2014	DT 193	SRR1965812
H143560443	34	Human	28/08/2014	PT U311	SRR1968126
H143560444	34	Human	28/08/2014	PT U323	SRR1967236
H143580507	34	Human	29/08/2014	None	SRR1957883

H143580513	19	Human	29/08/2014	DT 1	SRR1958437
H143580522	19	Human	29/08/2014	DT 8	SRR1958243
H143580525	34	Human	29/08/2014	None	SRR1957998
H143580526	213	Human	29/08/2014	RDNC	SRR1958128
H143580530	34	Human	29/08/2014	None	SRR3049344
H143580533	19	Human	29/08/2014	DT 193	SRR1958267
H143580535	19	Human	29/08/2014	DT 193	SRR1967221
H143580536	19	Human	29/08/2014	DT 99	SRR1957908
H143580538	34	Human	29/08/2014	PT U311	SRR1957831
H143580621	19	Human	29/08/2014	PT U314	SRR1957829
H143600533	34	Human	01/09/2014	DT 193	SRR3049729
H143600534	34	Human	01/09/2014	DT 120	SRR1967810
H143600536	34	Human	01/09/2014	DT 193	SRR1969973
H143600539	19	Human	01/09/2014	DT 104	SRR1958631
H143600540	34	Human	01/09/2014	PT U311	SRR1958601
H143600541	34	Human	01/09/2014	PT U311	SRR1958201
H143600542	34	Human	01/09/2014	DT 193	SRR1958563
H143600543	34	Human	01/09/2014	PT U311	SRR1958443
H143600544	323	Human	01/09/2014	RDNC	SRR1969955
H143600562	19	Human	01/09/2014	DT 193	SRR1970296
H143620791	34	Human	02/09/2014	PT U311	SRR1966947
H143620791	34	Human	02/09/2014	PT U311	SRR3049255
H143620793	34	Human	02/09/2014	DT 193	SRR1969126
H143620794	19	Human	02/09/2014	DT 193	SRR1965314
H143620798	34	Human	02/09/2014	DT 120	SRR1968496
H143620798	34	Human	02/09/2014	DT 120	SRR3049022
H143640556	34	Human	03/09/2014	DT 193	SRR1968416
H143640578	19	Human	03/09/2014	DT 104b	SRR1968898
H143640578	19	Human	03/09/2014	DT 104b	SRR3049314
H143640580	34	Human	03/09/2014	DT 193	SRR1965203
H143640580	34	Human	03/09/2014	DT 193	SRR3048843
H143640580	34	Human	03/09/2014	DT 193	SRR3284842
H143640583	34	Human	03/09/2014	DT 193	SRR1966773
H143640584	19	Human	03/09/2014	None	SRR3284944
H143640585	19	Human	03/09/2014	None	SRR1958404
H143640585	19	Human	03/09/2014	None	SRR3048705
H143660447	34	Human	04/09/2014	DT 193	SRR1960212

H143660449	34	Human	04/09/2014	DT 193	SRR1967114
H143660450	19	Human	04/09/2014	DT 120	SRR1958481
H143660462	19	Human	04/09/2014	DT 193	SRR1963093
H143660470	34	Human	04/09/2014	DT 193	SRR1959296
H143660473	34	Human	04/09/2014	DT 193	SRR1963458
H143660478	19	Human	04/09/2014	DT 135	SRR1958518
H143660478	19	Human	04/09/2014	DT 135	SRR1969896
H143660479	34	Human	04/09/2014	PT U323	SRR1958538
H143660479	34	Human	04/09/2014	PT U323	SRR1969457
H143660507	34	Human	04/09/2014	DT 193	SRR1965903
H143660508	34	Human	04/09/2014	DT 120	SRR1967287
H143680693	34	Human	05/09/2014	DT 120	SRR1957803
H143680693	34	Human	05/09/2014	DT 120	SRR1957990
H143680695	34	Human	05/09/2014	DT 193	SRR1958399
H143680695	34	Human	05/09/2014	DT 193	SRR3049120
H143680696	34	Human	05/09/2014	DT 193	SRR1958629
H143680696	34	Human	05/09/2014	DT 193	SRR3049850
H143700616	19	Human	08/09/2014	DT 101	SRR1968491
H143700617	34	Human	08/09/2014	DT 193	SRR1957778
H143700618	34	Human	08/09/2014	DT 193	SRR1957779
H143700633	34	Human	08/09/2014	PT U311	SRR1957785
H143720723	19	Human	09/09/2014	PT U312	SRR1969870
H143720725	19	Human	09/09/2014	DT 104	SRR1969947
H143720758	19	Human	09/09/2014	DT 193	SRR1968980
H143720760	19	Human	09/09/2014	DT 2	SRR1968871
H143720761	34	Human	09/09/2014	DT 193	SRR1968716
H143720762	34	Human	09/09/2014	DT 193	SRR1965898
H143720763	34	Human	09/09/2014	None	SRR1968178
H143720765	34	Human	09/09/2014	DT 193	SRR1967556
H143720766	34	Human	09/09/2014	DT 193	SRR1966993
H143720768	34	Human	09/09/2014	DT 193	SRR1969486
H143720769	34	Human	09/09/2014	DT 193	SRR1967499
H143740686	213	Human	10/09/2014	DT 81	SRR1957801
H143740703	19	Human	10/09/2014	None	SRR1957724
H143740704	34	Pig	10/09/2014	DT 193	SRR1957753
H143740705	34	Pig	10/09/2014	DT 193	SRR1967290
H143740708	34	Human	10/09/2014	DT 120	SRR1958570

H143740710	19	Human	10/09/2014	DT 193	SRR1965598
H143740713	19	Human	10/09/2014	PT U311	SRR1969097
H143740716	19	Human	10/09/2014	DT 44	SRR1970113
H143740718	34	Human	10/09/2014	DT 193	SRR1969540
H143740719	34	Human	10/09/2014	PT U311	SRR1966303
H143740720	34	Human	10/09/2014	PT U311	SRR1969666
H143740721	213	Human	10/09/2014	DT 80	SRR1968540
H143740722	34	Human	10/09/2014	DT 193	SRR1966306
H143740724	34	Human	10/09/2014	DT 193	SRR1966517
H143740725	19	Human	10/09/2014	DT 99	SRR1965815
H143740729	34	Unknown	10/09/2014	DT 193	SRR1957966
H143760527	34	Human	11/09/2014	DT 193	SRR1969791
H143760573	19	Human	11/09/2014	Untypable	SRR1966880
H143760577	19	Human	11/09/2014	DT 41b	SRR1969437
H143760578	34	Human	11/09/2014	None	SRR1967339
H143760579	19	Human	11/09/2014	DT 193	SRR1968338
H143760580	19	Human	11/09/2014	DT 193	SRR1965896
H143760581	34	Human	11/09/2014	DT 193	SRR1970246
H143760583	19	Human	11/09/2014	None	SRR1966034
H143760585	19	Human	11/09/2014	DT 8	SRR1958371
H143760590	34	Human	11/09/2014	DT 193	SRR1957858
H143760594	34	Human	11/09/2014	PT 8	SRR1958203
H143780536	34	Human	12/09/2014	DT 193	SRR1958084
H143780537	19	Human	12/09/2014	PT U302	SRR1957861
H143780538	213	Human	12/09/2014	DT 80	SRR1957825
H143780539	34	Human	12/09/2014	DT 120	SRR1958181
H143780540	34	Human	12/09/2014	DT 193	SRR1957989
H143780543	19	Human	12/09/2014	DT 1	SRR1958214
H143780564	19	Human	12/09/2014	DT 110	SRR1958670
H143800497	19	Human	15/09/2014	DT 193	SRR1958502
H143800498	19	Human	15/09/2014	PT U291	SRR1958173
H143800499	19	Human	15/09/2014	DT 46	SRR1958012
H143800500	34	Human	15/09/2014	DT 193	SRR1958472
H143800502	34	Human	15/09/2014	DT 193	SRR1958591
H143800504	19	Human	15/09/2014	DT 46	SRR1958459
H143820790	19	Human	16/09/2014	DT 8	SRR1965054
H143820791	19	Human	16/09/2014	DT 193	SRR1969106

H143820792	34	Human	16/09/2014	PT U323	SRR1969100
H143820794	19	Human	16/09/2014	PT U312	SRR1965135
H143820795	19	Human	16/09/2014	DT 104	SRR1966540
H143820797	19	Human	16/09/2014	DT 193	SRR1965373
H143820842	19	Human	16/09/2014	Untypable	SRR1967527
H143840498	323	Human	17/09/2014	PT U277	SRR1969122
H143840570	19	Human	17/09/2014	DT 46a	SRR1967874
H143840571	213	Human	17/09/2014	DT 80	SRR1968141
H143840575	34	Human	17/09/2014	None	SRR1967200
H143840576	19	Human	17/09/2014	PT U302	SRR1969433
H143860639	2078	Human	18/09/2014	DT 195	SRR1967107
H143860655	19	Human	18/09/2014	PT U312	SRR1968872
H143860656	34	Human	18/09/2014	PT U311	SRR1968177
H143860657	34	Human	18/09/2014	DT 120	SRR1967677
H143860658	34	Human	18/09/2014	DT 120	SRR1965418
H143880696	19	Human	19/09/2014	PT U311	SRR1969376
H143880701	323	Human	19/09/2014	PT U277	SRR1969726
H143880703	34	Human	19/09/2014	DT 193	SRR1969450
H143880705	19	Human	19/09/2014	DT 120	SRR1966722
H143880724	34	Pig	19/09/2014	DT 193	SRR1963501
H143880725	34	Pig	19/09/2014	DT 193	SRR1959251
H143880726	34	Pig	19/09/2014	DT 193	SRR1959255
H143900512	34	Human	22/09/2014	DT 193	SRR1963341
H143900570	34	Human	22/09/2014	DT 193	SRR1963523
H143900571	34	Human	22/09/2014	DT 193	SRR1960169
H143900574	19	Human	22/09/2014	PT U308	SRR1959284
H143920789	34	Human	23/09/2014	DT 193	SRR1958132
H143920792	19	Human	23/09/2014	RDNC	SRR1958015
H143920793	34	Human	23/09/2014	DT 7a	SRR1958465
H143920796	19	Human	23/09/2014	DT 104b	SRR1957971
H143920797	19	Human	23/09/2014	PT U302	SRR1958026
H143920800	34	Human	23/09/2014	PT U311	SRR1958024
H143920801	34	Human	23/09/2014	PT U311	SRR1958221
H143920802	34	Human	23/09/2014	DT 193	SRR1958048
H143920803	19	Human	23/09/2014	DT 101	SRR1958588
H143940533	34	Human	24/09/2014	DT 193	SRR1958285
H143940609	34	Human	24/09/2014	DT 193	SRR1958116

H143940629	34	Human	24/09/2014	DT 193	SRR1958596
H143940630	19	Human	24/09/2014	DT 8	SRR1958470
H143940635	34	Human	24/09/2014	DT 193	SRR1958071
H143940637	34	Human	24/09/2014	DT 193	SRR1958316
H143940638	34	Human	24/09/2014	DT 120	SRR1958331
H143940639	34	Human	24/09/2014	DT 193	SRR1958075
H143960298	34	Human	25/09/2014	PT U323	SRR1958449
H143960356	19	Human	25/09/2014	DT 104	SRR1957900
H143960357	34	Human	25/09/2014	DT 193	SRR1957940
H143960359	34	Human	25/09/2014	PT U323	SRR1958458
H143960360	19	Human	25/09/2014	PT U302	SRR1958444
H143960363	34	Human	25/09/2014	DT 193	SRR1958034
H143960367	19	Human	25/09/2014	PT U312	SRR1958189
H143980708	34	Human	26/09/2014	DT 193	SRR1958133
H143980710	34	Human	26/09/2014	DT 193	SRR1958346
H143980721	34	Neck flap	26/09/2014	DT 193	SRR1958016
H143980724	34	Neck flap	26/09/2014	DT 193	SRR1957915
H143980725	34	Neck flap	26/09/2014	DT 193	SRR1958548
H143980726	34	Neck flap	26/09/2014	DT 193	SRR1958185
H143980747	34	Human	26/09/2014	DT 120	SRR1958620
H143980748	19	Human	26/09/2014	PT U302	SRR1957902
H144000584	34	Human	29/09/2014	DT 193	SRR1966591
H144000589	19	Human	29/09/2014	DT 101	SRR1967939
H144000591	313	Human	29/09/2014	DT 2	SRR1969854
H144000593	19	Human	29/09/2014	DT 1	SRR1966367
H144000594	34	Human	29/09/2014	DT 120	SRR1967636
H144020656	34	Human	30/09/2014	DT 195	SRR3048796
H144020667	34	Human	30/09/2014	DT 193	SRR1967685
H144020673	19	Human	30/09/2014	RDNC	SRR1967884
H144020708	19	Human	30/09/2014	PT U302	SRR1958686
H144020710	34	Human	30/09/2014	DT 193	SRR1958355
H144020711	34	Human	30/09/2014	DT 193	SRR3049350
H144020714	34	Human	30/09/2014	None	SRR1958577
H144020715	19	Human	30/09/2014	DT 104	SRR3049950
H144020716	34	Human	30/09/2014	DT 193	SRR1958107
H144020717	19	Human	30/09/2014	DT 104	SRR1970277
H144020720	34	Human	30/09/2014	PT U323	SRR1958223

H144020722	19	Human	30/09/2014	DT 101	SRR3049485
H144040605	19	60.23504	01/10/2014	DT 8	SRR1966062
H144040606	19	60.23503	01/10/2014	DT 8	SRR1969552
H144040612	19	Human	01/10/2014	DT 104	SRR3048701
H144040613	34	Human	01/10/2014	DT 193	SRR1966196
H144040615	34	Human	01/10/2014	DT 193	SRR1968804
H144040617	34	Human	01/10/2014	DT 193	SRR3048873
H144040620	19	Human	01/10/2014	DT 104	SRR3049029
H144040621	19	Human	01/10/2014	DT 104	SRR3049557
H144040622	34	Human	01/10/2014	DT 193	SRR1965178
H144060558	19	Human	02/10/2014	PT U302	SRR1966019
H144060572	19	Human	02/10/2014	PT U302	SRR1958281
H144060575	19	Human	02/10/2014	DT 104	SRR3049551
H144060576	19	Human	02/10/2014	PT U302	SRR3048762
H144060577	34	Human	02/10/2014	PT U323	SRR1958513
H144080387	19	Human	03/10/2014	DT 104	SRR3049580
H144080426	19	Human	03/10/2014	None	SRR1965208
H144080428	19	Human	03/10/2014	None	SRR3049353
H144080429	19	Human	03/10/2014	None	SRR1966606
H144080430	34	Human	03/10/2014	DT 193	SRR1965771
H144080431	34	Human	03/10/2014	DT 193	SRR1969414
H144100484	19	Human	06/10/2014	DT 104	SRR1966620
H144100490	19	Human	06/10/2014	DT 161	SRR1966654
H144120507	34	Human	07/10/2014	DT 193	SRR1969452
H144120530	19	Human	07/10/2014	DT 191a	SRR1968848
H144120537	34	Human	07/10/2014	DT 193	SRR1970107
H144120538	19	Human	07/10/2014	None	SRR1965166
H144120539	34	Human	07/10/2014	PT U323	SRR1965737
H144120542	19	Human	07/10/2014	DT 41b	SRR1967722
H144120543	19	Human	07/10/2014	PT U302	SRR1967613
H144120556	19	Human	07/10/2014	DT 1	SRR1958394
H144120557	19	Human	07/10/2014	DT 104	SRR1957899
H144120558	19	Human	07/10/2014	PT U302	SRR3048519
H144120559	34	Human	07/10/2014	PT U323	SRR1957824
H144140610	19	Human	08/10/2014	DT 104	SRR1958306
H144140619	19	Human	08/10/2014	DT 193	SRR1958177
H144140627	34	Human	08/10/2014	PT U311	SRR1958304

H144140628	34	Human	08/10/2014	None	SRR1958556
H144140631	34	Human	08/10/2014	DT 195	SRR1958194
H144140632	34	Human	08/10/2014	DT 193	SRR1958260
H144140634	19	Human	08/10/2014	DT 104	SRR1958688
H144140636	19	Human	08/10/2014	PT U302	SRR1967124
H144140639	34	Human	08/10/2014	PT U323	SRR1958202
H144160425	19	Human	09/10/2014	DT 15a	SRR1958405
H144160430	34	Human	09/10/2014	DT 193	SRR1957818
H144160459	34	Human	09/10/2014	DT 120	SRR1967188
H144160463	34	Human	09/10/2014	PT U323	SRR1958237
H144160466	19	Human	09/10/2014	PT U302	SRR1969817
H144160467	34	Human	09/10/2014	None	SRR1966450
H144160468	19	Human	09/10/2014	DT 99	SRR1970247
H144160469	19	Human	09/10/2014	DT 9	SRR1968393
H144180582	19	Human	10/10/2014	DT 104	SRR1970208
H144180586	313	Human	10/10/2014	DT 2	SRR1967447
H144180588	19	Human	10/10/2014	DT 41b	SRR1958320
H144180590	19	Human	10/10/2014	PT U302	SRR1967565
H144180592	19	Human	10/10/2014	DT 193	SRR1967069
H144200657	19	Human	13/10/2014	DT 208	SRR1958288
H144200658	19	Human	13/10/2014	DT 1	SRR1958467
H144200659	34	Human	13/10/2014	DT 7a	SRR1958440
H144200661	34	Human	13/10/2014	DT 193	SRR1958053
H144220658	19	Human	14/10/2014	PT U302	SRR1957885
H144220659	19	Human	14/10/2014	Untypable	SRR1958317
H144220661	313	Human	14/10/2014	None	SRR1958036
H144220662	34	Human	14/10/2014	PT U323	SRR1958344
H144220667	34	Human	14/10/2014	DT 193	SRR1958388
H144220669	34	Human	14/10/2014	DT 193	SRR1958645
H144220671	34	Human	14/10/2014	DT 120	SRR1957846
H144220672	19	Human	14/10/2014	DT 8	SRR1958298
H144220673	34	Human	14/10/2014	DT 193	SRR1958146
H144220674	34	Human	14/10/2014	PT U311	SRR1958174
H144220675	19	Human	14/10/2014	DT 208	SRR1968809
H144220712	19	Green finch	14/10/2014	PT U277	SRR1969901
H144220720	19	Human	14/10/2014	DT 208	SRR1957884
H144220720	19	Human	14/10/2014	DT 208	SRR1958209

H144240607	19	Human	15/10/2014	DT 104b	SRR1958407
H144240609	34	Human	15/10/2014	DT 193	SRR1965899
H144240612	19	Human	15/10/2014	DT 104	SRR1957860
H144240613	34	Human	15/10/2014	DT 193	SRR1969770
H144240614	19	Human	15/10/2014	DT 193	SRR1967869
H144240634	19	Human	15/10/2014	DT 135	SRR1958652
H144240637	19	Human	15/10/2014	DT 208	SRR1966399
H144240638	19	Human	15/10/2014	DT 208	SRR1969270
H144260658	19	Human	16/10/2014	DT 208	SRR1966270
H144260669	19	Human	16/10/2014	PT U288	SRR1958402
H144260673	19	Human	16/10/2014	DT 104	SRR1966000
H144260674	19	Human	16/10/2014	DT 8	SRR1958120
H144260675	19	Human	16/10/2014	DT 193	SRR3048893
H144260676	19	Human	16/10/2014	RDNC	SRR1957906
H144280521	34	Human	17/10/2014	DT 193	SRR1967423
H144280569	19	Human	17/10/2014	DT 101	SRR1969748
H144280570	34	Human	17/10/2014	None	SRR3049622
H144280571	19	Human	17/10/2014	DT 104	SRR1969641
H144300569	34	Human	20/10/2014	DT 193	SRR3049202
H144300576	34	Human	20/10/2014	PT U311	SRR1965878
H144300578	34	Human	20/10/2014	PT U311	SRR1967258
H144300579	34	Human	20/10/2014	PT U311	SRR3049605
H144320727	34	Human	21/10/2014	DT 120	SRR1969352
H144320757	34	Human	21/10/2014	DT 193	SRR1968201
H144320758	19	Human	21/10/2014	DT 193	SRR1969387
H144320759	19	Human	21/10/2014	PT U302	SRR1968572
H144320760	19	Human	21/10/2014	DT 101	SRR1970298
H144320771	34	Human	21/10/2014	DT 195	SRR1966505
H144320772	34	Human	21/10/2014	PT U323	SRR1969508
H144320773	34	Human	21/10/2014	PT U323	SRR1966187
H144340671	34	Human	22/10/2014	DT 193	SRR1967627
H144340706	34	Human	22/10/2014	PT U311	SRR1966137
H144340707	19	Human	22/10/2014	DT 96	SRR1969429
H144340709	19	Human	22/10/2014	DT 104	SRR1966230
H144340709	19	Human	22/10/2014	DT 104	SRR1967674
H144340710	34	Human	22/10/2014	PT U323	SRR1965661
H144340711	34	Human	22/10/2014	DT 193	SRR1960180

H144340712	19	Human	22/10/2014	DT 1	SRR1960366
H144340714	34	Human	22/10/2014	DT 193	SRR1959223
H144340715	19	Human	22/10/2014	DT 93	SRR1959513
H144340716	19	Human	22/10/2014	DT 193	SRR1960075
H144340718	34	Human	22/10/2014	DT 193	SRR1963473
H144340719	34	Human	22/10/2014	PT U311	SRR1960310
H144360255	34	Human	23/10/2014	DT 193	SRR1958093
H144360256	19	Human	23/10/2014	PT U302	SRR1958124
H144360257	34	Human	23/10/2014	DT 193	SRR1958639
H144360258	19	Human	23/10/2014	PT U302	SRR1958061
H144360259	19	Human	23/10/2014	RDNC	SRR1963474
H144380544	313	Human	24/10/2014	DT 2	SRR1958303
H144380574	19	Human	24/10/2014	PT U311	SRR1958313
H144380583	19	NCTC11881B2D	24/10/2014	None	SRR1968588
H144380585	34	Human	24/10/2014	PT U323	SRR1958609
H144380586	34	Human	24/10/2014	PT U323	SRR1958483
H144400569	34	Human	27/10/2014	DT 193	SRR1965199
H144400598	34	Human	27/10/2014	DT 193	SRR1963406
H144400600	19	Human	27/10/2014	PT U302	SRR1969320
H144400601	34	Human	27/10/2014	DT 193	SRR1959385
H144420807	19	Human	28/10/2014	DT 208	SRR1968842
H144420810	19	Human	28/10/2014	DT 2	SRR1966604
H144420811	19	Human	28/10/2014	DT 193	SRR1969156
H144420896	19	Human	28/10/2014	DT 2	SRR1968712
H144420905	19	Human	28/10/2014	DT 104	SRR1966530
H144420906	19	Human	28/10/2014	DT 2	SRR1969649
H144420908	34	Human	28/10/2014	DT 193	SRR1965768
H144420910	34	Human	28/10/2014	DT 120	SRR1968753
H144420913	34	Human	28/10/2014	DT 193	SRR1966700
H144420914	34	Human	28/10/2014	DT 193	SRR1965627
H144420915	34	Human	28/10/2014	DT 193	SRR1965659
H144420916	34	Human	28/10/2014	DT 193	SRR1969683
H144420918	34	Human	28/10/2014	PT U329	SRR1965153
H144440743	34	Human	29/10/2014	DT 193	SRR1970010
H144440744	34	Human	29/10/2014	DT 193	SRR1958625
H144440750	34	Human	29/10/2014	DT 193	SRR1969217
H144440753	19	Human	29/10/2014	PT U288	SRR1969625

H144440765	34	Human	29/10/2014	DT 193	SRR1958100
H144460488	19	Human	30/10/2014	Untypable	SRR1957895
H144480567	34	Human	31/10/2014	DT 193	SRR1958532
H144480568	34	Human	31/10/2014	DT 193	SRR1958473
H144480569	34	Human	31/10/2014	DT 120	SRR1958415
H144500515	19	Human	03/11/2014	Untypable	SRR1957832
H144500519	19	Human	03/11/2014	PT U302	SRR1957965
H144500533	19	Human	03/11/2014	PT U302	SRR1969974
H144500534	19	Human	03/11/2014	DT 193	SRR1958241
H144500535	19	Human	03/11/2014	PT U302	SRR1965045
H144520686	19	Human	04/11/2014	DT 193	SRR1965745
H144520687	34	Human	04/11/2014	DT 193	SRR1969456
H144520703	19	Human	04/11/2014	DT 1	SRR1967739
H144520714	34	Human	04/11/2014	PT U323	SRR1968460
H144520715	19	Human	04/11/2014	DT 2	SRR1967918
H144520716	19	Human	04/11/2014	DT 104	SRR1965513
H144520719	19	Human	04/11/2014	DT 193	SRR1970121
H144520720	19	Human	04/11/2014	PT U312	SRR1967130
H144520723	34	Human	04/11/2014	DT 193	SRR1965393
H144520724	34	Human	04/11/2014	DT 193	SRR1965996
H144520725	34	Human	04/11/2014	DT 193	SRR1966847
H144540567	19	Human	05/11/2014	None	SRR1969628
H144540642	568	Human	05/11/2014	DT 56	SRR1967749
H144540644	34	Human	05/11/2014	DT 193	SRR1966366
H144540652	34	Human	05/11/2014	DT 193	SRR1968784
H144540654	34	Human	05/11/2014	DT 193	SRR1965421
H144540655	19	Human	05/11/2014	PT U302	SRR1965888
H144540656	19	Human	05/11/2014	DT 132	SRR1970138
H144540658	34	Human	05/11/2014	DT 193	SRR1968376
H144540660	19	Human	05/11/2014	DT 2a	SRR1969560
H144560569	19	Human	06/11/2014	DT 104	SRR1967389
H144560589	19	Unknown	06/11/2014	DT 132	SRR1958205
H144560590	34	Human	06/11/2014	DT 193	SRR1968676
H144560591	34	Human	06/11/2014	PT U329	SRR1969244
H144580647	34	Human	07/11/2014	DT 193	SRR1958004
H144580651	34	Human	07/11/2014	PT U311	SRR1958552
H144580657	34	Human	07/11/2014	PT 21	SRR1969960

H144600638	34	Human	10/11/2014	DT 193	SRR1965977
H144600640	2083	Human	10/11/2014	RDNC	SRR1965215
H144600656	19	Human	10/11/2014	DT 193	SRR1966929
H144600660	34	Human	10/11/2014	DT 193	SRR1970183
H144620685	34	Human	11/11/2014	DT 193	SRR1969910
H144620695	2083	Human	11/11/2014	DT 193	SRR1965712
H144620698	19	Human	11/11/2014	RDNC	SRR1969675
H144620699	34	Human	11/11/2014	DT 193	SRR1968504
H144620700	19	Human	11/11/2014	DT 2a	SRR1965814
H144620701	19	Human	11/11/2014	DT 1	SRR1968133
H144620718	34	Human	11/11/2014	None	SRR1967074
H144620719	34	Human	11/11/2014	DT 193	SRR1966170
H144620721	34	Human	11/11/2014	PT 11	SRR1966235
H144620722	34	Human	11/11/2014	DT 193	SRR1966429
H144640572	19	Unknown	12/11/2014	DT 132	SRR1960218
H144640573	19	Unknown	12/11/2014	DT 132	SRR1959469
H144640576	34	Human	12/11/2014	DT 193	SRR1965820
H144640577	19	Human	12/11/2014	PT U312	SRR1968717
H144640578	34	Human	12/11/2014	DT 193	SRR1966189
H144640579	34	Human	12/11/2014	DT 193	SRR1965469
H144640580	34	Human	12/11/2014	DT 120	SRR1963225
H144640583	19	Human	12/11/2014	DT 104b	SRR1965244
H144640584	2083	Human	12/11/2014	PT U288	SRR1966841
H144640585	34	Human	12/11/2014	DT 120	SRR1965928
H144660684	19	Human	13/11/2014	RDNC	SRR1960165
H144660687	19	Unknown	13/11/2014	DT 132	SRR1960382
H144660698	34	Human	13/11/2014	DT 193	SRR1963444
H144660699	19	Human	13/11/2014	PT U281	SRR1963380
H144660700	19	Human	13/11/2014	PT U302	SRR1963146
H144660701	34	Human	13/11/2014	PT U323	SRR1960285
H144680508	34	Human	14/11/2014	DT 193	SRR1967720
H144680543	19	Human	14/11/2014	None	SRR1968559
H144680544	34	Human	14/11/2014	None	SRR1965479
H144680548	34	Human	14/11/2014	DT 193	SRR1968397
H144680549	34	Human	14/11/2014	DT 120	SRR3286904
H144680550	34	Human	14/11/2014	DT 193	SRR1965129
H144680551	19	Human	14/11/2014	DT 101	SRR3286577

H144700423	34	Human	17/11/2014	DT 193	SRR1969497
H144700424	19	Human	17/11/2014	RDNC	SRR1970192
H144700426	19	Human	17/11/2014	DT 132	SRR1968820
H144720736	34	Human	18/11/2014	DT 193	SRR1968140
H144720785	34	Human	18/11/2014	DT 193	SRR1960321
H144720834	34	Human	18/11/2014	None	SRR1968281
H144720840	34	Human	18/11/2014	DT 193	SRR1967138
H144720845	34	Human	18/11/2014	DT 193	SRR1963461
H144720847	34	Human	18/11/2014	DT 193	SRR1966998
H144740651	19	Human	19/11/2014	DT 2	SRR1960987
H144740658	19	Human	19/11/2014	DT 15a	SRR1959433
H144740702	34	Human	19/11/2014	None	SRR1963330
H144740709	34	Human	19/11/2014	DT 193	SRR1960040
H144740710	19	Human	19/11/2014	DT 8	SRR1962554
H144740711	19	Human	19/11/2014	None	SRR1960314
H144740712	19	Human	19/11/2014	RDNC	SRR1960264
H144740718	19	Human	19/11/2014	DT 1	SRR1961771
H144740719	34	Human	19/11/2014	DT 193	SRR1963091
H144740720	19	Human	19/11/2014	DT 101	SRR1959412
H144740723	34	Human	19/11/2014	DT 179	SRR1960364
H144760703	19	Unknown food	20/11/2014	DT 132	SRR1960290
H144760704	19	Unknown food	20/11/2014	DT 132	SRR1963332
H144760705	19	Unknown food	20/11/2014	DT 132	SRR1959486
H144760707	19	Unknown food	20/11/2014	DT 132	SRR1960320
H144760708	19	Human	20/11/2014	DT 191	SRR1963451
H144760719	19	Unknown food	20/11/2014	DT 132	SRR1960081
H144780621	34	Human	21/11/2014	DT 195	SRR1967413
H144780661	19	Unknown	21/11/2014	DT 132	SRR1969571
H144780666	19	Human	21/11/2014	DT 132	SRR1963525
H144780666	19	Human	21/11/2014	DT 132	SRR1969215
H144780673	34	Human	21/11/2014	DT 120	SRR1968944
H144800748	34	Human	24/11/2014	DT 193	SRR1967311
H144800782	34	Human	24/11/2014	PT U311	SRR1969476
H144800784	34	Human	24/11/2014	DT 120	SRR1965802
H144820727	34	Human	25/11/2014	PT 21	SRR3049426
H144820733	34	Human	25/11/2014	PT U311	SRR1959460
H144840757	34	Human	26/11/2014	DT 120	SRR1969008

H144840760	34	Human	26/11/2014	PT U311	SRR1968207
H144840761	34	Human	26/11/2014	DT 120	SRR1969058
H144840762	34	Human	26/11/2014	DT 193	SRR1966098
H144840763	34	Human	26/11/2014	DT 193	SRR1968936
H144840764	34	Human	26/11/2014	PT U311	SRR1969297
H144880865	19	Human	28/11/2014	DT 12	SRR1963243
H144880866	19	Human	28/11/2014	None	SRR1963337
H144880867	34	Human	28/11/2014	DT 193	SRR1960074
H144880869	19	Human	28/11/2014	DT 41b	SRR1960532
H144880871	34	Human	28/11/2014	DT 193	SRR1963085
H144880873	34	Human	28/11/2014	DT 120	SRR1959221
H144900572	19	Human	01/12/2014	DT 104	SRR1963318
H144900575	19	Human	01/12/2014	DT 161	SRR1959493
H144900578	34	Human	01/12/2014	PT U323	SRR1959248
H144920410	34	Human	02/12/2014	PT U311	SRR1959233
H144920411	34	Human	02/12/2014	PT U311	SRR1963478
H144920412	34	Human	02/12/2014	PT U311	SRR1959497
H144920413	34	Human	02/12/2014	PT U311	SRR1969799
H144920618	34	Human	02/12/2014	None	SRR1960128
H144920624	34	Human	02/12/2014	DT 120	SRR1963276
H144920627	323	Human	02/12/2014	RDNC	SRR1963331
H144920628	19	Human	02/12/2014	Untypable	SRR1960375
H144920630	34	Human	02/12/2014	DT 120	SRR1960185
H144920634	34	Human	02/12/2014	DT 193	SRR1959234
H144920635	34	Human	02/12/2014	DT 120	SRR1960902
H144920636	34	Human	02/12/2014	PT U311	SRR1960362
H144920637	568	Human	02/12/2014	DT 56	SRR1959499
H144940487	34	Human	03/12/2014	PT U311	SRR1965720
H144940490	19	Human	03/12/2014	DT 41	SRR1963083
H144940553	34	Human	03/12/2014	DT 120	SRR1969401
H144940556	34	Human	03/12/2014	DT 120	SRR1960687
H144940560	34	Human	03/12/2014	DT 193	SRR1963098
H144960545	34	Human	04/12/2014	DT 120	SRR1967893
H144960547	34	Human	04/12/2014	DT 120	SRR1965532
H144980592	19	Human	05/12/2014	DT 208	SRR1965583
H144980717	34	Human	05/12/2014	DT 193	SRR1969136
H144980718	34	Human	05/12/2014	None	SRR1969098

H144980731	34	Human	05/12/2014	DT 120	SRR1966459
H144980732	34	Human	05/12/2014	DT 120	SRR1970065
H144980733	34	Human	05/12/2014	DT 120	SRR1968062
H144980734	34	Human	05/12/2014	DT 120	SRR1968643
H144980735	34	Human	05/12/2014	DT 120	SRR1966817
H145000544	34	Human	08/12/2014	DT 20a	SRR1960296
H145000545	19	Human	08/12/2014	PT U317	SRR1966708
H145020733	19	Human	09/12/2014	DT 191a	SRR1959462
H145020735	34	Human	09/12/2014	DT 193	SRR1965342
H145040692	34	Human	10/12/2014	DT 193	SRR1967196
H145040715	34	Human	10/12/2014	PT U311	SRR1968020
H145040717	34	Human	10/12/2014	DT 120	SRR1963199
H145040718	34	Human	10/12/2014	DT 120	SRR1969059
H145040721	34	Human	10/12/2014	DT 120	SRR1968877
H145040722	34	Human	10/12/2014	DT 120	SRR1968647
H145040728	34	Human	10/12/2014	DT 193	SRR1967566
H145060536	34	Human	11/12/2014	DT 193	SRR1960162
H145060537	34	Human	11/12/2014	DT 193	SRR1963291
H145060538	34	Human	11/12/2014	None	SRR1963250
H145060720	19	Human	11/12/2014	DT 104	SRR1963412
H145060728	19	Human	11/12/2014	Untypable	SRR1963253
H145080408	34	Human	12/12/2014	DT 193	SRR1959473
H145080409	34	Human	12/12/2014	DT 193	SRR1963070
H145080411	34	Human	12/12/2014	DT 120	SRR1963407
H145080413	19	Human	12/12/2014	DT 193	SRR1959318
H145080422	19	Human	12/12/2014	DT 12	SRR1960220
H145100624	34	Human	15/12/2014	DT 120	SRR1960147
H145100673	34	Human	15/12/2014	DT 40	SRR1959453
H145100674	19	Human	15/12/2014	DT 193	SRR1963532
H145100676	19	Human	15/12/2014	DT 104	SRR1959472
H145100680	34	Human	15/12/2014	DT 193	SRR1963542
H145100686	34	Human	15/12/2014	DT 195	SRR1960122
H145100804	34	Human	15/12/2014	None	SRR1960189
H145120591	323	Human	16/12/2014	DT 104	SRR1965592
H145120596	19	Human	16/12/2014	DT 2	SRR1960140
H145120597	19	Human	16/12/2014	DT 2	SRR1963115
H145120612	34	Human	16/12/2014	DT 193	SRR1959727

H145120615	34	Human	16/12/2014	DT 193	SRR1963079
H145120616	34	Human	16/12/2014	DT 193	SRR1960385
H145120619	34	Human	16/12/2014	PT U329	SRR1961508
H145120620	19	Human	16/12/2014	DT 18	SRR1959425
H145120625	34	Human	16/12/2014	None	SRR1960187
H145140590	34	Human	17/12/2014	DT 193	SRR1967167
H145140591	19	Human	17/12/2014	PT U302	SRR1965276
H145140592	34	Human	17/12/2014	DT 193	SRR1963353
H145140594	19	Human	17/12/2014	DT 1	SRR1968444
H145140595	34	Human	17/12/2014	PT U311	SRR1967689
H145140599	34	Human	17/12/2014	PT U311	SRR1963465
H145140600	34	Human	17/12/2014	DT 120	SRR1963313
H145140601	34	Human	17/12/2014	DT 120	SRR1967703
H145160571	19	Human	18/12/2014	DT 1	SRR1966474
H145160574	19	Human	18/12/2014	DT 104	SRR1966843
H145160676	19	Human	18/12/2014	DT 10	SRR1969048
H145180570	34	Human	19/12/2014	PT U302	SRR1967265
H145180710	34	Human	19/12/2014	DT 7a	SRR1969975
H145180715	34	Human	19/12/2014	PT U311	SRR1966304
H145180717	34	Human	19/12/2014	DT 193	SRR1966795
H145180718	34	Human	19/12/2014	None	SRR3048764
H145180719	34	Human	19/12/2014	DT 193	SRR1965328
H145180720	34	Human	19/12/2014	PT U323	SRR1968387
H145200602	19	Human	22/12/2014	DT 41	SRR3049744
H145200606	34	Human	22/12/2014	PT U311	SRR1969977
H145200609	19	Human	22/12/2014	DT 2	SRR3049441
H145200611	34	Human	22/12/2014	None	SRR3048988
H145220846	34	Human	23/12/2014	DT 193	SRR1961873
H145220852	34	Human	23/12/2014	DT 193	SRR1963287
H145220853	19	Human	23/12/2014	DT 1	SRR1960036
H145240540	34	Human	24/12/2014	PT U311	SRR1960355
H145240594	34	Human	24/12/2014	RDNC	SRR1959274
H150120636	2212	Human	30/12/2014	DT 193	SRR1959451
H150120639	19	Human	30/12/2014	DT 135	SRR1963111
H150120641	34	Human	30/12/2014	DT 193	SRR1963537
H150120645	34	Human	30/12/2014	DT 193	SRR1963496
H150120668	19	Human	30/12/2014	DT 1	SRR1960045

H150140575	34	Human	31/12/2014	DT 193	SRR1960985
H150140577	34	Human	31/12/2014	DT 193	SRR1960327
H150200808	19	Human	05/01/2015	PT U288	SRR1962604
H150200809	34	Human	05/01/2015	DT 193	SRR1959504
H150200814	19	Human	05/01/2015	None	SRR1963082
H150220500	2103	Human	06/01/2015	RDNC	SRR1963446
H150220501	34	Human	06/01/2015	DT 120	SRR1960315
H150220502	34	Human	06/01/2015	DT 120	SRR1960358
H150220503	19	Human	06/01/2015	DT 120	SRR1959363
H150220505	34	Human	06/01/2015	DT 193	SRR1963491
H150220507	34	Human	06/01/2015	DT 193	SRR1959507
H150220512	34	Human	06/01/2015	None	SRR1963223
H150220513	34	Human	06/01/2015	DT 193	SRR1963356
H150220521	19	Human	06/01/2015	DT 41	SRR1963271
H150240436	34	Human	07/01/2015	DT 193	SRR1963360
H150240437	34	Human	07/01/2015	PT U311	SRR1960053
H150240438	19	Human	07/01/2015	DT 41	SRR1959400
H150240440	34	Human	07/01/2015	DT 120	SRR1960311
H150240441	34	Human	07/01/2015	DT 120	SRR1960033
H150260321	34	Human	08/01/2015	DT 193	SRR1959242
H150260323	34	Human	08/01/2015	DT 193	SRR1963514
H150260324	34	Human	08/01/2015	PT U323	SRR1960131
H150260326	19	Human	08/01/2015	DT 104	SRR1959478
H150260329	34	Raw poultry	08/01/2015	DT 193	SRR1959505
H150260336	19	Human	08/01/2015	DT 2	SRR1970120
H150280543	19	Human	09/01/2015	DT 104	SRR1960186
H150280548	34	Human	09/01/2015	DT 193	SRR1960127
H150300415	19	Human	12/01/2015	PT U331	SRR1963471
H150300545	34	Human	12/01/2015	DT 193	SRR1966887
H150300551	34	Human	12/01/2015	DT 193	SRR1967563
H150300552	34	Human	12/01/2015	PT U311	SRR1969432
H150320538	34	Human	13/01/2015	PT U311	SRR1960138
H150320614	19	Human	13/01/2015	DT 41	SRR1970235
H150320692	19	Human	13/01/2015	DT 104	SRR1963378
H150320697	19	Human	13/01/2015	RDNC	SRR1960982
H150320702	34	Human	13/01/2015	DT 193	SRR1960217
H150320703	19	Human	13/01/2015	DT 104	SRR1960252

H150320704	34	Human	13/01/2015	DT 193	SRR1966121
H150320705	34	Human	13/01/2015	DT 120	SRR1963369
H150320711	34	Human	13/01/2015	DT 193	SRR1963327
H150320714	19	Human	13/01/2015	PT U288	SRR1963513
H150320715	34	Human	13/01/2015	DT 193	SRR1960689
H150340498	19	Human	14/01/2015	RDNC	SRR1960686
H150340501	34	Human	14/01/2015	DT 137	SRR1959467
H150340503	313	Human	14/01/2015	DT 2	SRR1968084
H150380351	19	Human	16/01/2015	DT 104	SRR1963345
H150400536	34	Human	19/01/2015	RDNC	SRR1959420
H150400537	34	Human	19/01/2015	DT 193	SRR1959120
H150400539	19	Human	19/01/2015	Untypable	SRR1963222
H150400540	34	Human	19/01/2015	DT 193	SRR1959380
H150440669	34	Human	21/01/2015	DT 193	SRR1968612
H150440671	19	Human	21/01/2015	DT 193	SRR1966150
H150440673	34	Human	21/01/2015	DT 193	SRR1968234
H150440674	34	Human	21/01/2015	DT 193	SRR1969858
H150460715	34	Human	22/01/2015	PT U323	SRR1969842
H150460717	19	Human	22/01/2015	PT U288	SRR1968762
H150460720	34	Human	22/01/2015	DT 193	SRR1966454
H150460724	34	Human	22/01/2015	DT 193	SRR1966112
H150480665	19	Human	23/01/2015	DT 104	SRR1967417
H150480666	34	Human	23/01/2015	DT 193	SRR1966416
H150520684	34	Human	27/01/2015	PT U311	SRR1967382
H150520687	34	Human	27/01/2015	DT 193	SRR1969237
H150540596	34	Human	28/01/2015	DT 120	SRR1969381
H150540602	34	Human	28/01/2015	DT 193	SRR1968886
H150540604	34	Human	28/01/2015	PT U311	SRR1966434
H150560629	34	Human	29/01/2015	DT 193	SRR1966531
H150560633	19	Human	29/01/2015	DT 132	SRR3048993
H150560638	19	Human	29/01/2015	RDNC	SRR1967936
H150560642	19	Human	29/01/2015	DT 41b	SRR1965280
H150580636	313	Raw product	30/01/2015	RDNC	SRR1968221
H150580642	19	Human	30/01/2015	DT 104	SRR1969373
H150580643	313	Human	30/01/2015	RDNC	SRR1967938
H150580648	568	Human	30/01/2015	DT 56	SRR1967253
H150600548	34	Human	02/02/2015	DT 104b	SRR1967348

H150620544	313	Human	03/02/2015	DT 193	SRR1969492
H150620545	313	Human	03/02/2015	DT 193	SRR1968269
H150620591	34	Human	03/02/2015	DT 193	SRR1970022
H150620593	34	Human	03/02/2015	DT 193	SRR1967194
H150640534	34	Human	04/02/2015	DT 193	SRR1968940
H150640559	19	Human	04/02/2015	DT 8	SRR1967003
H150640562	34	Human	04/02/2015	RDNC	SRR1968629
H150640564	34	Raw sausage	04/02/2015	DT 193	SRR1966871
H150640565	34	Human	04/02/2015	DT 193	SRR1969961
H150660556	19	Human	05/02/2015	DT 135	SRR1969536
H150660587	19	Human	05/02/2015	DT 10	SRR1968283
H150660588	34	Human	05/02/2015	DT 193	SRR1965517
H150660591	34	Human	05/02/2015	DT 193	SRR1967940
H150680611	34	Human	06/02/2015	DT 193	SRR1967474
H150700258	19	Human	09/02/2015	PT U288	SRR1966265
H150700259	19	Human	09/02/2015	PT U288	SRR1965201
H150700279	19	Human	09/02/2015	RDNC	SRR1967008
H150700280	19	Human	09/02/2015	DT 10	SRR1966404
H150700283	19	Human	09/02/2015	DT 135	SRR1965778
H150720593	19	IQA	10/02/2015	RDNC	SRR1969485
H150720596	34	Human	10/02/2015	DT 193	SRR1970116
H150720604	34	Human	10/02/2015	DT 193	SRR1969208
H150740512	19	Human	11/02/2015	DT 104	SRR1970045
H150740514	34	Human	11/02/2015	None	SRR1969180
H150740515	2079	Human	11/02/2015	Untypable	SRR1967523
H150740519	34	Human	11/02/2015	PT U323	SRR1968916
H150740521	19	Human	11/02/2015	PT U313	SRR1966143
H150760395	34	Human	12/02/2015	DT 193	SRR1968597
H150780479	19	Human	13/02/2015	DT 41	SRR1968223
H150780487	19	Human	13/02/2015	PT U302	SRR1969798
H150780489	19	Human	13/02/2015	RDNC	SRR1966410
H150780494	19	Human	13/02/2015	PT U302	SRR1965071
H150800485	19	Human	16/02/2015	RDNC	SRR1969440
H150800523	34	Human	16/02/2015	DT 193	SRR1966059
H150820630	34	Human	17/02/2015	DT 193	SRR1967558
H150820634	19	Human	17/02/2015	Untypable	SRR1969029
H150820638	19	Human	17/02/2015	DT 104	SRR1965809

H150820681	19	Human	17/02/2015	DT 44	SRR1968663
H150820682	19	Human	17/02/2015	DT 1	SRR1967493
H150840619	313	Human	18/02/2015	RDNC	SRR1966635
H150840621	568	Human	18/02/2015	DT 41	SRR1965287
H150840622	34	Human	18/02/2015	DT 193	SRR1966877
H150880555	19	Human	20/02/2015	RDNC	SRR1969165
H150880559	19	Raw sausage	20/02/2015	PT U288	SRR1965149
H150880560	19	Raw sausage	20/02/2015	PT U288	SRR1967027
H150880561	19	Raw sausage	20/02/2015	PT U288	SRR1966783
H150880563	34	Raw sausage	20/02/2015	DT 193	SRR1968525
H150900461	313	Human	23/02/2015	DT 56	SRR1966023
H150900562	313	Human	23/02/2015	RDNC	SRR1966593
H150900563	34	Human	23/02/2015	DT 193	SRR1966963
H150900564	34	Human	23/02/2015	DT 120	SRR1967554
H150900566	313	Human	23/02/2015	PT U274	SRR1968962
H150900568	34	Human	23/02/2015	DT 193	SRR1968960
H150900571	34	Human	23/02/2015	DT 193	SRR1969596
H150920652	34	Human	24/02/2015	DT 193	SRR1967873
H150920673	34	Human	24/02/2015	Untypable	SRR1969691
H150920675	34	Human	24/02/2015	DT 193	SRR1965634
H150920678	34	Human	24/02/2015	DT 193	SRR1969599
H150920681	34	Human	24/02/2015	DT 193	SRR1966157
H150940605	19	Human	25/02/2015	RDNC	SRR1968966
H150940620	313	Human	25/02/2015	PT U274	SRR1967438
H150940621	19	Human	25/02/2015	PT U312	SRR1968273
H150940623	34	Human	25/02/2015	DT 193	SRR1968091
H150940624	34	Human	25/02/2015	Untypable	SRR1968445
H150940627	34	Human	25/02/2015	DT 193	SRR1968450
H150940630	34	Human	25/02/2015	DT 193	SRR1970041
H150960615	19	Human	26/02/2015	DT 2	SRR1969860
H150960623	34	Human	26/02/2015	DT 193	SRR1968964
H150980442	34	Human	27/02/2015	DT 193	SRR3048561
H151000369	19	Human	02/03/2015	DT 12	SRR1966726
H151000372	19	Human	02/03/2015	DT 1	SRR1965174
H151020588	34	Human	03/03/2015	DT 193	SRR1970282
H151020623	34	Human	03/03/2015	PT U323	SRR1969498
H151020625	34	Human	03/03/2015	DT 193	SRR1967122

H151020626	34	Human	03/03/2015	Untypable	SRR1965858
H151020627	34	Human	03/03/2015	PT U323	SRR1965305
H151020629	19	UKNEQAS	03/03/2015	DT 12	SRR1968878
H151020630	34	Human	03/03/2015	DT 193	SRR1965821
H151020631	34	Human	03/03/2015	DT 193	SRR1965124
H151020632	34	Human	03/03/2015	DT 193	SRR1966239
H151020634	313	Human	03/03/2015	DT 2	SRR1968061
H151020638	34	Human	03/03/2015	PT U311	SRR1970114
H151040666	19	Human	04/03/2015	PT U312	SRR1970190
H151040668	34	Human	04/03/2015	DT 137	SRR1967234
H151040669	213	Human	04/03/2015	DT 131	SRR1969768
H151040672	19	Human	04/03/2015	DT 193	SRR1968876
H151060483	34	Human	05/03/2015	DT 193	SRR1968670
H151060484	34	Human	05/03/2015	Untypable	SRR1969001
H151060485	34	Human	05/03/2015	Untypable	SRR1965942
H151060493	34	Human	05/03/2015	DT 193	SRR1968847
H151060497	34	Human	05/03/2015	DT 120	SRR1969167
H151080426	34	Human	06/03/2015	Untypable	SRR1965751
H151080428	34	Human	06/03/2015	DT 120	SRR1968347
H151080429	19	Human	06/03/2015	DT 208	SRR1966917
H151080433	34	Human	06/03/2015	DT 193	SRR1965147
H151080437	34	Human	06/03/2015	Untypable	SRR1970091
H151100483	34	Human	09/03/2015	DT 193	SRR1969876
H151100485	34	Human	09/03/2015	DT 190	SRR1965497
H151100486	34	Human	09/03/2015	Untypable	SRR1970250
H151100487	34	Human	09/03/2015	Untypable	SRR1966820
H151100488	34	Human	09/03/2015	Untypable	SRR1965245
H151100488	34	Human	09/03/2015	Untypable	SRR3285491
H151100490	34	Human	09/03/2015	Untypable	SRR1969678
H151100491	34	Human	09/03/2015	DT 193	SRR1966079
H151120549	19	Human	10/03/2015	DT 8	SRR3049968
H151120553	34	Human	10/03/2015	DT 193	SRR3049457
H151120554	19	Human	10/03/2015	RDNC	SRR3049296
H151120555	34	Human	10/03/2015	Untypable	SRR3049555
H151120556	34	Human	10/03/2015	Untypable	SRR3049385
H151120557	34	Human	10/03/2015	Untypable	SRR3048985
H151120558	34	Human	10/03/2015	Untypable	SRR3049575

H151120559	34	Human	10/03/2015	Untypable	SRR3049482
H151120560	34	Human	10/03/2015	Untypable	SRR3049188
H151120561	34	Human	10/03/2015	Untypable	SRR3049535
H151120562	34	Human	10/03/2015	DT 120	SRR3049047
H151120568	34	Human	10/03/2015	DT 193	SRR3049536
H151140648	34	Human	11/03/2015	DT 193	SRR1965092
H151140654	19	Human	11/03/2015	DT 193	SRR1966813
H151140674	34	Human	11/03/2015	DT 120	SRR1969468
H151160603	34	Human	12/03/2015	DT 193	SRR1967718
H151180241	19	Human	13/03/2015	None	SRR1969327
H151180245	34	Human	13/03/2015	DT 193	SRR1969531
H151180246	34	Human	13/03/2015	Untypable	SRR1968704
H151180253	34	Human	13/03/2015	Untypable	SRR3049949
H151200223	19	Human	16/03/2015	PT U288	SRR1968605
H151200224	313	Human	16/03/2015	RDNC	SRR1969857
H151220453	19	Human	17/03/2015	DT 193	SRR1967247
H151220457	34	Human	17/03/2015	DT 193	SRR1967643
H151220461	34	Human	17/03/2015	Untypable	SRR1969719
H151220462	34	Human	17/03/2015	Untypable	SRR1966786
H151220463	34	Human	17/03/2015	Untypable	SRR1966099
H151220464	34	Human	17/03/2015	Untypable	SRR1968099
H151220465	34	Human	17/03/2015	Untypable	SRR1969618
H151220466	34	Human	17/03/2015	Untypable	SRR1968549
H151220467	34	Human	17/03/2015	DT 193	SRR1966883
H151220468	34	Human	17/03/2015	DT 193	SRR1966044
H151220469	34	Human	17/03/2015	Untypable	SRR1966834
H151220470	34	Human	17/03/2015	DT 193	SRR1970314
H151220472	568	House sparrow	17/03/2015	DT 56	SRR1965823
H151220474	568	Reed bunting	17/03/2015	DT 56	SRR1966622
H151240457	19	Human	18/03/2015	DT 193	SRR1968405
H151280436	34	Human	20/03/2015	DT 120	SRR3049127
H151280438	19	Human	20/03/2015	DT 104	SRR3049680
H151280509	34	Human	20/03/2015	DT 193	SRR3049975
H151280512	34	Human	20/03/2015	DT 193	SRR3049101
H151280513	34	Human	20/03/2015	PT U323	SRR3049297
H151300474	34	Human	23/03/2015	DT 193	SRR3049890
H151300475	34	Human	23/03/2015	Untypable	SRR1965836

H151300476	34	Human	23/03/2015	Untypable	SRR1968057
H151300477	34	Human	23/03/2015	Untypable	SRR1966155
H151300479	34	Human	23/03/2015	DT 193	SRR1967561
H151300480	34	Human	23/03/2015	Untypable	SRR1967212
H151300482	34	Human	23/03/2015	RDNC	SRR3049889
H151320496	34	Human	24/03/2015	DT 193	SRR1966185
H151320498	19	Human	24/03/2015	PT U288	SRR1967485
H151320500	19	Human	24/03/2015	DT 12a	SRR3049209
H151320504	34	Human	24/03/2015	DT 193	SRR1966131
H151340523	34	Human	25/03/2015	DT 193	SRR1965442
H151340531	34	Human	25/03/2015	PT U323	SRR3048757
H151340533	34	Human	25/03/2015	DT 195	SRR3049027
H151340538	34	Human	25/03/2015	DT 120	SRR3049041
H151360443	34	Human	26/03/2015	DT 195	SRR3048819
H151360458	34	Human	26/03/2015	DT 193	SRR3049102
H151380459	34	Human	27/03/2015	DT 193	SRR3048923
H151380461	34	Human	27/03/2015	DT 195	SRR3049823
H151380467	34	Human	27/03/2015	DT 193	SRR3049682
H151400464	34	Human	30/03/2015	DT 193	SRR3049454
H151420580	19	Human	31/03/2015	DT 104b	SRR3049451
H151420583	34	Human	31/03/2015	Untypable	SRR3049466
H151420588	34	Human	31/03/2015	DT 193	SRR3049971
H151440507	34	Human	01/04/2015	DT 193	SRR3286591
H151440566	19	Human	01/04/2015	DT 120	SRR3286566
H151540860	19	Human	08/04/2015	DT 104	SRR3286646
H151540864	34	Human	08/04/2015	DT 120	SRR3286617
H151540868	34	Human	08/04/2015	Untypable	SRR3286592
H151540879	34	Human	08/04/2015	DT 193	SRR3286599
H151540883	34	Human	08/04/2015	DT 193	SRR3286627
H151540888	34	Human	08/04/2015	DT 193	SRR3286678
H151540890	34	Human	08/04/2015	DT 193	SRR3286676
H151540892	34	Human	08/04/2015	DT 120	SRR3286587
H151560674	34	Unknown	09/04/2015	DT 193	SRR3286634
H151560676	34	Unknown	09/04/2015	DT 193	SRR3285356
H151560680	19	Human	09/04/2015	PT U331	SRR3286667
H151560685	34	Human	09/04/2015	Untypable	SRR3286631
H151580353	19	Human	10/04/2015	DT 10	SRR3286633

H151620337	19	Human	14/04/2015	PT U309	SRR3285351
H151620362	19	Human	14/04/2015	DT 104	SRR3286672
H151640374	34	Human	15/04/2015	DT 193	SRR3286664
H151640376	19	Human	15/04/2015	PT U302	SRR3286616
H151640382	19	Human	15/04/2015	DT 191a	SRR3286660
H151640383	34	Human	15/04/2015	Untypable	SRR3286641
H151640411	34	Human	15/04/2015	DT 120	SRR3285446
H151660628	34	Human	16/04/2015	Untypable	SRR3286658
H151660629	34	Human	16/04/2015	Untypable	SRR3286673
H151680561	34	Human	17/04/2015	DT 193	SRR3286696
H151700443	34	Human	20/04/2015	DT 120	SRR3286597
H151700448	34	Human	20/04/2015	DT 193	SRR3286604
H151700449	19	Human	20/04/2015	DT 93	SRR3285378
H151700453	34	Human	20/04/2015	DT 193	SRR3286570
H151720436	34	Human	21/04/2015	DT 193	SRR3286671
H151720439	313	Human	21/04/2015	DT 56	SRR3286620
H151720461	313	Human	21/04/2015	DT 56	SRR3285401
H151720643	19	Human	21/04/2015	DT 104	SRR3284714
H151720657	34	Human	21/04/2015	DT 193	SRR3286640
H151720671	19	Human	21/04/2015	DT 104	SRR3286623
H151740324	34	Human	22/04/2015	Untypable	SRR3286580
H151740325	34	Human	22/04/2015	Untypable	SRR3286588
H151740327	34	Human	22/04/2015	Untypable	SRR3284828
H151760296	34	Human	23/04/2015	Untypable	SRR3286608
H151760308	34	Human	23/04/2015	Untypable	SRR3286612
H151800443	34	Human	27/04/2015	Untypable	SRR3323033
H151820415	34	Human	28/04/2015	Untypable	SRR3284731
H151820416	34	Human	28/04/2015	Untypable	SRR3285483
H151820500	34	Human	28/04/2015	DT 193	SRR3286637
H151820505	19	Human	28/04/2015	DT 104	SRR3286578
H151820511	19	Human	28/04/2015	DT 104	SRR3322097
H151820523	34	Human	28/04/2015	DT 193	SRR3322411
H151820524	34	Human	28/04/2015	DT 120	SRR3285419
H151840449	19	Human	29/04/2015	RDNC	SRR3286607
H151840455	34	Human	29/04/2015	DT 120	SRR3286600
H151840468	34	Human	29/04/2015	PT U311	SRR3284753
H151840476	19	Human	29/04/2015	DT 104b	SRR3286644

H151860632	19	Human	30/04/2015	DT 1	SRR3286662
H151920509	34	Human	05/05/2015	PT U311	SRR3286704
H151920511	34	Human	05/05/2015	PT U311	SRR3286648
H151940558	19	Human	06/05/2015	PT U288	SRR3285442
H151960469	19	Human	07/05/2015	RDNC	SRR3286848
H151980548	19	Human	08/05/2015	DT 1	SRR3286688
H152020819	34	Human	12/05/2015	Untypable	SRR3286723
H152020821	34	Human	12/05/2015	DT 193	SRR3286868
H152020822	19	Human	12/05/2015	DT 42	SRR3286819
H152020836	19	Human	12/05/2015	DT 132	SRR3321718
H152020838	99	Human	12/05/2015	DT 193	SRR3286705
H152020863	19	Human	12/05/2015	DT 87	SRR3322723
H152040583	34	Human	13/05/2015	DT 193	SRR3322329
H152060596	34	Human	14/05/2015	DT 193	SRR3286727
H152060607	34	Human	14/05/2015	DT 120	SRR3286709
H152060609	19	Human	14/05/2015	PT U298	SRR3286821
H152080598	34	Raw meat	15/05/2015	DT 193	SRR3286706
H152080602	34	Human	15/05/2015	DT 193	SRR3286699
H152080603	34	Raw meat	15/05/2015	DT 193	SRR3286683
H152080608	19	Human	15/05/2015	DT 135	SRR3322320
H152100636	568	Human	18/05/2015	DT 56	SRR3286707
H152120540	34	Human	19/05/2015	DT 193	SRR3286689
H152120542	34	Human	19/05/2015	DT 193	SRR3284765
H152120543	19	Human	19/05/2015	PT U288	SRR3285053
H152120554	34	Human	19/05/2015	DT 193	SRR3286686
H152180624	19	Human	22/05/2015	PT U298	SRR3284723
H152180625	34	Human	22/05/2015	PT U302	SRR3286703
H152180628	34	Human	22/05/2015	DT 193	SRR3286714
H152240454	19	Human	27/05/2015	RDNC	SRR3286695
H152240468	34	Human	27/05/2015	DT 193	SRR3322420
H152240488	34	Human	27/05/2015	DT 193	SRR3322960
H152260543	34	Human	28/05/2015	DT 193	SRR3286729
H152280417	19	Human	29/05/2015	DT 41	SRR3284823
H152280418	19	Human	29/05/2015	PT U302	SRR3286779
H152280421	19	Human	29/05/2015	RDNC	SRR3286813
H152300644	19	Human	01/06/2015	DT 1	SRR3321902
H152300647	34	Human	01/06/2015	PT U329	SRR3286776

H152300650	34	Human	01/06/2015	PT U329	SRR3286815
H152320501	34	Human	02/06/2015	DT 120	SRR3286875
H152320524	19	Human	02/06/2015	PT U302	SRR3286858
H152320527	34	Human	02/06/2015	DT 193	SRR3286899
H152320528	34	Human	02/06/2015	PT U329	SRR3286895
H152320546	34	Human	02/06/2015	DT 120	SRR3322591
H152320547	34	Human	02/06/2015	PT U329	SRR3286913
H152340669	19	Human	03/06/2015	DT 104	SRR3286893
H152340700	19	Human	03/06/2015	PT U302	SRR3284851
H152340701	19	Human	03/06/2015	DT 10	SRR3286872
H152340702	19	Human	03/06/2015	DT 10	SRR3286883
H152340712	34	Human	03/06/2015	DT 120	SRR3286905
H152340713	34	Human	03/06/2015	DT 193	SRR3284715
H152340799	213	EMRU	03/06/2015	DT 1	SRR3286840
H152360402	19	Human	04/06/2015	DT 193	SRR3286862
H152360403	34	Human	04/06/2015	PT U329	SRR3286892
H152360408	19	Human	04/06/2015	DT 116	SRR3286912
H152380439	19	Human	05/06/2015	DT 193	SRR3285469
H152380466	19	Cat	05/06/2015	PT U330	SRR3322031
H152380468	19	Human	05/06/2015	PT U302	SRR3286865
H152380479	19	Human	05/06/2015	PT U302	SRR3285354
H152400594	34	Human	08/06/2015	DT 193	SRR3286888
H152400631	19	Human	08/06/2015	PT U288	SRR3323019
H152400632	34	Human	08/06/2015	DT 193	SRR3322096
H152400637	34	Human	08/06/2015	PT U329	SRR3286837
H152400637	34	Human	08/06/2015	PT U329	SRR3286903
H152420810	34	Human	09/06/2015	Untypable	SRR3286896
H152420817	19	Human	09/06/2015	DT 1	SRR3286915
H152420820	34	Human	09/06/2015	PT U329	SRR3285425
H152440505	34	Human	10/06/2015	DT 193	SRR3285445
H152440529	34	Human	10/06/2015	DT 120	SRR3321731
H152440533	19	Human	10/06/2015	DT 104	SRR3286927
H152460484	34	Human	11/06/2015	DT 193	SRR3285461
H152460491	34	Human	11/06/2015	PT U329	SRR3286851
H152460492	34	Human	11/06/2015	PT U311	SRR3284783
H152480461	19	Human	12/06/2015	PT U307	SRR3286894
H152480473	19	Human	12/06/2015	RDNC	SRR3285155

H152520807	34	Human	16/06/2015	DT 7a	SRR3286864
H152520822	34	Human	16/06/2015	Untypable	SRR3286884
H152520823	19	Human	16/06/2015	DT 193	SRR3286897
H152520830	19	Human	16/06/2015	DT 116	SRR3285349
H152540595	19	Human	17/06/2015	DT 193	SRR3286854
H152540598	34	Human	17/06/2015	PT U329	SRR3285080
H152540610	19	Human	17/06/2015	RDNC	SRR3285451
H152540619	19	Human	17/06/2015	DT 66a	SRR3322370
H152580650	34	Human	19/06/2015	RDNC	SRR3322094
H152580658	19	Human	19/06/2015	DT 46	SRR3286965
H152600555	34	Human	22/06/2015	DT 193	SRR3285347
H152600561	34	Human	22/06/2015	Untypable	SRR3286922
H152620495	34	Human	23/06/2015	DT 193	SRR3286925
H152640594	313	Human	24/06/2015	RDNC	SRR3323011
H152660386	19	Human	25/06/2015	PT U288	SRR3322701
H152660397	19	Human	25/06/2015	DT 135	SRR3322667
H152700573	34	Human	29/06/2015	PT U302	SRR3321856
H152700576	34	Human	29/06/2015	PT U329	SRR3285456
H152700580	568	Human	29/06/2015	DT 56	SRR3286918
H152700586	19	Human	29/06/2015	DT 120	SRR3286929
H152700587	19	Human	29/06/2015	DT 12	SRR3322091
H152720601	19	Human	30/06/2015	DT 46	SRR3285288
H152720602	19	Human	30/06/2015	DT 46	SRR3286940
H152720605	19	Human	30/06/2015	DT 8	SRR3286963
H152720608	19	Human	30/06/2015	DT 104	SRR3286973
H152720613	19	Human	30/06/2015	DT 85	SRR3284814
H152720615	19	Human	30/06/2015	DT 104b	SRR3286955
H152720617	34	Human	30/06/2015	PT U311	SRR3286945
H152720637	19	Human	30/06/2015	DT 116	SRR3322099
H152720640	19	Human	30/06/2015	PT U311	SRR3285397
H152720641	19	Raw material spice	30/06/2015	DT 193	SRR3285367
H152720642	34	Human	30/06/2015	DT 120	SRR3286950
H152720643	19	Human	30/06/2015	RDNC	SRR3286956
H152720648	19	Human	30/06/2015	DT 120	SRR3286949
H152720655	P3572	Human	30/06/2015	DT 120	SRR3321997
H152720661	19	Human	30/06/2015	PT U312	SRR3285452

H152720664	34	Human	30/06/2015	DT 193	SRR3286901
H152740235	19	Human	01/07/2015	DT 193	SRR3286920
H152740236	19	Human	01/07/2015	PT U298	SRR3286932
H152740238	34	Human	01/07/2015	DT 193	SRR3286958
H152740240	34	Human	01/07/2015	Untypable	SRR3284717
H152740273	19	Human	01/07/2015	DT 8	SRR3285243
H152760467	34	Human	02/07/2015	DT 193	SRR3285380
H152760469	34	Human	02/07/2015	DT 120	SRR3286926
H152760473	313	Human	02/07/2015	RDNC	SRR3284779
H152820418	19	Human	07/07/2015	DT 9	SRR3286967
H152840417	34	Human	08/07/2015	DT 193	SRR3322972
H152860554	19	Human	09/07/2015	DT 101	SRR3322710
H152860555	34	Human	09/07/2015	DT 120	SRR3286969
H152880399	34	Human	10/07/2015	DT 120	SRR3321855
H152900244	19	Human	13/07/2015	DT 104b	SRR3286974
H152900267	19	Human	13/07/2015	PT U278	SRR3322628
H152920486	19	Human	14/07/2015	DT 40	SRR3322970
H152920498	19	Human	14/07/2015	PT U278	SRR3322721
H152940468	313	Human	15/07/2015	DT 2	SRR3285477
H152940469	19	Human	15/07/2015	RDNC	SRR3284752
H152980464	19	Human	17/07/2015	RDNC	SRR3285368
H153000432	19	Human	20/07/2015	DT 104	SRR3285437
H153020667	19	Human	21/07/2015	RDNC	SRR3285434
H153040541	34	Human	22/07/2015	PT U329	SRR3285471
H153120511	19	Human	28/07/2015	DT 193	SRR3285399
H153120515	19	Human	28/07/2015	PT U331	SRR3285430
H153140404	34	Human	29/07/2015	DT 193	SRR3285421
H153160274	34	Human	30/07/2015	DT 193	SRR3284845
H153160280	34	Human	30/07/2015	PT U323	SRR3284818
H153220457	34	Human	04/08/2015	DT 193	SRR3284739
H153260506	34	Human	06/08/2015	PT U323	SRR3285468
H153260510	19	Human	06/08/2015	DT 12	SRR3285455
H153300540	34	Human	10/08/2015	RDNC	SRR3284825
H153400468	19	Human	17/08/2015	DT 104	SRR3284813
H153400470	19	Human	17/08/2015	DT 8	SRR3284768
H153440473	19	Human	19/08/2015	DT 12	SRR3284830
H153440473	19	Human	19/08/2015	DT 12	SRR3285440

H153440508	34	Human	19/08/2015	DT 120	SRR3285383
H153460592	19	Human	20/08/2015	DT 104	SRR3284841
H153460595	34	Human	20/08/2015	DT 193	SRR3285473
H153460616	34	Human	20/08/2015	DT 104b	SRR3284834
H153480514	34	Human	21/08/2015	DT 193	SRR3285364
H153520446	34	Human	25/08/2015	DT 120	SRR3285478
H153520451	19	Human	25/08/2015	DT 193	SRR3284829
H153520457	34	Human	25/08/2015	DT 193	SRR3285366
H153520466	19	Human	25/08/2015	PT U312	SRR3284693
H153520481	19	Human	25/08/2015	Untypable	SRR3284686
H153540331	34	Human	26/08/2015	Untypable	SRR3285358
H153560137	34	Human	27/08/2015	Untypable	SRR3284809
H153560143	19	Human	27/08/2015	PT U302	SRR3285420
H153640677	34	Human	02/09/2015	DT 195	SRR3285374
H153640680	19	Human	02/09/2015	DT 193	SRR3285402
H153640688	19	Human	02/09/2015	PT U302	SRR3285413
H153680363	34	Human	04/09/2015	DT 193	SRR3285406
H153700324	34	Human	07/09/2015	DT 193	SRR3284758
H153700446	34	Human	07/09/2015	PT U329	SRR3285422
H153740470	34	Human	09/09/2015	DT 120	SRR3284744
H153820390	19	Human	15/09/2015	RDNC	SRR3285359
H153840505	34	Human	16/09/2015	DT 193	SRR3285242
H153840508	34	Human	16/09/2015	PT U311	SRR3285405
H153840513	34	Human	16/09/2015	DT 193	SRR3285458
H153840518	34	Human	16/09/2015	PT U311	SRR3284760
H153840520	34	Human	16/09/2015	Untypable	SRR3284810
H153940639	19	Human	23/09/2015	PT U302	SRR3285435
H153960420	19	Human	24/09/2015	DT 104	SRR3284677
H153960432	19	Human	24/09/2015	DT 101	SRR3284839
H153980204	19	Human	25/09/2015	DT 195	SRR3284807
H154040409	19	Human	30/09/2015	DT 1	SRR3284716
H154060438	19	Human	01/10/2015	RDNC	SRR3284847
H154080147	19	Human	02/10/2015	DT 15a	SRR3285499
H154100393	34	Human	05/10/2015	Untypable	SRR3284725
H154120570	34	Human	06/10/2015	DT 120	SRR3285282
H154140372	34	Human	07/10/2015	DT 120	SRR3284711
H154160457	34	Human	08/10/2015	DT 193	SRR3285360

H154220780	19	Human	13/10/2015	DT 35	SRR3285490
H154220788	34	Human	13/10/2015	Untypable	SRR3284675
H154240483	34	Human	14/10/2015	DT 120	SRR3284741
H154280562	34	Human	16/10/2015	PT U311	SRR3285348
H154300382	34	Human	19/10/2015	PT U329	SRR3284734
H154320744	19	Human	20/10/2015	DT 41	SRR3284682
H154320745	19	Human	20/10/2015	DT 1	SRR3285379
H154360432	19	Human	22/10/2015	DT 1	SRR3285489
H154360435	34	Human	22/10/2015	DT 193	SRR3285427
H154460357	34	Human	29/10/2015	DT 193	SRR3285400
H154500452	34	Human	02/11/2015	None	SRR3284730
H154520625	19	Human	03/11/2015	None	SRR3284681
H154520647	34	Human	03/11/2015	None	SRR3285252
H154520679	34	Human	03/11/2015	None	SRR3285487
H154620518	34	Human	10/11/2015	None	SRR3284812
H154660433	19	Human	12/11/2015	None	SRR3285346
H154700345	19	Human	16/11/2015	None	SRR3284743
H154720547	34	Human	17/11/2015	None	SRR3285398
H154720549	34	Human	17/11/2015	None	SRR3284832
H154720556	34	Human	17/11/2015	None	SRR3285279
H154760207	34	Human	19/11/2015	None	SRR3285432
H154780417	19	Human	20/11/2015	None	SRR3284747
H154800359	34	Human	23/11/2015	None	SRR3284831
H154820551	19	Human	24/11/2015	None	SRR3285384
H154920516	34	Human	01/12/2015	None	SRR3284691
H154940333	19	Human	02/12/2015	None	SRR3284727
H154960546	19	Human	03/12/2015	None	SRR3285428
H154960760	34	Sewer swab	03/12/2015	None	SRR3284748
H155020271	34	Old spray bottle	08/12/2015	None	SRR3285454
H155060512	19	Human	10/12/2015	None	SRR3284722
H155060514	19	Human	10/12/2015	None	SRR3284811
H155120548	34	Human	15/12/2015	None	SRR3284719
H155140536	34	Dirty used oven cloth	16/12/2015	None	SRR3284819
H155140539	34	Dirty used oven cloth	16/12/2015	None	SRR3284780
H155140542	34	Human	16/12/2015	None	SRR3285055

H155160053	19	Human	17/12/2015	None	SRR3285438
H155220490	34	Human	22/12/2015	None	SRR3284816
H155360251	34	Human	31/12/2015	None	SRR3285460
H160120678	34	Human	05/01/2016	None	SRR3285376
H160140786	19	Human	06/01/2016	None	SRR3284821
H160160247	34	Human	07/01/2016	None	SRR3285466
H160220779	34	Swab of WHB	12/01/2016	None	SRR3285414
H160220781	34	Human	12/01/2016	None	SRR3284778
H160240656	34	Human	13/01/2016	None	SRR3285396
H160300259	34	Human	18/01/2016	None	SRR3285281

Appendix 5 List of isolates tested for lysogenic conversion with mTmV (Chapter 5). The isolate name and serovar of the isolates are reported together with their phage type, sequence type and accession number/reference, when available.

Isolate	**Serovar**	**Phage type**	**Sequence type**	**Accession number/ Reference**
166A	Anatum	-	-	(Burns, 2015)
793C	Derby	-	-	(Burns, 2015)
783A	Derby	-	-	(Burns, 2015)
3387A	Derby	-	-	(Burns, 2015)
3264A	Derby	-	-	(Burns, 2015)
2617A	Derby	-	-	(Burns, 2015)
2591A	Derby	-	-	(Burns, 2015)
2582A	Derby	-	-	(Burns, 2015)
2577A	Derby	-	-	(Burns, 2015)
1918B	Derby	-	-	(Burns, 2015)
1097A	Derby	-	-	(Burns, 2015)
3108A	Dublin	-	-	(Burns, 2015)
3149B	Dublin	-	-	(Burns, 2015)
2593A	Infantis	-	-	(Burns, 2015)
799D	Infantis	-	-	(Burns, 2015)
3337A	London	-	-	(Burns, 2015)
3373A	London	-	-	(Burns, 2015)
3657A	London	-	-	(Burns, 2015)

3661C	London	-	-	(Burns, 2015)
1985A	Stanley	-	-	(Burns, 2015)
3426A	Tennessee	-	-	(Burns, 2015)
3447A	Tennessee	-	-	(Burns, 2015)
3598A	Tennessee	-	-	(Burns, 2015)
01960-05	Typhimurium	U288	19	ERR029231
03-715	Typhimurium	DT2	-	ERR028071
11020-1996	Typhimurium	DT193	19	ERR028308
2882-06	Typhimurium	DT99		ERR024407
4179-2001	Typhimurium	DT121	19	ERR028301
D23580	Typhimurium	-	ST313	FN424405.1
L01157-10	Typhimurium	DT8	-	-
NCTC13348	Typhimurium	DT104	19	HF937208.1
S07676-03	Typhimurium DT56	DT56	568	ERR029222
SL1344	Typhimurium DT44	DT44	19	FQ312003.1
SO7292-07	Typhimurium U308	U308	19	ERR029234
CVM19633	Schwarzengrund	-	-	CP001127.1
SGSC4901	Enteritidis	PT4		
287/91	Gallinarum			AM933173
LK5	Enteritidis	PT8		ERS673772
SL476	Heidelberg	-	-	NC_011083.1

Appendix 6 Polymorphisms in mTmV prophage sequences detected in *S*. Derby, *S*. California and *S*. Brandenburg isolates in comparison with the reference sequence of S04698-09 (accession number NZ_LN999997.1) described in Chapter 5. The strain name, the nucleotide position and the type of the polymorphism are reported as well as the reference and alternative sequence and the locus affected.

Isolate	**Position**	**Type of mutation**	**Ref.**	**Alt.**	**Product (ORF n.)**
DerbySA64	4273	snp	A	C	-
Brandenburg SA20113174	9300	snp	G	A	Dam protein (00338)

DerbySA64	12027	snp	G	T	Hypothetical protein (00342)
DerbySA64	12991	snp	T	C	-
DerbySA64	18203	snp	T	G	Phage terminase (00351)
DerbySA64	18799	snp	G	C	Hypothetical protein (00352)
DerbySA64	19445	snp	C	A	Phage portal protein (00353)
Derby 2014LSAL02547	19795	snp	G	A	Phage portal protein (00353)
Derby 2014LSAL02547	19834	snp	T	C	Phage portal protein (00353)
Brandenburg SA20113174	21064	snp	A	C	Phage capsid protein (00355)
Brandenburg SA20113174	21097	snp	G	A	Phage capsid protein (00355)
Brandenburg SA20113174	21113	complex	CTA	TTG	Phage capsid protein (00355)
Brandenburg SA20113174	21133	snp	C	T	Phage capsid protein (00355)
Brandenburg SA20113174	21172	snp	T	C	Phage capsid protein (00355)
Brandenburg SA20113174	21193	snp	C	T	Phage capsid protein (00355)
Brandenburg SA20113174	21207	complex	AA	CG	Phage capsid protein (00355)
Brandenburg SA20113174	21226	complex	TAA	CAC	Phage capsid protein (00355)
Brandenburg SA20113174	21256	snp	T	C	Phage capsid protein (00355)
Brandenburg SA20113174	21319	snp	T	C	Phage capsid protein (00355)
DerbySA64	21335	snp	G	T	Phage capsid protein (00355)
Brandenburg SA20113174	21427	snp	T	C	Phage capsid protein (00355)

Brandenburg SA20113174	24466	ins	G	GC	Phage tail sheath protein (00361)
Derby CFSA231	24466	ins	G	GC	Phage tail sheath protein (00361)
California CD-SL01	24466	ins	G	GC	Phage tail sheath protein (00361)
DerbySA64	24466	ins	G	GC	Phage tail sheath protein (00361)
Derby SA2003	24466	ins	G	GC	Phage tail sheath protein (00361)
Derby 2014LSAL02547	24466	ins	G	GC	Phage tail sheath protein (00361)
Brandenburg SA20113174	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
Derby CFSA231	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
California CD-SL01	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
DerbySA64	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
Derby SA2003	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
Derby 2014LSAL02547	24561	complex	ACCG	CCGA	Hypothetical protein (00362)
DerbySA64	24738	ins	A	AGG	-
Brandenburg SA20113174	24880	del	TG	T	Hypothetical protein (00362)
Derby CFSA231	24880	del	TG	T	Hypothetical protein (00362)
California CD-SL01	24880	del	TG	T	Hypothetical protein (00362)
DerbySA64	24880	del	TG	T	Hypothetical protein (00362)
Derby SA2003	24880	del	TG	T	Hypothetical protein (00362)
Derby 2014LSAL02547	24880	del	TG	T	Hypothetical protein (00362)

Brandenburg SA20113174	24924	del	AACG	A	Phage tail sheath protein (00363)
Derby CFSA231	24924	del	AACG	A	Phage tail sheath protein (00363)
California CD-SL01	24924	del	AACG	A	Phage tail sheath protein (00363)
DerbySA64	24924	del	AACG	A	Phage tail sheath protein (00363)
Derby SA2003	24924	del	AACG	A	Phage tail sheath protein (00363)
Derby 2014LSAL02547	24924	del	AACG	A	Phage tail sheath protein (00363)
DerbySA64	25355	ins	C	CG	Phage tail sheath protein (00363)
DerbySA64	25841	snp	C	T	Phage tail tube protein (00364)
DerbySA64	26280	ins	C	CT	Hypothetical protein (00365)
DerbySA64	26907	ins	G	GAAA	Minor tail protein (00366)
Derby 2014LSAL02547	27131	snp	T	C	Minor tail protein (00366)
Brandenburg SA20113174	27362	snp	G	A	Minor tail protein (00366)
Derby SA2003	27362	snp	G	A	Minor tail protein (00366)
Derby 2014LSAL02547	27362	snp	G	A	Minor tail protein (00366)
Brandenburg SA20113174	27400	snp	C	A	Minor tail protein (00366)
Derby SA2003	27400	snp	C	A	Minor tail protein (00366)
Derby 2014LSAL02547	27400	snp	C	A	Minor tail protein (00366)
DerbySA64	28823	snp	A	G	Hypothetical protein (00367)

DerbySA64	30745	ins	C	CG	Phage late control gene D (00369)
DerbySA64	31378	ins	G	GGAAAAA	Phage late control gene D (00369)
DerbySA64	32449	snp	C	G	Hypothetical protein (00373)
Brandenburg SA20113174	33160	snp	G	C	Baseplate J-like protein (00372)
Derby CFSA231	33160	snp	G	C	Baseplate J-like protein (00372)
California CD-SL01	33160	snp	G	C	Baseplate J-like protein (00372)
DerbySA64	33160	snp	G	C	Baseplate J-like protein (00372)
Derby SA2003	33160	snp	G	C	Baseplate J-like protein (00372)
Derby 2014LSAL02547	33160	snp	G	C	Baseplate J-like protein (00372)
DerbySA64	33886	ins	A	AC	Hypothetical protein (00373)
DerbySA64	34481	ins	C	CGT	Hypothetical protein (00373)
DerbySA64	35076	snp	C	G	Caudovirales tail fiber assembly protein (00375)
DerbySA64	35228	snp	G	C	Caudovirales tail fiber assembly protein (00375)

Appendix 7 Diverse *S. enterica* serovars isolated from Irish pig farms (Burns, 2015) sequenced to check for the presence of *sopE* and mTmV (Chapter 5). The isolate name, serovar, farm and visit (either first or second sampling date), *sopE* and mTmV presence are reported. *sopE* and mTmV were detected with SRST2 (Inouye et al., 2014) using the sequences of S04698-09 as reference (accession number NZ_LN999997.1).

Strain name	**Serovar**	**Farm**	**Visit**	***sopE***	**mTmV**
166A	Anatum	A	1	absent	absent
2158A	Derby	C	2	present	present

547A	Derby	C	1	present	present
551B	Derby	C	1	present	present
693D	Derby	D	1	present	present
1097A	Derby	G	1	absent	absent
1918B	Derby	B	2	absent	absent
2577A	Derby	E	2	absent	absent
2582A	Derby	E	2	absent	absent
2591A	Derby	E	2	absent	absent
2617A	Derby	E	2	absent	absent
3264A	Derby	I	2	absent	absent
3387A	Derby	I	2	absent	absent
3637A	Derby	J	2	absent	absent
783A	Derby	E	1	absent	absent
793C	Derby	E	1	absent	absent
3108A	Dublin	H	2	present	absent
3149B	Dublin	H	2	present	absent
3254A	Dublin	I	2	absent	absent
2593A	Infantis	E	2	absent	absent
799D	Infantis	E	1	absent	absent
3337A	London	I	2	absent	absent
3373A	London	I	2	absent	absent
3558A	London	J	2	absent	absent
3647A	London	J	2	absent	absent
3657A	London	J	2	absent	absent
3661C	London	J	2	absent	absent
1758A	Orion	A	2	absent	absent
1985A	Stanley	B	2	absent	absent
3426A	Tennessee	J	2	absent	absent
3447A	Tennessee	J	2	absent	absent
3598A	Tennessee	J	2	absent	absent
1086C	Typhimurium Copenhagen	G		absent	absent
1339A	Typhimurium Copenhagen	I		absent	absent

1367A	Typhimurium Copenhagen	I		absent	absent

Appendix 8 Sequences of *S*. Derby isolates from clinical infections in the UK assessed for the presence of *sopE* and mTmV (Chapter 5). The isolate name, source and year of isolation are reported together with mTmV and *sopE* presence/absence which was determined using SRST2 (Inouye et al., 2014). The detection of polymorphic sequences in comparison with the reference sequences of S04698-09 (accession number NZ_LN999997.1) is marked with an asterisk (*).

Isolate	**Source**	**Year of Isolation**	**Accession number**	**mTmV**	***sopE***
H133220765	Food	2013	SAMN03168795	absent	absent
H121420373	Human	2012	SAMN03168898	absent	absent
H122800357	Human	2012	SAMN03168996	absent	absent
H124760756	Food	2012	SAMN03169293	absent	absent
H123540655	Human	2012	SAMN03169449	absent	absent
44675	Human	2014	SAMN03465613	absent	absent
53980	Human	2014	SAMN03465799	present*	present*
50140	Human	2014	SAMN03465805	present*	present
54035	Human	2014	SAMN03465833	present*	present
57670	Human	2014	SAMN03465874	present*	present
46624	Human	2014	SAMN03466185	absent	absent
73629	Food	2014	SAMN03468472	absent	absent
78606	Food	2014	SAMN03468482	absent	absent
78607	Food	2014	SAMN03468561	absent	absent
78608	Food	2014	SAMN03468606	absent	absent
69183	ND	2014	SAMN03468618	absent	absent
71536	Human	2014	SAMN03469032	absent	absent
78610	Food	2014	SAMN03469625	absent	absent
78609	Food	2014	SAMN03469727	absent	absent
73649	Human	2014	SAMN03469823	absent	absent
73630	Food	2014	SAMN03469896	absent	absent
73628	Food	2014	SAMN03469931	absent	absent
84423	Food	2015	SAMN03475410	absent	absent
97854	Human	2015	SAMN03475480	absent	absent

73204	Food	2014	SAMN03475653	absent	absent
7398	Food	2014	SAMN03475761	absent	absent
9454	Human	2014	SAMN03475766	absent	absent
25182	Human	2014	SAMN03475802	absent	absent
69281	Human	2014	SAMN03475923	absent	absent
87401	Food	2015	SAMN03476602	absent	absent
91968	Human	2015	SAMN03476659	absent	absent
53000	Food	2014	SAMN03476852	absent	absent
98271	Human	2015	SAMN03477011	absent	absent
73203	Food	2014	SAMN03477122	absent	absent
67571	Human	2014	SAMN03477296	absent	absent
73201	Food	2014	SAMN03477328	absent	absent
95032	Human	2015	SAMN03478455	absent	absent
75994	Human	2014	SAMN03479094	present*	present
84422	Food	2015	SAMN03479189	absent	absent
9250	ND	2014	SAMN03479309	absent	absent
69894	Human	2014	SAMN03479472	present*	present
80598	Human	2015	SAMN03479636	absent	absent
18636	Human	-	SAMN04362904	absent	absent
15233	Human	-	SAMN04362912	absent	absent
56931	Human	-	SAMN04362973	absent	absent
43911	Human	-	SAMN04363677	absent	absent
1511	Human	-	SAMN04363702	absent	absent
181077	Human	2015	SAMN04600250	absent	absent
117270	Human	2015	SAMN04600422	absent	absent
122646	Human	2015	SAMN04600548	absent	absent
330501	Food	2016	SAMN06247639	absent	absent
329908	Food	2016	SAMN06278344	absent	absent
264138	Human	2016	SAMN06680365	absent	absent
318351	Human	2016	SAMN06680371	absent	absent
241785	Human	2016	SAMN06680376	present*	present
227035	Human	2016	SAMN06680377	present*	present
261432	Human	2016	SAMN06680380	absent	absent
341373	Human	2017	SAMN06680383	absent	absent
275941	Human	2016	SAMN06680385	present*	present
173552	Human	2015	SAMN06680388	absent	absent
120576	Human	2015	SAMN06680391	absent	absent

322283	Human	2016	SAMN06680392	present*	present
198252	Human	2015	SAMN06680393	absent	absent
243841	Human	2016	SAMN06680394	absent	absent
304435	Human	2016	SAMN06680415	absent	absent
260043	Human	2016	SAMN06680461	absent	absent
169633	Human	2015	SAMN06680462	present*	present
260046	Human	2016	SAMN06680463	absent	absent
301659	Human	2016	SAMN06680467	absent	present*
152469	Human	2015	SAMN06680468	absent	absent
212745	Human	2016	SAMN06680481	absent	absent
232778	Human	2016	SAMN06680483	absent	present
283182	Human	2016	SAMN06680484	absent	absent
190645	Human	2015	SAMN06680485	absent	absent
205616	Human	2016	SAMN06680490	absent	absent
205615	Human	2016	SAMN06680491	present*	present
163120	Human	2015	SAMN06680497	absent	absent
302890	Human	2016	SAMN06680498	absent	absent
280310	Human	2016	SAMN06680499	absent	absent
271333	Human	2016	SAMN06680500	absent	absent
293044	Human	2016	SAMN06680501	absent	absent
267755	Human	2016	SAMN06680502	present*	present
115262	Human	2015	SAMN06680503	present*	present
262734	Human	2016	SAMN06680504	absent	absent
320993	Human	2016	SAMN06680505	absent	absent
109911	Human	2015	SAMN06680506	present*	absent
320327	Human	2016	SAMN06680507	absent	absent
273570	Human	2016	SAMN06680508	present*	present
129502	Human	2015	SAMN06680527	absent	absent
289056	Human	2016	SAMN06680528	absent	absent
196424	Human	2015	SAMN06680529	absent	absent
265324	Human	2016	SAMN06680530	absent	absent
330531	Human	2017	SAMN06680531	absent	absent
173721	Human	2015	SAMN06680532	absent	absent
170294	Human	2015	SAMN06680533	present*	present
140702	Human	2015	SAMN06680534	absent	absent
320273	Human	2016	SAMN06680535	absent	absent
223521	Human	2016	SAMN06680536	absent	absent

248576	Human	2016	SAMN06680537	absent	absent
168342	Human	2015	SAMN06680538	absent	absent
171284	Human	2015	SAMN06680539	absent	absent
221652	Human	2016	SAMN06680547	absent	absent
164936	Human	2015	SAMN06680548	absent	absent
282430	Human	2016	SAMN06680549	absent	absent
113041	Human	2015	SAMN06680550	present*	present
265285	Human	2016	SAMN06680551	absent	absent
178848	Human	2015	SAMN06680552	absent	absent
103725	Human	2015	SAMN06680581	absent	absent
367239	Human	2017	SAMN07155141	absent	absent
252404	Food	2016	SAMN07674568	present*	present
255574	Food	2016	SAMN07674569	present*	present
376310	Human	2017	SAMN07674570	present*	present*
385795	Human	2017	SAMN07674571	present*	present
373952	Human	2017	SAMN07674572	absent	present
389602	Human	2017	SAMN07674573	absent	absent
87407	Food	2015	SAMN07674574	absent	absent