Genome Evolution of *Salmonella enterica* serovar Typhimurium

A thesis submitted to the School of Biological Sciences at the University of East Anglia in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Oliver James Daniel Charity

03.2021

Contents

Abstract

Salmonella enterica serovar Typhimurium (*S.* Typhimurium) is a food-borne pathogen which accounts for ~25 % of non-typhoidal *Salmonella* infections, usually through faecal-oral transmission. To maintain food safety and avoid unnecessary human and animal disease, understanding how pathogens such as *S.* Typhimurium evolve and their requirements for persistence is paramount. This study investigated genomic evolution of *S.* Typhimurium. First, a recently acquired novel genomic island was characterised that initiated successful clonal expansion of pandemic monophasic *S.* Typhimurium ST34, *Salmonella* genomic island 4 (SGI-4). SGI-4 was discovered to be self-transferable and enhance resistance to heavy metals. This was consistent with the epidemiology of monophasic *S.* Typhimurium ST34 strains as they are most frequently isolated from pigs and copper is used as a growth promoter in pig farming. Next, the phage sensitivity of *S.* Typhimurium lineages was investigated through associating phage sensitivity with whole genome sequence (WGS). DT193 that is characterised by resistance to all Anderson typing scheme phage preparations was the most frequent majority phage type, consistent with selection for phage resistance occurring many times throughout *S.* Typhimurium. The monophasic *S.* Typhimurium ST34 population was then investigated using large-scale WGS-based statistical analyses. A novel prophage termed mTmII was associated with continuation of the monophasic ST34 lineage and acquisition was accompanied by an increase in phage resistance. How phage predation shapes the evolution and persistence of *S.* Typhimurium pathovars in host animal niches was then investigated and a novel phenomenon enabling rapid resistance to phage predation through polymorphic gene variation discovered, including a base-specific deletion of serogroup B1 O-antigen polymerase encoding *wzy*, coupled with recombination mediated reversion of the locus in a mixed population of a single strain.

List of Figures

List of Tables

Abbreviations & Definitions

Abi: Abortive infection; altruistic cell death of a minority of a bacterial population to inhibit phage spread after infection.

AMR: Anti-microbial resistance

Antigen: A target for antibody binding, usually on the cell surface.

Biphasic: Consisting of 2 flagella protein encoding genes, which can be 'on' or 'off'; single cells exhibiting phase 1 or 2, but both existing in a population (Bi- preferred over Di-; Bi- being two in Latin, the more commonly used language for scientific terminology).

bp: base pairs, with standard increase of 10^3 interval metric prefixes, Kb (kilobase) Mb (megabase).

BrEX: BacteRiophage EXclusion; an innate, widespread, restriction modification based, anti-phage system.

cas /Cas: CRISPR associated gene/or protein, with standard terminology; italicised non-capital letter referring to the gene; capital first letter, non-italicised referring to the expressed protein.

Cattle: Animals of the species *Bos taurus* .

CRISPR: Clustered, regularly interspaced, short palindromic repeats; an adaptive anti-phage system.

CIMES: *cis*-integratng mobilizable elements; mobile genetic elements which retain sequence specific attachment sites and can be transferred by other mobile element machinery.

Δ (delta) bitscore: Quantified effect of an altered protein sequence compared to a reference; for quantifying amino acid substitutions, HDCS, and gene degradation.

DNA: Deoxyribonucleic acid.

DT: Definitive type (phage type).

E. coli: Escherichia coli.

Effector protein: A protein secreted from one cell to another which modulates activity within the recipient cell.

Genomic island: A mobile genetic island which has integrated into a genome, typically those observed frequently enough to be considered a genomic feature due to positive selection.

GWAS: Genome-wide association.

HDCS: Hypothetically disrupted coding sequences; nucleotide sequences that display traits such as multiple premature stop codons and hypothetical amino acid changes that disrupt protein function, hypothetically inhibiting protein function through truncated or tertiary structure altered amino acid sequences.

h: hours.

HGT: Horizontal gene transfer.

ICE: Integrative and conjugative element; chromosomally integrating elements which self-transfer through a T4SS.

IS: Insertion sequence element; a transposable DNA element that excises and integrates, usually consisting of 2 or 3 genes and possible co-transferred genes.

iNTS: Invasive non-typhoidal *Salmonella.*

l : Litre, with standard metric prefixes, pl(pico), nl(nano), μl(micro), ml(mili).

LRT: Likelihood ratio test; for testing model fit by comparing log-likelihoods, usually nested data where it is unknown if heavily parametrized models are more significant.

M: Moles per litre

m: Meter, with standard 10^3 interval metric prefixes, pm(pico) ,nm(nano), μm(micro), mm(mili), km(kilo).

MDR : Multi-drug resistant.

MGE: Mobile genetic element; any section of DNA which can mobilise by some mechanism e.g. phage, ICEs, CIMEs, IMEs, plasmids, IS elements.

MIC: Minimum inhibitory concentration; the minimum concentration required of an antimicrobial agent to inhibit growth of a micro-organism.

ML: Maximum likelihood; estimating unknown parameters for observed data to maximise the likelihood function.

MLST: Multi-locus sequence type(-ing); a method for intra-species discrimination, typically based on amplicon sequencing of 7 housekeeping genes.

mol: Moles, with reference to 1 mole of carbon-12 weighing 12 grams and containing $6.02214076 \times 10^{23}$ carbon molecules, with standard 10^{-3} interval decreasing metric prefixes mmol(mili), μmol(micro), nmol(nano), pmol(pico).

Monophasic: Consisting of only phase 1 flagella gene protein expression, usually FliC, suggesting deletion of the phase 2 encoding locus *fljAB*.

Monophasic *S.* Typhimurium ST34 : Specified monophyletic pandemic lineage of a flagellar phase 1 (FliC) monophasic *Salmonella enterica* subspecies *enterica* (I) serovar 4,[5],12:i:- of sequence type (ST)34, distinct from other monophasic *S.* Typhimurium lineages, e.g. the 'Spanish clone' which is monophasic but not ST34.

MRCA: Most recent common ancestor.

n: Quantity; number of (strains).

Neothilization: The transition of humans from hunter-gatherer culture to farming and agriculture, which occurred ~6,000 years ago.

NTS: Non-typhoidal *Salmonella.*

ORF: Open reading frame.

Pathovariants: Closely related isolates of an infectious disease which have distinct lineages that cause varying disease states.

Patristic distance: Quantified phylogenetic core SNP distance between two strains; a smaller cladal patristic distance indicative of increased clonality and therefore positive selection.

PCR: Polymerase chain reaction.

Phage: Bacteriophage, for this study specifically tailed phage of the order *Caudovirales.*

Phage typing: Discrimination of serovars into groups (phage types), by challenging a strain with a panel of phages; pathogen surveillance technique used for outbreak detection.

Poultry: Animals of the order Galliformes, particularly *Gallus gallus*.

qPCR: Quantitative polymerase chain reaction (a.k.a. quantitative real-time polymerase chain reaction, here to avoid confusion, rtPCR: reverse transcription PCR; qPCR: real-time/quantitative PCR).

Ri: Resistance index of a phage type, in this study with respect to the Anderson phage typing scheme.

RNA: Ribonucleic acid.

s: seconds, with standard decrease of 10^3 interval metric prefixes ms (mili), μs (micro), ns (nano).

S. bongori: Salmonella bongori.

S. enterica: *Salmonella enterica*.

Serovar: A subdivision of a subspecies based on cell surface structures through challenge with different antibodies (anti-sera), producing agglutination if the antibody target is present; the process known as serotyping (*Salmonella* typing with serum antibodies).

SGI: *Salmonella* Genomic Island.

SLE: SGI-4-like elements.

SPI: *Salmonella* Pathogenicity Island.

spp: species.

SNP(s): Single nucleotide polymorphism(s).

ST: Sequence type; as determined through amplicon sequencing of 7 housekeeping genes.

S. Typhimurium: *Salmonella enterica* subspecies *enterica* (I) serovar Typhimurium.

SVC: *Salmonella* containing vacuole.

T3SS: Type III secretion system; a tubular protein injection-like apparatus, usually for secreting effector proteins into eukaryotic cells.

T4SS: Type IV secretion system; protein machinery for conjugative DNA transfer through DNA binding, translocation to pilus and pilus transfer into recipient cells.

T6SS: Type VI secretion system; intracellular membrane attached protein machinery for transferring proteins and puncturing holes within other bacteria, analogous to a contractile phage tail.

TA: Toxin-antitoxin; dual systems with many different functions, for example genome stability and abortive infection after phage detection.

Typing phage: Phage of the Anderson typing scheme for *S.* Typhimurium

U: Undefined (phage type).

WGS : Whole genome sequence(s).

WT: Wild type; a naturally existing allele of a gene, or unaltered ancestral allele within a lineage.

XDR : Extensive-drug resistant.

Zoonotic: Pathogen transmission from an animals to humans, (plural: Zoonoses).

Acknowledgements

I would like to acknowledge: i) the United Kingdom Research and Innovation council and Quadram Institute Bioscience for funding this project; ii) Norwich Bioscience Institutes staff for training opportunities and keeping the institutes & laboratories running; iii) Public Health England and the Animal and Plant Health Protection Agency for their collaboration on this project, particularly Tim Dallman, Elizabeth de Pinna, and Liljana Petrovska; iv) Robert Kingsley for his unwavering support and always making time for discussion when I barge into his office looking for advice; v) Mark Reuter for initiating my molecular biology training; vi) Mark Kirkwood and Priscilla Branchu for being impervious to ridiculous questions and giving sound advice toward the beginning of my PhD and help with data and experiments; vii) Matt Bawn for sharing his ideas, data, useful review feedback, informatics advice, and keeping me grounded; viii) Gaetan Thilliez for guidance, thoughts, support, computer handywork, and molecular biology advice, despite not formally part of my supervisory team; ix) Jennifer Tanner for sharing ideas, expert laboratory help, experimental guidance, and molecular biology advice; x) Luke Acton and Haider Al-Khanaq for expert lab help and testing my hypotheses with; ix) Eleonora Tassinari for expert advice, methods, trees, and coffee breaks; xi) Martin Lott and Michaela Matthews for expert computer programming help and losing at pub quizzes (except for when we won the *Salmonella* round); xii) Anne Charity and Fiona Newberry for exhibiting saint-like patience toward my temperamentality during this time, xiii) Archer Charity, Nicholas Charity, David Charity, Julia Charity, Katherine Seton, Daniel Yara, Adam Waring, Alex Harris, and Joseph Bond for being supportive friends & family through this period of my life, helping in a plethora of different ways.

Chapter I: Introduction

I.1 *Salmonella* taxonomy

Salmonellae are motile, Gram-negative, facultative anaerobic, non-spore forming, rod shaped bacteria, with characteristic peritrichous flagella (1). *Salmonellae* are of the phylum Proteobacteria, class Gammaproteobacteria, order Enterobacterales (2), and family Enterobacteriaceae (3). *Salmonellae* cells are typically 0.7-1.5 by 2.0-5.0 μm in size. The genus consists of two species: *Salmonella enterica* and *Salmonella bongori*, formally subspecies V (4). *S. enterica* contains seven well characterised subspecies: *enterica* (I), *salamae* (II), *arizonae* (IIIa), *diarizonae* (IIIb), *houtenae* (IV), *indica* (VI), and subspecies VII (5) (Figure I.1). A study of the whole genome sequence (WGS) of 926 isolates of *Salmonella enterica* discovered up to nine extant subspecies, indicating three novel subspecies (4). A study found that subspecies *houtenae* is composed of two lineages arising though hybridisation events, now termed *houtenae* A and *houtenae* B (6). The study also reported subspecies *salamae* presents two lineages now designated *salamae* A and *salamae* B. *Salmonella* is further subtyped into >2500 serovars using the Kaufman-White serotyping scheme, based on antigenic differences using anti-sera that bind lipopolysaccharide (O-antigen), flagella (H-antigen), and a capsular, heat-labile, polysaccharide, somatic virulence antigen (Vi antigen) present in serovars Typhi, Paratyphi C, and some Dublin (7).

I.2 Evolution of the genus *Salmonella*

"Nothing in biology makes sense, except in the light of evolution" – Theodore Dobzhansky (8).

I.2.1 Divergence from *Escherichia coli*

Salmonella spp. and *Escherichia spp.* are estimated to have shared a common ancestor between 10-100 million years ago (9-12). This is evident from nucleotide sequence identity and estimated molecular clock rate (2). WGS alignment of *S.* Typhimurium LT2 and *E.coli* K12, prototypical strains from these genera, showed 71 % shared genome coverage with nucleotide sequence identity of 89.83 % (13). Their common ancestor was probably a commensal enteric bacterium of early Cretaceous to late Miocene cold-blooded (ectothermic) animals, possibly retaining evolutionary fitness through enteric modulation and

adaptation. Enteric bacteria are commonly facultatively anaerobic (14). Divergence of *Salmonella* and *Escherichia* may have begun with horizontal transfer of a ~40kb pathogenicity island designated *Salmonella* Pathogenicity Island 1 (SPI-1), by the *Salmonellae* lineage (18). Further divergence was accompanied through acquisition of SPI-4, SPI-5, type 1 fimbriae (T1F), long polar fimbriae (*lpf*), *ttr* and *eut* (Figure.I.1) (15-21). SPI-1 acquisition allowed adaptation to a new intracellular niche resulting in enteric pathogenicity through type three secretion system (T3SS) mediated effector protein secretion, allowing host cell modulation and invasion (22), prompting inflammation. Type one fimbriae, long polar fimbriae, and SPI-4 encoded giant adhesin (*siiE*) promote adhesion to epithelial cells of the intestine, enhancing cellular attachment and allowing T3SS mediated host cell uptake (23, 24).The *ttr* locus encodes five genes, *ttrACBSR,* which encode protein machinery for reduction of tetrathionate, a common molecule produced during gut inflammation. *Salmonella* utilises tetrathionate as a terminal electron acceptor in the respiratory electron transport chain, producing thiosulfate (25). Inflammation associated oxygen free radicals regenerate tetrathionate from thiosulfate, producing more tetrathionate, resulting in a positive feedback loop aiding growth and gut domination (21, 25). The acquisition of 17 gene *eut* locus facilitated utilisation of ethanolamine as a carbon and nitrogen source (26). This contributed to establishing *Salmonella's* pathogenic lifestyle by generating competition with host commensals and epithelial cells (27). Both *S. bongori* and *S. enterica* encode these features, evincing that transfer into an ancestor occurred early in their evolution before speciation (Figure I.1).

I.2.2 Evolution of *Salmonella enterica* subspecies

Evolution of *S. enterica* was likely initiated through acquisition of SPI-2, increasing capacity for intracellular phagocyte survival and replication (15, 28, 29), but most strains associated with human and livestock infection belong to subspecies *enterica* (I). Evolutionarily this may be associated with acquisition of ~100 subspecies I specific genes, including those encoded on CS54, an island involved in intestinal colonisation. This island encodes proteins such as fibronectin binding outer membrane protein ShdA which is thought to increase basic case reproductive number allowing prolonged faecal shedding in warm blooded animals (30). Long polar fimbriae is present in *S. bongori* and subspecies I, but deleted from an ancestor of subspecies II, IIIa, IIIb, IV, and VI (Figure I.1, red lines) (15). Subspecies I

divergence is also associated with acquisition of type one fimbriae (31), and plasmid encoded virulence loci, *spv* (32, 33). Subspecies I, VI, II, and IIIb are proposed to have emerged after IIIa, IV, and VII (38), by the acquisition of *fljAB-hin* loci encoding a second flagella protein (FljB), DNA-invertase (hin), and a protein (FljA) capable of inhibiting expression of the phase 1 flagella gene FliC (34, 35). Hin inverts the *fljAB* promoter inhibiting expression of FljB and removing inhibition of FliC expression (36, 37). This phase variation renders subspecies I, VI, II and IIIb 'biphasic' in flagella gene expression, with IIIa, IV, VIII and bongori only encoding FliC. How biphasicity is beneficial is unknown. Deletion of the phase 2 locus from subspecies I serovars such as Typhimurium (I:4,[5],12:i:-) and Paratyphi B (I:4,12:b:-) had no apparent detriment to virulence or host range (34, 38, 39). This arouses suspicion that it is beneficial in certain situations, such as if FliC becomes compromised through use as a receptor by phage or recognition by immune systems, but can be lost without apparent loss of fitness, questioning our understanding of the requirements for phase variable flagellar proteins (40, 41).

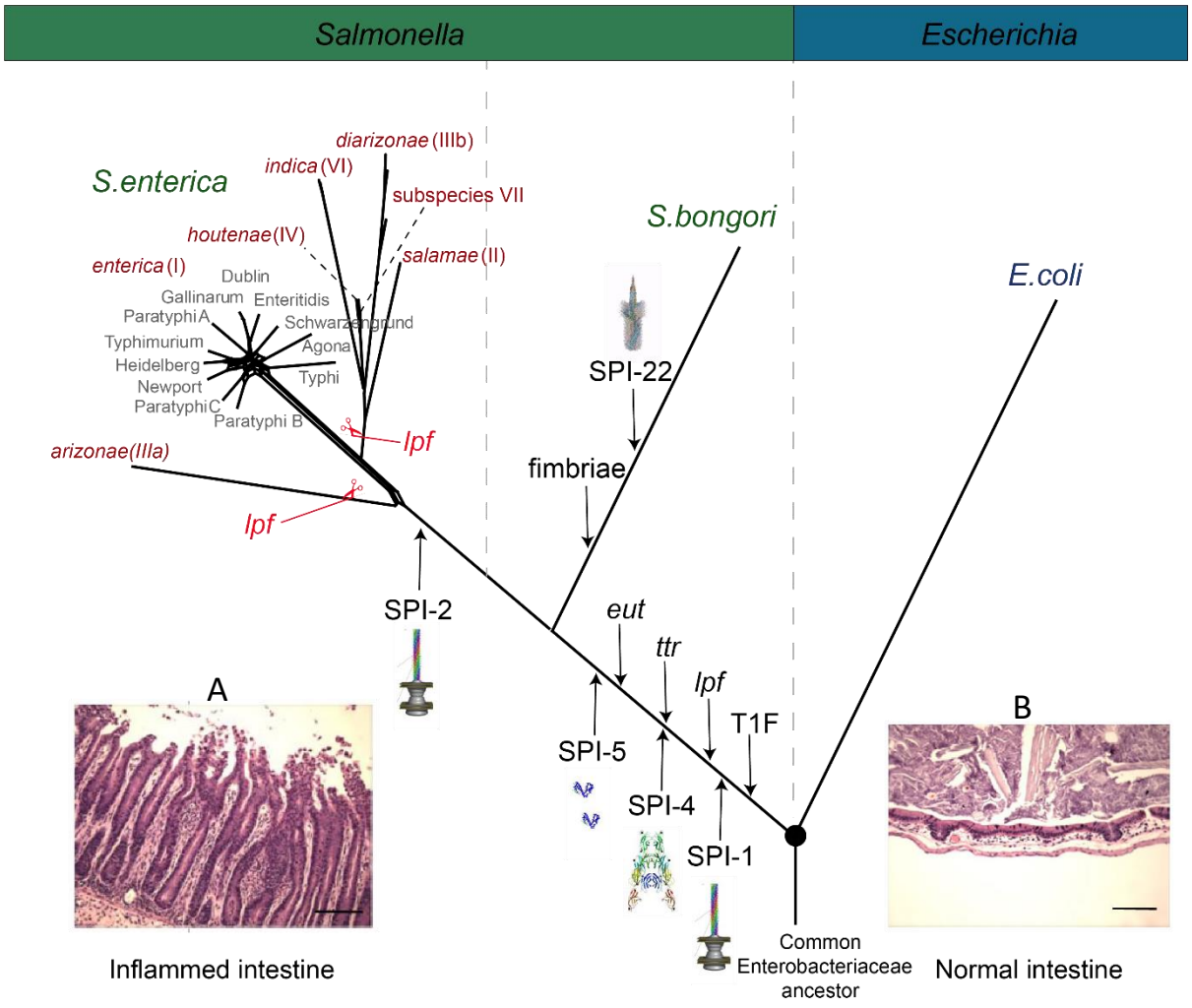


Figure.I.1 Divergence of the *Salmonella* and *Escherichia* genera from a common ancestor. The phylogenetic tree was constructed using sequence variation in 7 gene fragments used for MLST for *E.coli*, *S. bongori*, *S. enterica* subspecies and 12 serovars from subspecies I. Arrows indicate acquisition of a genomic locus, and the structure of the type of secretion system is shown with each SPI acquisition. The red line specifies loss of 3 fimbrial regions from subspecies II, III, IV, VI & VII, which are present in subspecies I and *S. bongori*. Red font indicates subspecies lineages, and grey font indicates serovars. (A) Stained histological cross-section of gut tissue after *Salmonella* infection. (B) Stained histological cross-section of large intestine tissue in absence of infection.

1.2.3 Evolution of subspecies I serovars

Of 2557 known serovars of *Salmonella,* 1531 (59.8 %) are from subspecies I (42). That subspecies I evolved as a pathogen of warm-blooded animals may explain this disproportionate percentage due to increased likelihood of sampling. Subspecies I has three host range categories: host restricted serovars, host adapted serovars, and host generalist serovars (43-47). Host restriction refers to a high degree of host adaptation which has restricted the lineage to circulating in the population of a single host (48). Host adaption is often associated with changes in pathogenicity, albeit not an essential criterion. Variation in serovar host adaptation creates zoonotic potential for *Salmonellae* as well as for other bacterial pathogens, where the species barrier is crossed. This has implications for disease severity, as a host adapted pathogen may exhibit increased virulence with hosts to which they are not adapted (48, 49).

1.2.3.1 Host restricted serovars

S. Typhi is a prototypical host-restricted serovar, and its origins of circulating in humans is associated with Neothilization (50). WGS analysis of host-restricted serovars *S.* Typhi and *S.* Paratyphi A showed hypothetically disrupted coding sequences (HDCS), a hallmark of host adaptation and restriction (51). Adaptation negatively selects genes resulting in HDCS, for example, within genes encoding biosynthetic pathways for molecules which are abundant in the host, due to the energy cost of expressing genes not required in the niche (44, 52). WGS of human restricted serovar Typhi strain CT18 showed >200 HDCS compared to *S.*Typhimurium LT2 (51). Human host adaptation is also associated with addition of extra

genes through HGT (51). *S.* Typhi CT18 encoded 601 extra genes compared with *S.* Typhimurium LT2 notably including SPI-7 (51). SPI-7 encodes the Vi antigen which separates *S.* Typhi outer membrane bound structures from the innate immune system, reducing the host inflammatory response and aiding dissemination (43). Serovar Gallinarum causes disseminated typhoid-like disease and dysentery in poultry (53). It has two distinct variants forming separate linages, termed 'biovars', Pullorum and Gallinarum (54). The former produces dysentery-like disease, and the latter typhoid-like disseminated disease (54). Evolutionary traits which probably began *S.* Gallinarum host restriction included unique coding sequences within SPI-19 introduced via recombination, and single nucleotide polymorphisms (SNPs) in SPI-1 & 2 effectors (55).

I.2.3.2 Host adapted serovars

Well characterised and frequently isolated animal adapted serovars of subspecies I include Dublin, Choleraesuis and Paratyphi C, which are cattle (Dublin) and pig (Choleraesuis and Paratyphi C) adapted respectively (54, 56, 57). Due to their host adaptation these serovars are of great concern to their respective agricultural industries and producers of related food products. They are considered host-adapted as they initiate disease in humans as well as their host reservoirs (58, 59). *S.* Dublin causes disseminated disease in cattle, displaying host adapted traits of genome degradation and HGT (60). *S.* Dublin is pathogenic to humans via zoonoses (61), and infections often result in invasive disease primarily in immunocompromised people (62). Evolution of *S.* Dublin to cattle adapted lifestyle included recombination within a plasmid similar to that of other serovars, where 10 Kb of content was deleted, and an 11 Kb region acquired encoding specific fimbriae and virulence genes such as *vagC* and *vagG* (63, 64). *S.* Dublin also acquired T6SS encoding SPI-19, where it is thought to be involved in competitive colonisation through secreting toxic proteins into competing bacteria (65). SPI-19 is observed in *S.* Dublin, *S.* Gallinarum, and *S.* Enteritidis, and the three serovars share a common ancestor, suggesting a single acquisition perhaps driving evolution of the lineage containing the three serovars (66). Serovar Choleraesuis causes disease in pigs and is pig adapted, circulating widely in pigs as a reservoir (48). It also causes severe invasive disease in humans, with 74.1 % of Choleraesuis infections isolated from blood (48). Serovar Choleraesuis exists within a lineage sharing a common ancestor with Paratyphi C. WGS of

Choleraesuis strain SC-B67 revealed 151 HDCS, most of which encoded genes involved in chemotaxis signalling pathways and over expression of AcrAB through truncation of coding sequence of regulator *acrR* due to a premature stop codon providing ciprofloxacin resistance (67). *S.* Paratyphi C exhibits diverse genomic structures with signatures of primary pig adaptation; isolate WGSs are closely related with serovar Choleraesuis (68-70). This suggests these serovars shared a common ancestor undergoing pig adaptation, but *S.* Paratyphi C diverged. This may have involved filling a niche in different hosts while undergoing accumulation of HDCS, possibly explaining its observed pathogenicity (59). Accordingly, phylogenetic analysis suggested these serovars shared a common ancestor, being more closely related to each other than to other subspecies I serovars (59).

I.2.3.3 Host generalist serovars

Many serovars are commonly isolated from multiple hosts, and are considered generalists (56). Two serovars considered broad host range are Enteritidis and Typhimurium (56). *S.* Enteritidis shared a recent common ancestor with *S.* Gallinarum, and accordingly has been associated with poultry (71). Evolutionarily, colonisation of chickens with *S.* Enteritidis requires specific fimbriae; fimbriae diversity being a common theme throughout *Salmonellae* evolution (72). *S.* Enteritidis infections are highly associated with contaminated eggs (71). Mechanisms, such as reinforcing the bacterial cell wall and upregulating expression of proteins for repair, are hypothesised to allow Enteritidis egg contamination and survival, suggesting novel microevolution during this process (71). Both serovars can infect a wide range of hosts, but with varying disease severity, including asymptomatic carriage (73). The broad host nature of *S.* Enteritidis is exemplified through observation of isolates within mice that inhabit layer flock coups, also suggesting secondary methods of *S.* Enteritidis transmission between animals (74). Both *S.* Enteritidis and *S.* Typhimurium adapt to hosts, and due to this both serovars contain lineages which exhibit human invasive disease, particularly in immunocompromised people, children, and the elderly (39, 75, 76). Invasive *S.* Enteritidis has been linked to global locations with low-income settings, also suggesting separate lineages with novel microevolutionary features (47). *S.* Typhimurium is more typically a host generalist serovar, which is exemplified by its isolation from disease cases of all domestic

animals (56). Despite *S.* Typhimurium typifying a host generalist serovar, some lineages exhibit host adaptation, for example in pigeons (77), pigs (78), and humans (79-81).

I.2.4 Incidence of *S. enterica* serovars and estimated burden

Salmonellae are a major problem in low and middle income countries where sanitation is poor, with higher incidence of human-adapted typhoidal and paratyphoidal serovars Typhi, Sendai, Paratyphi A, B, and C, which circulate globally causing ~2,500,000 infections resulting in 65,000-200,000 deaths each year from invasive, disseminated disease (82).

Non-Typhoidal *Salmonella* (NTS) are a significant global burden (82-84). The estimated incidence of NTS was 93.8 million cases per year world-wide in 2010, with 155,000 deaths (84). This amounts to significant economic burden from mortality and h lost through sickness and medical expenses (85-87). Invasive Non-typhoidal *Salmonella* (iNTS) also contribute to global economic loss, and more significant morbidity, especially in low-income countries (88-90). These *Salmonellae* are mostly lineages of *S.* Enteritids or *S.* Typhimurium, and are commonly associated with prior disease in poultry (47, 91).

In the United States of America (USA) ~40,000 cases are reported each year to the Centre for Disease Control, (CDC), Atlanta, Georgia (92). From culture-confirmed and unconfirmed infections the CDC estimated *Salmonella* infections cause 15,000 hospitalisations and 400 deaths annually in the US (93). The most prevalent serovar in the USA in 1995 was *S.* Enteritidis (I,1,[9],12:gm:1,2) accounting for 24.7 % of infections, closely followed by *S.* Typhimurium with 23.5 % of infections. Together with serovars Newport (6.2 %) and Heidelberg (5.1 %) these constituted 60 % of all *Salmonella* infections at this time. This serovar dominance was also distinct in Germany, with *S.* Enteritidis (64.3 %) and *S.* Typhimurium (19.9 %) making up 80 % of all isolates reported to the Robert Koch Institute in 2004 (94).

In the United Kingdom (UK), Public Health England (PHE) had 104,549 confirmed *Salmonella* cases from 1990-2013, suggesting more than 1 million infections when including unreported cases (95). *S.* Typhimurium is prevalent in the United Kingdom (UK) (57). A resurgence of *S.* Enteritidis was observed from 1980-2001 mostly infecting people through contaminated eggs and chicken (95). This

dominance was likely due to asymptomatic disease in chickens when colonised with *S.* Enteritidis, enabling undetected spread (96, 97). In 2018 *S.* Typhimurium and *S.* Enteritidis accounted for 51.3 % of all laboratory confirmed cases (Figure I.2) (56, 57). In UK animals *S.* Typhimurium is more widespread than any other serovar, being second most common in cattle and sheep after serovar Derby and IIIb,61:k:1,5,[7], respectively (FigureI.2, below) (56). *S.* Typhimurium, including its monophasic variants, are the most common serovars isolated in pigs (79.9 %), and isolation of *S.* Typhimurium from all animal sources has risen by 119 % from 2015-2018 (56).I

In Europe Salmonellosis was the second most common foodborne infection 2010-2017 with 1,050 outbreaks and 89,162 confirmed cases in 30 countries in 2013 (98, 99). Cases decreased from 2010-2014, but then increased 2014-2017 with 92,649 and 0.17 % case fatality rate in 2017 (98, 99). Notification rate was eight times higher in young children (0-4 years) than adults (25-64). There is a strong seasonal trend toward increased cases in July-August each year (99). The most recent emergence in Europe is of a multi-drug resistant (MDR), flagellar monophasic (I, 4,[5],12:i:-) variant of *S.* Typhimurium (monophasic *S.* Typhimurium ST34), which has become a pandemic (100-107). The number of recorded human cases of 4,[5],12:i:- reached a peak in 2013, has since slightly lowered, but remains a prominent concern.

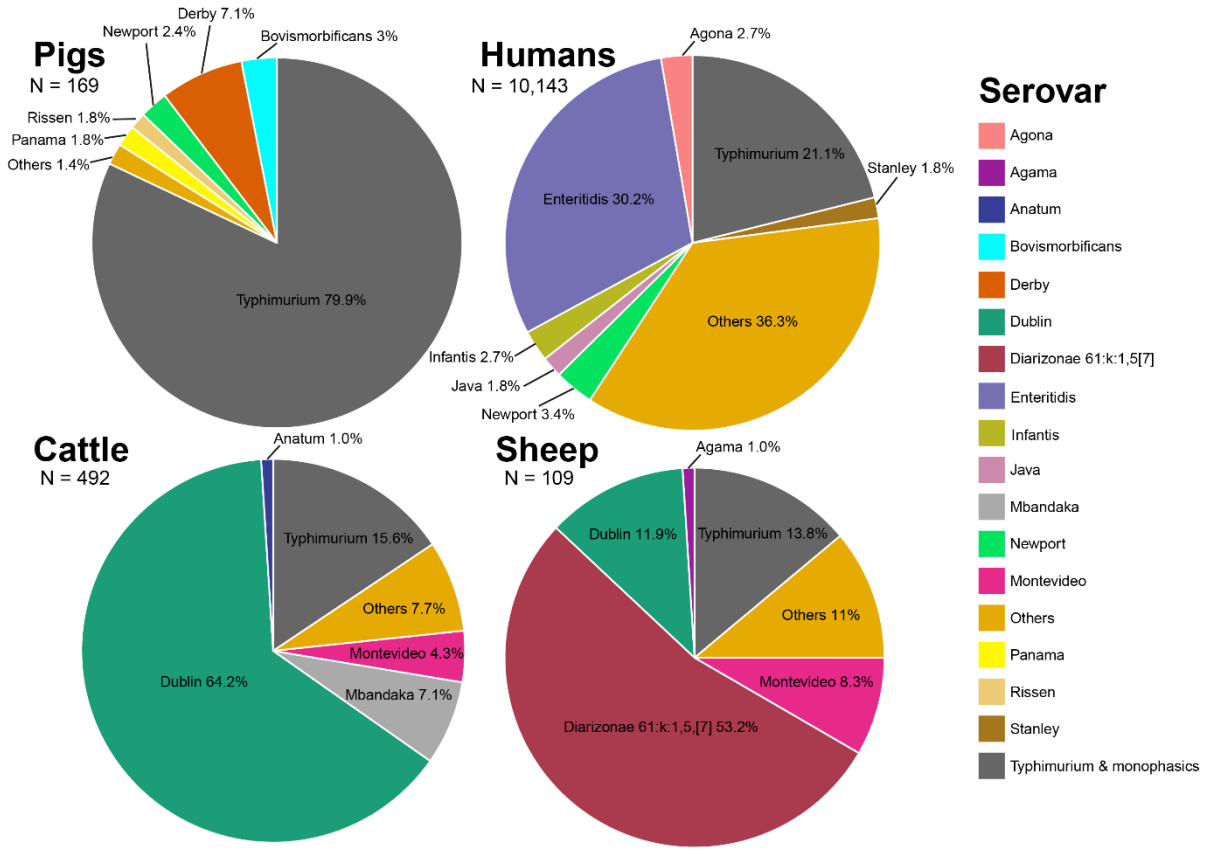


Figure I.2 Percent of *Salmonella* serovars isolated from various sources. Percentages of *Salmonella* serovars isolated from humans (top right), pigs (top left), cattle (bottom left), and sheep (bottom right). Adapted from APHA & PHE, (2019) (56).

I.3 *Salmonella enterica subspecies enterica serovar* Typhimurium

S. Typhimurium and its monophasic variants are present as a monophyletic subclade of *S. enterica*, with a similarly broad range of hosts as that described for *subspecies* I, despite ~1500 core genome SNPs separating the most distantly related strains (39). *S.* Typhimurium is associated with a wide range of disease syndromes amongst its hosts, including minor symptomatic carriage to disseminated disease. Here, the literature on key aspects of *S.* Typhimurium has been discussed, such as human pathogenesis, measures of control, surveillance, and population structure.

I.3.1 *S.* Typhimurium pathogenesis

Infection with *S.* Typhimurium usually causes self-limiting gastroenteritis and faecal shedding for continued faecal-oral transmission (108). Infection is typically through consumption of contaminated

food, lasting two to seven days (109). However, severe disease can occur in children, the elderly & immunocompromised patients (83). Outcome of infection also depends on pathovariants arising from separate lineages of *S.* Typhimurium and its monophasic variants.

Pathogenicity of *S.* Typhimurium is dependent on two T3SS encoded on separate loci: SPI-1 and SPI-2. There are six well characterised SPIs in *S.* Typhimurium (Table I.1), one in typhoidal serovars containing the Vi-antigen (SPI-7), and >20 recorded across the genus (110-117).

The infective dose of orally ingested *S.* Typhimurium can vary but is estimated at ~10^6 cells. This is sufficient to survive gastric acidity enabling access to intestinal epithelium and colonisation of the small and large intestine (118). During colonization the SPI-1 T3SS is utilized for invasion of epithelial cells (119, 120). Several microbial factors induce the immune response. Flagellin and fimbriae are recognised by Toll-like receptors five and two respectively, which initiate transcription factor NF-kB (121, 122). The secretion of SPI-1 effector proteins stimulates nucleotide binding and oligomerization domain (NOD)-like receptors which activate NF-kB and MAPK pathways increasing inflammation. *Salmonella* secretion of toxins such as lipopolysaccharide (LPS) lipid A further promotes inflammation. Proinflammatory cytokines result in influx of neutrophils and transepithelial migration resulting in damage to the mucosal barrier, net fluid loss and inflammatory diarrhoea. Inflammation ultimately helps clear infection, but also promotes spread of the pathogen.

Pathogenicity Island	Size (kb)	Secretion system	Functions	Reference
SPI-1	40	Type 3	Invasion of epithelial cells, SVC development, effectors for actin rearrangement, membrane ruffling, induce interleukin-8 and pathogen provoked epithelial chemoattractant secretion	(123) (124) (125)

SPI-2	40	Type 3	Survival in macrophages, inhibiting SVC-lysosome fusion, inhibits endocytic trafficking, evading macrophage NADPH oxidase-dependant killing, encodes effector, chaperone, and translocon proteins	(126) (113) (127) (123)
SPI-3	17	None	Intramacrophage survival via MgtC and MgtB magnesium transporters	(128) (129)
SPI-4	27	Type 1	Epithelial cell adhesion, *Salmonella* intestinal infection proteins (SiiA-F), SiiE non-fimbrial adhesion protein	(129) (130) (131)
SPI-5	8	None	Encodes SopB effector secreted by SPI1 T3SS, PipB SVC protein translocated by SPI-2 T3SS	(132) (129)
SPI-6	59	Type 6	Gastrointestinal colonisation, T6SS mediated host microbiota killing	(133) (134) (135)

Table I.1 Features and functions of *Salmonella* Pathogenicity islands. Description of *Salmonella* pathogenicity islands commonly associated with strains of *S.* Typhimurium and its monophasic variants. Adapted from Hurley et al. 2014 (117).

Out-competing host microbiota is crucial for pathogenesis. Colonisation can be dependant on the existing microbiota, with certain bacteria and microbiome-derived metabolites shown to provide colonisation resistance against *S.* Typhimurium in a murine model (136). *Salmonella* exploit the inflammatory environment to gain a growth advantage using inflammation associated molecules including nitrate and tetrathionate as electron acceptors (137). *S.* Typhimurium utilises an antibacterial type six secretion system (T6SS) to puncture host microbes and secrete effector toxins, aiding intestinal establishment (138-140).

Intracellular infection prolongs inflammation, providing a competitive advantage for *Salmonellae,* making *Salmonella* intracellular survival a fundamental part of pathogenesis (141). *S.* Typhimurium is able to survive within many mammalian cell types and regularly resides in macrophages, m-cells, dendritic cells , and epithelial cells (142). Internalisation of phagocytic cells is through receptor-mediated phagocytosis. When in contact with non-phagocytic cells, such as an epithelial cell, invasion initiates with adhesion through fimbrial and non-fimbrial adhesins such as BapA, SiiE, ShdA and MisL, surface lipopolysaccharides (LPS), and lysine methylation of surface exposed flagella FljB and FliC (143). SPI-1 is crucial for adhesion, invasion and early-phase post invasion (144). After adhesion the SPI-1 T3SS translocates effector proteins through the host cell membrane. Expression of SPI-1 genes are spatiotemporally controlled for expression at optimal times and host intestinal long chain fatty acids modulate SPI-1 T3SS expression (145, 146). Effector proteins SopB, SopE and SopE2 help mediate epithelial invasion by activation of RHO GTPase-dependant actin rearrangements at the host membrane causing membrane ruffling (147). Internalization is further promoted by SipA and SipC through direct actin binding at the T3SS insertion site (translocon). SipA promotes bundling and actin polymerisation, and SipC bundles and nucleates actin at the translocon site. The result is *Salmonella* uptake via macropinocytosis, forming the *Salmonella* Containing Vacuole (SVC), in which *Salmonella* cells survive and replicate. The host membrane morphology is subsequently restored by effector SptP (148, 149). In the SVC SopB recruits RAB5 and phosphatidylinositol-3-kinase VPS34 generating phosphatidylinositol-3-phosphate that decreases acidic lipids on the SVC interrupting lysosome fusion, enabling SVC persistence (150).

SPI-2 is essential for intracellular macrophage survival and allows disease progression and phagocytosis evasion. SPI-2 encodes a second T3SS that secretes proteins across the SVC membrane, as well as effector proteins SseG, SseF, SifA, SipA, PipB2, and SseJ (151). These effectors contribute to SVC maturation and formation of *Salmonella*-induced filaments (*Sif)* that are tubulovesicular structures generated during microbial replication (152). Recently, spoT has been recognised as essential to promote intracellular phagosome virulence programs through synthesising signalling molecule guanosine pentaphosphate (pppGpp) from guanosine tetraphosphate (ppGpp) (153). Within SVCs a

proportion of the *Salmonella* become non-replicating 'persister' cells through acetylation of charged tRNA by a toxin molecule that inhibits protein synthesis and allows survival in phagocytic SVCs (154). Subsequently the acetylated tRNAs are reverted by an antitoxin, and protein synthesis restarts. This bacteriostatic propensity allows resurgence of infection after antibiotic treatment, causing chronic and reoccurring disease.

I.3.2 Current measures of S. Typhimurium control, prevention & surveillance

Prevention of disease reduces economic burden and reduces mortality in people and livestock (155). Existing measures for disease prevention include generating and practising health and safety guidelines for people who handle livestock and food, vaccinations against pathogens, and surveillance to monitor the prevalence and spread of existing and emerging pathogens (155-157). Agricultural methods of reducing livestock infection and retaining yields of meat and dairy products in the European Union included antibiotic use until being banned in 2006 (158, 159). Since then changes in animal husbandry have possibly contributed to emergence of different lineages of S. Typhimurium. The use of copper as an antimicrobial and growth promoter in livestock feed superseded antibiotic use (160). Vaccines are also effective at preventing pathogen spread, but no effective vaccine for S. Typhimurium has been produced. S. Typhimurium is controlled in Europe and USA through pathogen surveillance techniques, outbreak detection, and epidemiology. Constant monitoring of infections in humans and animals allows source attribution and subsequent management of how pathogens are persisting and spreading within these sources.

Serotyping is still commonly used to group *Salmonella* isolates (70, 161). *Salmonella* serovars are constantly under surveillance in Europe including the UK (56, 99), For discrimination of within serovar variation of S. Typhimurium, techniques such as phage typing, Pulse-Field Gel Electrophoresis (PFGE), polymerised amplicon sequencing techniques, and more recently WGS have been utilised (162, 163).

I.3.2.1 *S.* Typhimurium phage typing

WGS for surveillance of *S.* Typhimurium in the UK is becoming more routine (164). Prior to WGS techniques such as phage typing were commonly used for outbreak detection (165). Phage typing is a

phenotype-based method which involves determining sensitivity of indicator *Salmonella* strain by overlaying bacterial lawns with a panel of phage lysates (162). Resistant *Salmonella* grow as an 'opaque' lawn of bacteria in the presence of the phage, various levels of sensitivity are indicated by 'semi-confluent lysis', or complete killing generates plaques resulting in 'confluent lysis' (Appendix I.1 for phage typing table and terminology). This leaves behind a certain sensitivity pattern, termed a "phage type" which can be definitive (DT) or undefined (U), with each type having a designated number. Three *S.* Typhimurium phage typing schemes have been developed, the most recent and diverse of which was by Anderson (1977) (165-167). The Anderson phage typing scheme relies on the outcome of challenging isolates of *S.* Typhimurium with 30 different phages producing patterns of resistance and varying degrees of lysis (165). Phage types have been split into more than 300 patterns that include 209 definitive types (DTs: DT1 - DT209), >118 undefined types (Us: U210 – U327), including a further 8 types differentiated between an additional 7 typing phage when resistance to all standard 30 phages is observed. Due to the prevalence of *S.* Typhimurium in the UK, the *S.* Typhimurium phage typing scheme has been used extensively for outbreak detection since the1950s (162). The prevalence of *S.* Typhimurium phage types in the UK has varied since the 1970s, with consecutive most frequently isolated strains exhibiting mostly DT9 (1960-1975), DT204/DT49 (1975-1990), DT104 (1990-2005), and DT193/DT120 (2005-present) as described previously (Figure I.4). These dominant clones are usually isolated within a 'background noise' of various phage types, indicating that there are many lineages of *S.* Typhimurium existing in different environmental niches, and that phage types can be variable within and between *S.* Typhimurium lineages. This important distinction expresses another limitation of phage typing - that phage types are frequently polyphyletic. DT193 is a polyphyletic type that is characterised by resistance to all standard typing phage preparations (168) (169, 170). Resistance occurs over time when a bacterium is repeatedly exposed to a stressor; but the specific phages used for typing are not in the wild to actively select for their resistances, consistent with selection toward a general increase phage resistance occurring naturally, as observed by multiple emergences of DT193, (39, 169).

Few studies have attempted to link phage type to genetic data. Pang et al. (2011) produced SNP profiles using 44 SNPs from previously generated data and six available genomes to analyze 215 isolates,

which grouped into 33 SNP profiles with four distinct phylogenetic clusters (171). Fourteen phage types differentiated into multiple SNP profiles in different phylogenetic clusters, suggesting phage types arose independently multiple times (171). The evolutionary relationship between phage types and the molecular basis of interactions with typing phage have not been well described. One study sought to determine relatedness and molecular characteristics of the Anderson typing phages (172). They differentiated phage lysates into 8 proposed genealogical groups (A-H) after propagation on available corresponding Anderson scheme Propagating Bacteria (PBA1-PBA32) and subsequent ecoRI restriction digest and electrophoresis for molecular weight comparison of genomic fragments. It was determined that most of the Anderson scheme typing phages are Lambdoid p22-like (173).

Differences between phage types result from specific phage resistance profiles. This is the profile which can hypothetically be observed by viewing genome encoded components of bacterial phage resistance mechanisms within isolates. More detail on molecular mechanisms which may be coveying phage lysis and resistance profiles for phage types is in section 1.3.4, below.

I.3.2.2 Other sub-typing techniques

Pulsed-Field Gel Electrophoresis (PFGE) is a restriction-enzyme-based surveillance method widely used for attempted molecular characterization of *S.* Typhimurium and has useful discriminatory power. PulseNet is an international network which allows comparison of results (174). These factors make it a valuable tool for species subtyping and understanding epidemiological dynamics of pathogens like *S.* Typhimurium, for which it has been extensively used (175). Gene amplification techniques utilize polymerase chain reaction (PCR) to amplify a specific genomic section for comparison. One commonly used for *S.* Typhimurium is multi-locus variable number tandem repeat analysis (MLVA), and a more recent development utilizes CRISPR spacers – CRISPR typing (176). MLVA exploits a bacterial genetic trait discovered through whole genome sequencing – many bacteria contain multiple regions with short repeated DNA motifs (177). Weill et al. (2014) showed CRISPR polymorphisms strongly correlate with serotype, subtype and multi-locus sequence type in *Salmonella*, and subsequently developed a typing method using PCR amplification, gel electrophoresis, and PCR product sizing for each of the two CRISPR loci present in *Salmonella* spp. (176). This has been sparingly used in recent studies. A study

from China determined that monophasic *S.* Typhimurium had a CRISPR-type of TST4, although 14.5% of TST4 CRISPR types were present in other *S.* Typhimurium isolates, questioning its validity as a marker for this lineage (178). MLST has been used globally for typing and epidemiology of a variety of pathogenic bacteria (179).

I.3.3 Population structure of *S.* Typhimurium

S. Typhimurium has been considered a model for broad-host-range pathogens (180). Most research has been undertaken with a few strains – 4/74, SL1344, LT2 and ATCC14028. Whether any significant genotypic diversity existed within this serovar was unknown prior to the advent of WGS, with limited diversity between aforementioned reference strains not representing the extant population. It has since been discovered that *S.* Typhimurium is widely spread in zoonotic reservoirs, and in many cases phage types can be associated with specific animal hosts (180). *S.* Typhimurium's broad, natural distribution outside humans, livestock, and wild avian species is evident, for example, with description of turtle adapted *S.* Typhimurium causing turtle renal disease, and an observation within hedgehogs (181, 182).

Phage type	Sequence type	Host range & epidemiology	References
DT9	ST19	Broad-host-range, cattle reservoir, MDR, epidemic 1960-1975	(183)
DT204	ST19	Broad-host-range, cattle reservoir, MDR, epidemic 1970-1985	(183)
DT104	ST19	Broad-host-range, cattle reservoir, MDR, pandemic 1990-2010	(183)
DT193/DT120	ST34	Broad-host-range, porcine reservoir, MDR, pandemic 2005 onwards	(183) (184)
U288	ST19	Porcine reservoir, MDR, epidemic in pigs, rare in humans	(180)
DT8	ST19	Duck and geese associated, several human outbreaks	(180) (185)
DT56var	ST568	Wild bird associated	(180) (186) (187)
DT2/DT99	ST98/ST128	Pigeon associated, endemic in pigeons	(46) (188)
DT160	ST19	Wild birds, extended multi-host outbreak in New Zealand	(189)

ND	ST313	Invasive disease in sub-Saharan Africa, XDR lineages, showing human adapted signatures	(190) (79)

Table I.2 Distinct *S.* Typhimurium lineages and their host ranges. *S.* Typhimurium phage and sequence types associated with animal hosts and niches. Adapted from Branchu et al. (2018) (180)

The diversity and population structure of *S.* Typhimurium from public health WGS data in the UK was reported using 131 representative isolates chosen to encompass the genotypic diversity of this pathogen, identifying signatures of genotypic diversity in more detail through analysis on a reduced dataset (39). Bayesian hierarchical analysis of the phylogenetic relationship identified two higher order (first order) clades designated α and β, the former associated with livestock animals containing three clades of successful epidemic clones, and the latter more associated with wild avian species, with the exception of an epidemic lineage of majority phage type DT204 (Figure I.3). Three ST36 isolates were excluded from the analysis due to high divergence, these isolates possibly representing the divergent G1 clade observed by Zhang et al., (191).

I.3.3.1 Clade α

Clade α (Figure I.3 (A) α) contained nine third order clades, with a high proportion of clade α exclusive pig isolates, and other lineages mainly associated with livestock. Three notable clade α lineages included a porcine epidemic clade that has infrequent reports of human infections (U288, α 12), DT104 associated pandemic lineage (DT104, α 15) which gained traction in the 1990s before reducing in isolation numbers, and the current monophasic *S.* Typhimurium ST34 pandemic clade (*S.* I:4,[5],12:i:-, α 17). Clade α has two possible epidemic clades not well reported within the literature; one of majority type DT193 (α 8) associated with pigs and cattle, and a pig, cattle, and human infection associated clade sharing a common ancestor with the U288 complex (Figure I.3 (A) α 11). There are distinct differences between clade α and β. Clade α lineages tend to be equipped with more AMR genes and fewer HDCS in virulence genes, usually retaining the ancestral allele. This is possibly a counter-signal of host adaptation, as the opposite is observed in host restricted serovars such as *S.* Typhi and *S.* Paratyphi A, which accumulate HDCS in non-essential virulence genes for human invasive lifestyle (192). Clade α has significantly fewer deleterious genes as suggested by a reduced Δbitscore, but also significantly reduced invasiveness index, consistent with mutations separating the clades having

increased likelihood of invasive disease in clade β (Figure I.3 (C)). Conversely, there are reports of iNTS infections in Vietnam from isolates which belong to clade α17, monophasic *S.* Typhimurium ST34, that had regained their phase 2 flagellar locus (193). Clade α also characteristically lacks a copy of gene *sseK* (*sseK3*) that encodes a T3SS effector that glycosylates host protein arginine residues, reducing NF-κB dependant immune responses. It is likely *sseK3* was introduced horizontally to clade β, where it appears recessive to *sseK2* due to more frequently becoming a HDCS (194).

1.3.3.1.1 Porcine associated U288 *S.* Typhimurium (α12)

S. Typhimurium of majority phage type U288 (Figure I.2, α12) form a distinct third order clade within the top of clade α, with S01960-09 as their prototypical reference strain, as used by Bawn et al., (2020) (39). Strains with this phage type have emerged as a persistent MDR epidemic within pigs in the UK, and there is a report of U288 associated outbreaks in Denmark resulting in the death of elderly patients (195, 196). Strains of the U288 clade appear to have some discrepancies compared to other clade α lineages - each of the other clades are hypothesised to contribute to both animal disease epidemics and substantial human population epidemics, with clear evidence for this in the DT104 clade (α8, red branches), and monophasic *S.* Typhimurium ST34 clade (α17). The U288 (α12) clade is hypothesised to have accumulated HDCS resulting in genome degradation, producing a reduced observation of human outbreaks and a reduced risk to food safety (78). The presence of MDR within the clade can be hypothesised as a prerequisite for bacterial survival due to past prophylactic use of antibiotics within livestock pigs, as well as their continued use for treatment of pig infections (197-200).

1.3.3.1.2 Pandemic DT104 *S.* Typhimurium (α15)

A clone with chromosomally integrated MDR phage typed as DT104 (ACSSuT; AMR genes displayed in Figure I.3), emerged in the early 1990s through cattle associated zoonoses (201). By 1992 it had become a pandemic with culture confirmed human infections in 29 countries (202). There was a clear rise and fall in isolations from 1992 to 2001 in the British Isles and Eastern Europe (202). Data from Eastern Asia shows only a dramatic decrease from 20 % of *S.* Typhimurium infection isolations exhibiting DT104 during 1996-1997, to 1-2 % from 1998 onwards, with no preceding data (202). Data from other world regions displayed an increasing trend with some plateaus from 1992 to 2001, except

Australasia which displayed no occurrences (202). There has since been a global decrease in the frequency of isolating DT104, but the lineage remains isolated from both human and animal infections (203). A study utilising WGS, phylogenetics, and Bayesian predictions from temporally structured WGS isolates estimated that a susceptible DT104 emerged in 1948 (95 % CI: 1934-1962), before acquiring the 13 Kb MDR region into an existing genomic region, now known as SGI-1, around 1972 (95 % CI: 1972-1988) (204).

SGI-1 has since spread across bacterial classes becoming a vast contributor to world-wide antimicrobial resistance (205-207). This demonstrates how crucial controlling and understanding *S.* Typhimurium is in a wider context: due to their ability to succeed - expanding globally within years and between hosts - a platform is also given for detrimental chromosomal elements such as Plasmids, ICE, IME, and phages to follow suit.

I.3.3.1.3 Pandemic monophasic DT193/DT120 *S.* Typhimurium ST34 (I:4,[5],12:i:-) (α17)

During the early 2000s an MDR (typically ASSuT; AMR genes shown in Figure I.3) clone emerged in Europe with a distinct sequence type of ST34 (208). It was initially associated with pigs and human infections which probably resulted from contamination through the pork production chain (168). The clonal lineage manifested with varying phage types, mostly DT193 and DT120, and was characteristically flagellar phase 1 monophasic, although distinct from a monophasic clone which had emerged in Spain at a similar time (208-210). A retrospective WGS analysis of UK epidemic isolates from 2005-2010 revealed that the flagellar phase 2 locus (*fljAB*) had been replaced by an MDR encoding transposon, a novel 87Kb genomic island had been acquired encoding heavy metal resistance genes, and sporadic acquisition of an uncommon virulence factor, *sopE*, had been gained several times, usually associated with a prophage designated mTmV (184). Curiously, isolates also displayed lack of the serovar specific virulence plasmid, pSLT. The genomic island was termed *Salmonella* genomic island-4 (SGI-4) and encoded genes hypothetically increasing resistance to silver, arsenic, and copper, all of which have been used as antimicrobials and growth promoters in farming and agriculture, suggesting an origin for the selective pressures driving evolution of this clone (211-217). Since initial outbreaks in Europe, the clone has now expanded globally mostly maintaining reservoirs in pigs but displaying an exceptional

host-range of domestic and wild animals, including isolations from pigs, tomato products, beef products, poultry, cattle, and deceased striped dolphins (178, 218-222). This global pandemic clone now encompasses sub-lineages which encode resistance to carbapenems, including an occurrence of IncX3 plasmid-borne *bla*$_{NDM-5}$ (223) and resistance to last-line-of-defence antibiotic *coli*stin through plasmid encoded *mcr-1, mcr-3,* & *mcr-5* (107, 224-226). This lineage has become of great clinical concern due to displaying invasive infections in Vietnam after regaining its flagellar phase 2 locus (193), and due to isolation from blood of infected children in China (227). This is contrary to the isolates studied by Bawn et al., (2020) which displayed a low invasiveness index (Figure I.3 (E)). It is likely these discrepancies are a testament to the ability of the lineage to adapt rapidly, and perhaps invasive disease aided by the addition of further AMR genes encoded on an IncH2 plasmid preventing treatment and allowing systemic infection (193). The ST34 lineage is an ever-increasing burden and other possible factors required for such successful clonal expansion remain elusive, while the clone remains the most dominant *S. enterica* globally.

I.3.3.2 Clade β

Clade β comprises mostly wild avian associated isolates, except for cattle associated MDR lineage DT204/DT49, but DT49 was also highly associated with poultry in the UK in the 1980s (Figure I.3 (A)) (228). Clade β encodes more HDCSs than clade α, with exception of DT204/DT49 clade isolates (Figure I.3 (A)). There are more occurrences of virulence gene HDCSs and a greater clade Δbitscore but a varied bitscore deviation, suggesting sustained variation throughout the clade. These factors, bar the ultimate one, are typical signatures of host adaptation. The invasiveness index of clade β is significantly greater than clade α, suggesting isolates from clade β more frequently cause disseminated disease (Figure I.3 (E)).

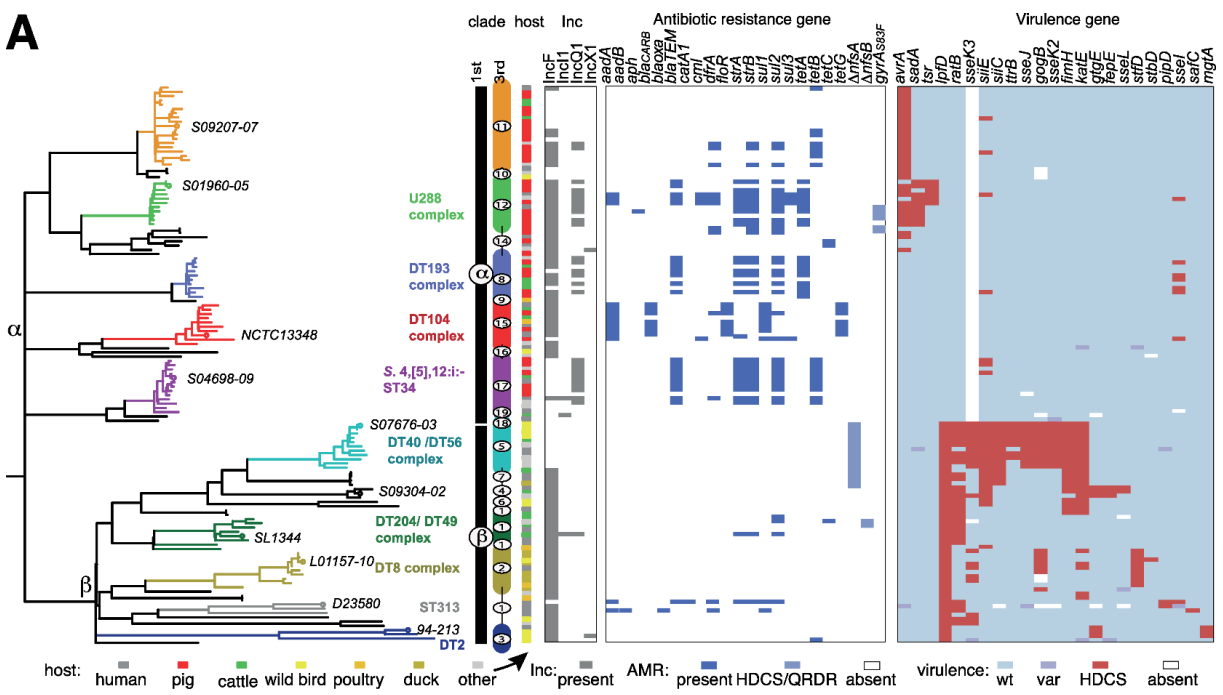

A
clade host Inc
Antibiotic resistance gene
Virulence gene
1st
3rd
IncF
IncI1
IncQ1
IncX1
aadA
aadB
aph
blaCARB
blaOXA
blaTEM
catA1
cmlA
dfrA
floR
strA
strB
sul1
sul2
sul3
tetA
tetB
tetC
tetG
ΔmfsB
gyrASer83Phe
avrA
sadA
tsr
pfiD
ratB
sseK3
silE
silC
ttrB
sseJ
gogB
sseK2
fimH
katE
gtgE
tepE
sseL
stfD
stbD
pipD
sseI
safC
mgtA
S09207-07
S01960-05
NCTC13348
S04698-09
S07676-03
S09304-02
SL1344
L01157-10
D23580
94-213
DT2
U288 complex
DT193 complex
DT104 complex
S. 4,[5],12:i:- ST34
DT40 /DT56 complex
DT204/ DT49 complex
DT8 complex
ST313
α
β
host: human pig cattle wild bird poultry duck other
Inc: present
AMR: present HDCS/QRDR absent
virulence: wt var HDCS absent
B
number of HDCS
strain
clade α
clade β
C
p<0.05
clade Δbitscore
D
bitscore deviation (% proteome)
E
p<0.05
invasiveness index
F
clade α
clade β
100 alleles
HierCC HC100 [No. of genomes]

Figure I.3 *S.* Typhimurium population structure. (A) Maximum-likelihood phylogenetic tree reconstructed using 17,832 core genome SNP sites (A, left) displaying AMR gene presence or absence (A, middle), virulence gene alleles and HDCS (A, right). (B) Quantified HDCS of 10 representative isolates from 10/19 third clade order clades. (C) Third order clade Δbitscores to show deleterious genetic information. (D) Third order clade bitscore deviations as a percentage of the proteome compared to a reference genome. (E) Third order clade's invasiveness index. Colours correspond to those used on the branches of third order clades in the phylogenetic tree (A). (F) A collection of 14,478 genomes from Enterobase representing the global diversity of *S.* Typhimurium, with the 131 isolates from (A) displayed with yellow outlines. Isolates of ST36 were also excluded from this analysis. Genomes from (F) were grouped according to their Enterobase HierCC:100 cluster definition (core E Burst Group), where neighbours must be no more than 100 core genome MLST alleles apart, displayed by different coloured circles as shown in the legend. A clear distinction of clades α and β can be observed with a large, non-UK dataset, reinforcing pertinence of analyses which dissect *S.* Typhimurium into two distinct clades. Adapted from Bawn et al., (39).

I.3.3.2.1 Pigeon associated DT2 *S.* Typhimurium (β 3)

Isolates causing disease in feral pigeons (*Columba livia*) are mainly ST128 or ST98, and mostly phage typed DT2 or DT99 (39, 77). This lineage has signatures of host restriction, exemplified by high association with pigeons. Kingsley et al., (2013) characterised genomic and transcriptomic adaptations within this lineage (77). Twenty two HDCS, which the authors termed pseudogenes, were observed within WGS of DT2 strain 94-213 compared with WGS of SL1344 of the DT204 complex, a DT104 strain, and D23580 of the ST313 clade (77). The transcriptional landscape of 94-213 also manifested with an altered response compared with SL1344 at 42^{0}C, this being the avian homeostatic temperature (77). The DT2 lineage had 3 HDCS in virulence genes identified by Bawn et al., (2020) (Figure I.3 (A) β 3), consistent with niche redundancy of the encoded proteins. The DT2 lineage also exhibits ~6 % proteome bitscore deviation and high cladal Δbitscore (~0.04), suggesting a degree of genome degradation, congruent with previous studies (Figure I.3 (D)). DT2 lineage strain 94-213 had the third highest number of HDCS (~100 genes) compared with other strains of *S.* Typhimurium (Figure I.3

(B)). There is a case report of *S.* Typhimurium DT2 causing an abdominal cavity abscess in a dog in Turkey (229). Whether pigeon adapted *S.* Typhimurium can produce severe disease in canines or DT2 is a polyphyletic phage type is undetermined, and possible oversampling from pigeon reservoirs may be skewing our understanding of the phage type.

I.3.3.2.2 Passerine bird associated DT40/DT56 *S.* Typhimurium (β 5)

Clade β 5 includes strains mostly from wild birds (passerine birds; of the order Passeriformes) as well as three cattle isolated strains and two human infection isolated strains (Figure I.3 (A)). Strains in this clade are well known to cause severe disease in passerine birds such as finches and sparrows in Australia, Asia, Europe and North America including a report from 1957 (187, 230-234). The strains are mostly phage typed DT56(var) and DT40, with no plasmid replicons, suggesting loss of pSLT. The lineage exhibits more HDCSs than any other lineage (101 HDCS, Figure I.3 (B)), consistent with a high degree of host adaptation, consistent with the observed association of the strains with seasonal outbreaks in passerine birds reported in the literature (187, 233, 235). The presence of a human isolated strain within the clade and other literature reports of human infection suggest sporadic zoonotic transmission from wild bird reservoirs to humans (187). There are also literature reports of 'spill over' to domestic cats during seasonal outbreaks, suggesting this lineage is not host restricted (235). Further adaptation is evident from accumulation of HDCS within ~10 virulence genes, albeit with slight variation amidst isolates, suggesting redundancy when causing disseminated illness while mostly undergoing transmission between passerine birds.

I.3.3.2.3 Cattle associated DT204 epidemic *S.* Typhimurium (β 1)

Between 1975-1990 strains phage typed DT204 and DT49 were dominant in the UK and Germany (169, 236, 237) (Figure I.3 (A), clade β 1). Strains from the clade had an MDR profile of resistance to ampicillin, chloramphenicol, streptomycin, sulphonamides, and trimethoprim (ACSSuT), and were assumed to originate from one clonal expansion due to similar phage type profiles (169). Strains were isolated in high numbers from cattle, suggesting this was a main reservoir, with zoonotic infections possibly from contaminated beef products (236). Strains with these phage types were discovered in other hosts, suggesting the clonal complex was broad-host-range (236). A review identified a second clone,

phage typed DT204c, that emerged from DT204 and subsequently expanded (46, 238). How these types of isolates were related genetically is unknown due to the limited methods available at the time.

I.3.3.2.4 ST313 human invasive and poultry associated *S.* Typhimurium

Isolates sequence typed ST313 emerged in Sub-Saharan Africa, causing invasive, blood-borne infection in immunocompromised people (239). This clade consists of two lineages with independent genotypes (79). The origin of the clade has been postulated as avian, as they group within avian associated clade β, and can cause invasive disease in experimentally infected chickens (91). Lineage I is not associated with invasive infection in humans, is genetically distinct, and the causative agent of gastroenteritis cases in the UK (240). Lineage II has a distinct genotype encoding two novel prophages (BTP-1 and BTP-5), a specific plasmid (pBT1), and exhibits adaptive traits such as HDCS accumulation similar to that seen in *S.* Typhi (79). The lineage has spread from Africa, with reports from South America (81). An African sub-lineage with further adaptive signatures and extensive drug resistance (XDR) has recently emerged in the Democratic Republic of the Congo (80).

I.3.3.2.5 Duck associated DT8/DT30 *S.* Typhimurium (β2)

Isolates from this lineage are associated with ducks and infrequent human disease. More considerable outbreaks occurred in the UK and Ireland between 2009-2011 (185, 241, 242). From 1992-2009 a handful of outbreaks of *S.* Typhimurium DT8/DT30 infections had been reported in England and Wales. Two substantial outbreaks of *S.* Typhimurium associated with duck egg consumption occurred in Europe; during 2010 81 confirmed cases occurred in England & Northern Ireland and 34 confirmed cases from August 2009 to February 2011 in the Republic of Ireland (241, 243). Five hospitalisations and one death were recorded as a result. An investigation revealed farms linked to the cases were positive for *S.* Typhimurium DT8 (241). These were the first outbreaks of salmonellosis due to duck egg consumption since introduction of the current UK surveillance system in 1992. The first known outbreak was due to lightly cooked pudding made with duck eggs which hospitalised >50 people (244). Despite possible host adaptation the lineage appears to retain full human pathogenicity. Isolates from both outbreaks were phage typed DT8 or DT30 having the same multi-locus sequence type (ST19). They were phylogenetically closely related with 106 core-genome SNPs separating the most distantly related

DT8 sequences in a retrospective WGS analysis (242). Both phage types have also been isolated from the same duck egg (245). The difference between the two types is gain of resistance to 10 typing phage preparations in DT30. A possible cause for a switch from DT8 to DT30 from WGS has been hypothesised (246). This hypothesis was drawn from differences in WGS data between 3 DT8 isolates and 1 DT30 isolate and featured acquisition of an ICE with loss of *S.* Typhimurium virulence plasmid pSLT in DT30. This equated to a different array of restriction modification (RM) systems, with DT30 missing the type II restriction enzyme M.EcoGIX, and other unspecified altered phage interactions hypothesised from the DT30 ICE genetic material.

I.3.4 Successive clonal replacement of dominant MDR broad-host-range *S.* Typhimurium

The most frequently isolated clones of *S.* Typhimurium appear to emerge and replace each other every ~15 years (183). Each of these clones are successful, broad-host-range, widespread clonally expanded lineages that encode MDR (Figure I.4). The reasons for these clonal replacements are not understood. It was noticed that replacements occurred between clones from cattle populations in Germany and the UK, but the most recent clonal replacement is highly associated with pigs, suggesting different origins for different successful clones (238, 247-249). Surveillance through phage typing began in the 1950s and toward the 1960s it became apparent that increasing proportions of *S.* Typhimurium phage typed DT9 were isolated, with subsequent replacement of this clone exhibiting mostly phage types DT204/DT49 (Figure I.2(A) β 1). This was dominant until spread of a DT104 clone characterised by acquisition of ~13kb multi-drug resistance (MDR) cassette within *Salmonella* genomic island 1 (SGI-1) (FigureI.3 (A) α15) (250, 251). The most recent emergence is clonal expansion of MDR monophasic *S.* Typhimurium ST34 (Figure I.3 (A) α17), possibly due to factors such as SGI-4 (252, 253).

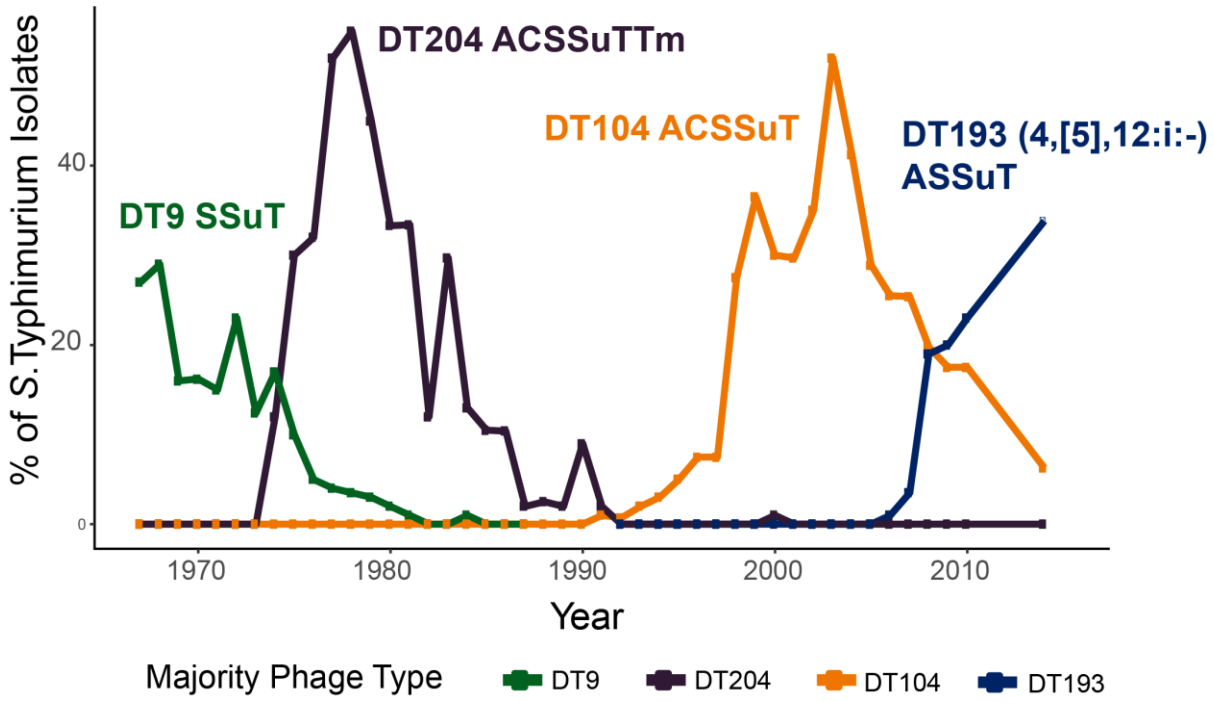


Figure I.4. Successive clonal expansion and replacement of *S.* Typhimurium clones 1965-2018. Successive clonal replacements of *S.* Typhimurium lineages expressed as a percentage of clones isolated over time. Each line is labelled with the phage type observed most frequently within each clonal complex and common MDR profiles. A: ampicillin, C: chloramphenicol, S: streptomycin, Su: sulphonamides, T: tetracycline, Tm: trimethoprim.

I.4 Mechanisms of microevolution

Lineages of *S.* Typhimurium, excluding ST36, exist within ~1500 core genome SNPs (39). Variation in closely related bacteria is due to microevolution (184). There are 2 main types of microevolution: SNPs and recombination. Recombination includes within genome rearrangements, as well as integration of new horizontally transferred material, usually through naked DNA uptake (transformation), phage mediated gene acquisition (transduction), and T4SS mediated uptake of genetic material (conjugation). These mechanisms exhibit crosstalk; interacting, merging, and reshaping each other, generating genomic plasticity which underlies the variation in closely related bacteria.

I.4.1 Single nucleotide polymorphisms

Transfer of polymorphisms to daughter cells is termed vertical gene transfer and the acquisition of mutations is gradual, generating comparatively slow evolution due to successive replication cycles

required for sufficient accumulation of SNPs to counter selective pressures. SNPs can be due to a phenomenon known as genetic drift or generated from selection pressure to modify a gene with individual SNPs, such as SNPs in phage receptor encoding genes which may alter the bacterial host susceptibility range (254). Genetic material can be under negative selection resulting in gene degradation due to redundancy, commonly observed with host adapted serovars and lineages of *S.* Typhimurium (43). When under negative selection through pressure against conservation of a gene it is common to see multiple disrupting SNPs and introducing of stop codons inhibiting deleterious genetic material – a HDCS - previously termed pseudogenes.

I.4.2 Horizontal gene transfer

The most rapid form of bacterial evolution is through horizontal gene transfer (HGT) (255). HGT can be mediated through phage transduction, DNA uptake through transformation, and acquisition of plasmids and ICE through conjugation. Transfer of this type allows exchange of large amounts of genetic material that can happen between a wide range of organisms, including inter-kingdom exchange (256, 257). Acquired advantageous elements can include genes for virulence, antimicrobial resistance, heavy metal resistance, and biosynthetic pathways, enabling bacteria to rapidly gain complex functions (258). This form of gene flow contributes extensively to evolution in *S.* Typhimurium: *Salmonella*e's distinguishing pathogenicity islands were horizontally transferred, and acquisition of new elements can produce strains which take over as dominant clones (18, 259-262). HGT aided the development of SGI-1 in DT104 which enabled it to become dominant in the UK in the 1990s; likewise SGI-4 from the current DT193/DT120 monophasic epidemic was hypothetically horizontally transferred and seems to have initiated its clonal expansion (103, 263). Phages also contribute extensively to HGT mediated genome variation in *S.* Typhimurium (39).

Plasmids and ICE that self-transfer within Gram-negative bacteria usually encode a T4SS with large protein pilus and contribute to gain of complex functions in one evolutionary event (264, 265).

1.4.2.1 Plasmids

Plasmids are circular, extrachromosomal DNA elements that can horizontally transfer between bacterial phyla (266). They can self-transfer, or require helper plasmids or chromosomal genes for transfer, typically borrowing a T4SS from other elements (267). They are frequently associated with antimicrobial-resistance genes and have spread antibiotic resistance considerably among pathogen populations (268-271). Plasmids encode specific DNA sequences used to initiate their replication, called replicons. If a plasmid enters a cell that encodes the same replicon as a native plasmid then the plasmids compete for replication machinery regulated through antisense RNAs and iterons that produce variable plasmid copy numbers. This produces a phenomenon where only one plasmid will transfer to daughter cells, known as plasmid incompatibility (271, 272). Strong evolutionary selection pressures will maintain plasmids which convey a significant selective advantage for the host. *S.* Typhimurium has a ~90kb serovar specific virulence F-type plasmid, termed pSLT, which was probably acquired during evolution of subspecies I, with subsequent recombination of the virulence operon, *spv,* into the chromosomes of subspecies II, IIIa, IV, and VII (273). More recently plasmids encoding MDR are rife within *S.* Typhimurium lineages, consistent with strong pressure from antibiotics used for disease treatment and previous prophylaxis in agricultural animals (107, 274-279). Notable examples include: pSTM6-275 in monophasic *S.* Typhimurium ST34 which encodes MDR and heavy metal resistance (276); emergence of various *coli*stin resistance encoding plasmids carrying *mcr-1, mcr-3*, and *mcr-5*, associated with incompatibility groups IncHI2, IncI2, and IncX4 respectively, also in monophasic *S.* Typhimurium ST34 (225), and; an IncHI2 plasmid, pSTm-ST313-II.1 associated with extensive drug resistant *S.* Typhimurium ST313 lineage II.1 (80).

1.4.2.2 Integrative elements

DNA elements which transpose themselves within the same genome and among other cells can vary in size. The smallest insertion elements usually encode one or two genes, including a transposition gene for excision and integration to separate genomic loci (280). These can be replicative, undergoing replication when excising, leaving behind a copy within the locus, or conservative, where the original sequence is conserved after excision (281, 282). The largest integrative elements can be >600kb and may encode a range of beneficial material for the host, but usually rely on host replication machinery

(283). Exceptions to this rule include Actinobacteria ICEs which self-replicate (284). ICE are overwhelming contributors to global antibiotic resistance evolution and dissemination, with some AMR genes being exclusively ICE associated (285). A recent study suggested that nearly half of all bacterial genomes contain an ICE, and a further study suggested that diversity in T4SS is far greater in ICE than in plasmids, contrary to previous beliefs (283, 286). If an integrative element does not encode machinery for self-transfer but self-integrate, these are usually designated integrative mobilizable elements (IMEs). IMEs require a helper T4SS from an ICE or plasmid, which may co-transfer with the helper element (287). Also, mainly observed within *Streptococcus agalactiae,* are elements which have lost both their integration and conjugative modules, but retain sequences used for recognition during transfer (attP and attB) (288). These elements are termed cis-integrating mobilizable elements (CIMES).

Integration into chromosomal locations by phage, IMEs CIMES, and ICE are largely through tyrosine recombinases (289). These recombine the integrating DNA recognising a specific attachment site that appears at the combined distal ends of the integrating DNA fragment (attP), the host chromosomal site having near identical sequence (attB) and integrating via crossover holiday junctions (290).

Non-replicative ICE life cycles have common characteristics (291). Usually ICE undergo vertical transfer to daughter cells and without stimulation limited quantities of transfer occurs in a subset of the population (292). Stimulators of excision are commonly DNA and protein stressors, where the selfish nature of the ICE to maintain itself manifests, transferring out of the stressed donor cell and into a recipient (Figure I.4). ICE transfer is initiated through self-encoded excision machinery, such as XerC/XerD recombinases (293), or phage excision-like machinery (294), with binding at attachment sites that generate a circularised, extrachromosomal ICE (Figure I.5) (295, 296). A relaxase protein, orthologous to *traI* from F-plasmids, then nicks one strand at a DNA site called the origin of transfer (oriT). Single stranded binding proteins protect the ssDNA cargo, and the DNA-protein complex is localised to a type-4 coupling protein and secreted through a T4SS pilus into a recipient cell. Integration then occurs, provided the recipient encodes a complementary attachment site (attB). Attachment sites are frequently encoded in slow-evolving DNA sites, such as tRNA-encoding regions and tRNA biosynthesis genes (297). ICEs also require maintenance, and frequently encode mechanisms to ensure

that daughter cells of the host retain the ICE, such as toxin-antitoxin systems (TA) (298). This creates extra layers of diversity - TA systems are well known to have multiple functions (299).

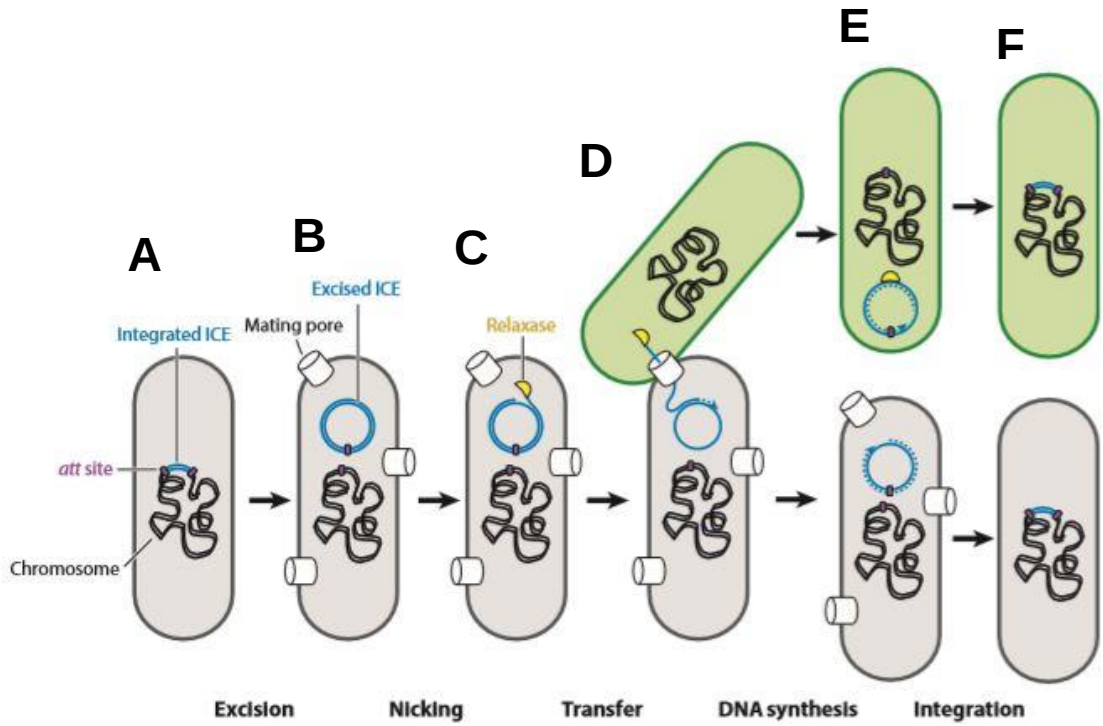


Figure I.5 ICE life cycle. (A) Transfer genes of an ICE are typically under repression or inactive, keeping the element maintained in the chromosome. (B) One of many stressors or signals, such as DNA intercalating molecules, may stimulate excision. Excision stimulation is dependent on the ICE regulatory system, which will stimulate expression and construction of a mating pore consisting of a T4SS pilus and coupling proteins under various conditions. Excision is undertaken through DNA binding at DNA attachment sites via crossover junction. The chromosomal attachment site is known as attB, and ICE attachment site attP. Excision generates a fully extrachromosomal circular ICE DNA molecule. (C) A relaxase protein nicks and unwinds the ICE DNA, covalently bonding to the 5' end, producing transfer DNA (T-DNA) and partitioning a single strand in a manner akin to rolling circle plasmid replication. (D) The T-DNA localises to the mating pore and is transferred into a donor cell. (E) The relaxase forms a covalently closed, circular ssDNA, and its complementary strand is synthesised. The resulting dsDNA is the substrate for chromosomal integration, and the ICE DNA integrates into donor and into recipient

chromosomes, making this process non-conservative transposition. The donor cell is shown (grey), recipient (green), ICE DNA (blue), and relaxase protein (yellow). From Johnsen e al., (2015)(291)

Notable integrative elements in *Salmonellae* include all SPIs and SGIs (18, 263, 265, 300). When an integrated island has considerable host benefit and is under strong positive selection the genes encoding transfer of the islands become depleted, creating a fixed genomic feature (287). This is likely the case for SPI-1 and SPI-2, which appear to have been horizontally transferred but lost self-transferability and become part of the core genome (18, 22). Other notable integrative elements in *Salmonellae* include: SPI-7, a putative ICE that encodes the Vi antigen of typhoidal serovars such as Typhi; SGI-1, an IME that encodes MDR in pandemic DT104 *S.* Typhimurium, and; SGI-4 of monophasic *S.* Typhimurium ST34, a hypothetical MGE that encodes heavy metal resistance genes (250, 265, 301).

I.4.3 Introduction to bacteriophage

Bacteriophage (phage) are viruses which infect and survive within bacterial populations. Phage have been co-evolving with bacteria since early forms of life and out-number bacteria by an estimated tenfold (302). Phage are the most abundant biological entity in the biosphere, and consequently play an important role in determining numbers of marine bacteria, greatly affecting aspects of our ecosystem such as nutrient cycles (303). The diversity of phages can be observed through the frequency of novel genes found in their genomes (304). Phage have a characteristic two-stage lifestyle, which involves lysogenic integration into the bacterial genome with vertical transfer to daughter cells, and an active lytic phase where they lyse bacterial cells and horizontally transfer to a new host (Figure I.6).

I.4.4 Bacteriophage transduction and bacterial genome plasticity

Phage play a fundamental role in bacterial evolution through horizontal gene transfer, phage mediated genome plasticity, recombination with the genome, and producing strong selection pressures for development of bacterial anti-phage mechanisms (305). *Salmonella* infecting phage have been shown to transduce various important genes (Table I.3).

Protein type	Function	Gene	Phage	Reference
O-antigen modification	Glycosylation	*rbf*	ε 34	Wright, 1971 (306)

O-antigen modification	Length determination	*GtrC*	BTP-1	Kintz et al., 2015 (307)
Effector protein	T3SS effector	*SopE*	SopEΦ	Mirold et al., 1999 (308)
Effector protein	T3SS effector	*GogB*	Gifsy-1	Coombes et al., 2005 (309)
Effector protein	T3SS effector	*SseI*	Gifsy-2	Figueroa-Bossi & Bossi, 1999 (310)
Effector protein	T3SS effector	*SspH1*	Gifsy-3	Figueroa-Bossi et al., 2001 (311)
Intracellular survival enzyme	Superoxide dismutase	*SodC-1*	Gifsy-2	Figueroa-Bossi & Bossi, 1999(310)
Intracellular survival enzyme	Superoxide dismutase	*SodC-3*	Fels-1	Figueroa-Bossi et al., 2001(311)
Intracellular survival Enzyme	Superoxide dismutase	*NanH*	Fels-1	Figueroa-Bossi et al., 2001(311)
Other	Antivirulence gene	*GrvA*	Gifsy-2, Fels-2	Ho & Slauch, 2001(312)

Table I.3. Genes horizontally transferred into *S.* Typhimurium through phage transduction.

Most characterized *Salmonella* phages are from the tailed, dsDNA order *Caudovirales,* which contains the viral families Podoviridae, Myoviridae, and Siphoviridae. Each family has corresponding tail morphology: short, medium, and long respectively. Due to recombination being a fundamental part of the phage life cycle, they often cross-recombine sections of their genomes, making them hypervariable and difficult to class in a traditional taxonomic manner.

I.4.5 Bacteriophage life cycle

The *Caudovirales* infection model is largely the same across the order (313) (Figure I.7). When encountering a host bacterium the tail section binds to a cell surface receptor and uses an injectisome-like mechanism to deliver dsDNA (314). The tail punches a hole through the bacterial plasma membranes and cell wall, and the genome passes down the tail into the cell. Genes are expressed from

transcripts produced by host machinery and ribosomes. *Caudovirales* genomes are generally replicated via use of concatemers which are put together to form the mature genome (315, 316). Capsid proteins come together to form a prohead, which holds the genome (317). The prohead is then matured by capsid subunit cleavage forming an icosahedral head with 5-fold symmetry. The tail consists of helix-based proteins with 6-fold symmetry and is either constructed separately and joined to the connecter or constructed directly onto the head. After maturation the cell is enzymatically lysed by endolysins and holins, releasing mature virions into the external milieu.

Phage λ has been well characterized and has a genetic structure is similar to P22 (Figure I.6, below)(318). In short, N and cro are immediately translated, which allows recombination and integration, as well as transcription of cII and cIII. CII is constantly degraded by cellular proteases, such as those in the RecA-dependant DNA damage response (319). If levels of cII are low, transcription from P_L and P_R continue and virion replication ensues, with resulting cell lysis. If cII concentration increases it binds O_L and O_R, inhibiting expression of replication of genes, enabling action of integration and recombination proteins, producing lysogeny. This is mostly true for P22, except P22 encodes two repressor proteins (318).

Several *Salmonella* phages have been characterized (table I.4 below). A widely used and studied phage in *Salmonella* research is λ-like P22, which has its own suggested family of P22-like phages (table I.4 below). P22 is of the short-tailed family *Podoviridae*. It is commonly used for transduction in *Salmonella* mutagenesis and therefore well characterized (320-326). Some *S.* Enteritidis typing phages have been characterized (327). There is little information about *S.* Typhimurium typing phages, although they are likely related to P22 (327).

Phage	Family	Features	References
SfV	μ-like	Originally discovered *Shigella flexneri*	(328)
ST64B	μ-like	A mosaic phage, *S.* Typhimurium DT64 lysogen	(329)
P27	μ-like	Shiga-toxin encoding, lysogenic antigen conversion	(330) (331)

μ	μ -like	Well characterised phage, ancestral protein fold enables immune evasion	(332-334)
SopEΦ	P2-like	SopE virulence factor encoding, common *S.* Typhimurium lysogen	(335) (336)
PSP3	P2-like	*Salmonella* P-2 like phage, frequent lysogen	(337) (336) (338)
Fels-2	P2-like	Common *S.* Typhimurium lysogen	(13, 339, 340) (336)
BTP5	P2-like	ST313 specific, P2-like structual genes but possible novel phage biology	(190) (341)
P2	P2-like	Common *E. coli* prophage, prototypical lysogen, no recA cleavage, R-type pyocin like encoding gene, helps 'pirate' phage P4 through sharing functional genes allowing P4 propagation	(342-344) (345)
Gifsy-1	λ -like	Common *S.* Typhimurium lysogen	(13, 339, 340)
Gifsy-2	λ -like	Common *S.* Typhimurium lysogen	(13, 339, 340)
λ	λ -like	*E. coli* phage well characterised lysis, lysogeny, & regulation	(346)
Fels-1	λ -like	Common *S.* Typhimurium lysogen	(13, 339, 340)
ST104	P22-like	MDR pandemic *S.* Typhimurium DT104 specific lysogen	(347)
P22	P22-like	Lysogenic antigen conversion, commonly used for molecular research in *E. coli & S. enterica*	(348) (349) (350)

ST64T	P22-like	Lysogenic phage type conversion, *S.* Typhimurium DT64 lysogen	(351)
ε 34	P22-like	Serovar Anatum specific	(352)
BTP1	P22-like	ST313 specific, O-chain length modifying, high spontaneous lysis induction	(190) (307) (341)
ES18	P22-like	Prototypical generalised transducing phage	(353)
ε 15	P22-like	Serovar converting phage	(354)
T7	T7-like	Well characterised lytic phage, encodes inhibitor of BReX defence	(355) (356)
SP6	T7-like	Dual host specificity through tailspike rotation	(357) (358)
Felix 01	Felix 01-like	Broad-host-range effective lysis	(173, 359)

Table I.4 Features of characterised *Salmonella* phages and their grouping phages. Adapted from Kropinski et el., (2007)(360)

I.4.6 Bacterial Phage Resistance Mechanisms

Bacteria have evolved intricate methods of bacteriophage defence through constant evolutionary pressure of phage-killing resulting in billions of years of antagonistic co-evolution (361-364). Phage evolve counter mechanisms which pressures evolution of further resistance mechanisms in bacteria, resulting in the world's most ancient conflict – a biological evolutionary arms race between bacteria and phage (365). Currently described methods of *S.* Typhimurium phage defence comprise both innate and adaptive systems (366-368). Innate systems include restriction modification (366), super infection exclusion (369), abortive infection (370), and BrEX (371). The adaptive system is a type 1E, clustered, regularly interspaced, short palindromic repeat (CRISPR) system with CRISPR associated (*cas*) genes (368). A common and effective innate defence which is highly indicative of phage host-range includes phase variation or single nucleotide polymorphisms in genes encoding cell surface structures (372). Phage require receptors which are constitutively present on the bacterial cell surface for successful

attachment and entry to continue the lytic cycle or undergo lysogeny. They therefore frequently utilise molecules which are important for the host. A common example is use of lipopolysaccharide (LPS) (373).

I.4.6.1 Receptor polymorphisms and the *S.* Typhimurium lipopolysaccharide O-antigen

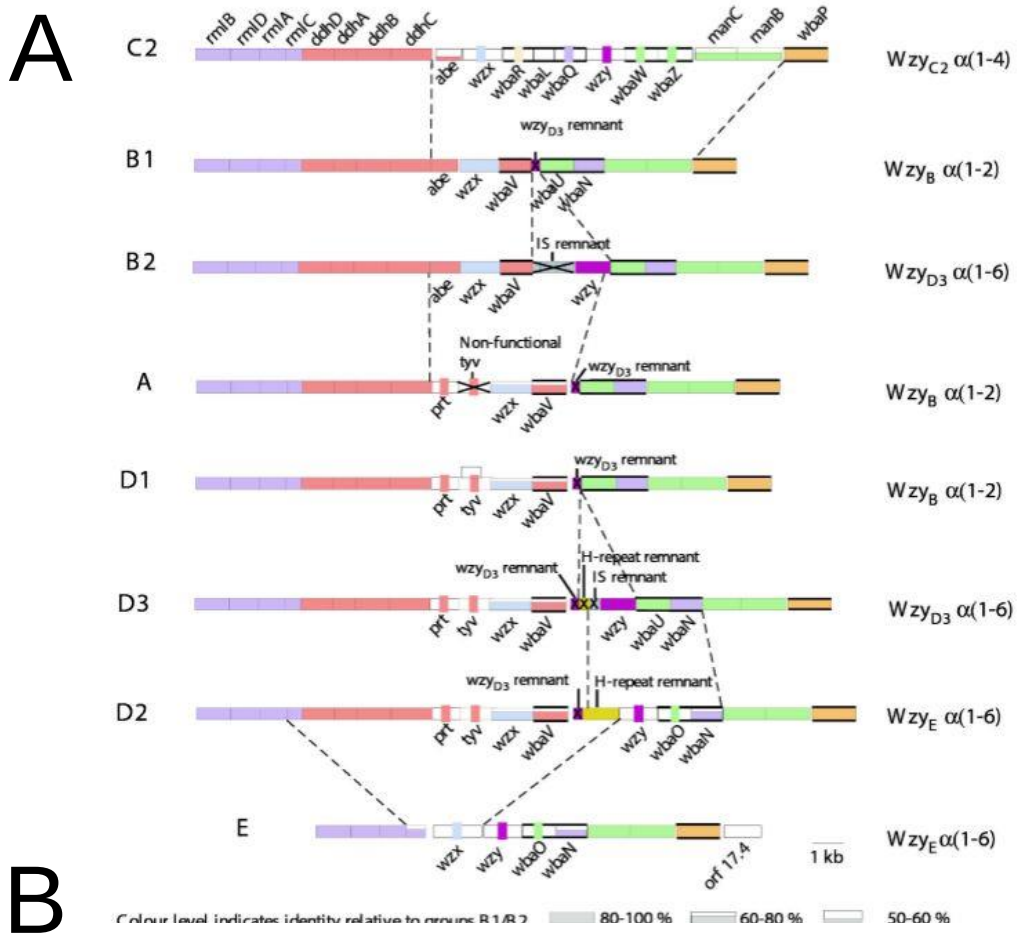
Receptor alterations can inhibit phage adsorption, preventing phage DNA entry. To limit phage propagation and deny phage receptor binding, bacteria adapt the structure of surface molecules, such as outer membrane proteins or the flagella, or change their 3-dimensional orientation (374). This can be through coding sequence mutations, phase variation, or receptor altering genes encoded on transferred MGEs.

A commonly used receptor for *Salmonella* phage, such as p22, is the lipopolysaccharide O-antigen (375). It consists of a membrane bound phospholipid called lipid A, sugar-phosphate inner core, outer core polysaccharide, and repeating oligosaccharide O-antigen unit which usually consist of 3-6 sugar residues (376). *S.* Typhimurium (I:4,[5],12:i:1,2) phenotypically exhibits antigens O4, O5, and O12. One O-unit core contains:

[→(2)mannose(α1→4)rhamnose(α1→3)galactose(1) →]

including an (α1→3) abeqouse side chain linked to the mannose residue, where numbers in brackets display the carbon position of linked bonds. The abeqouse residue is variably O-acetylated, producing a variable O5 factor, indicated by brackets, which is dependent on gene *oafA*, The gene has a frequent deletion of 7bp sequence repetition resulting in a frameshift and O5- phenotype (377, 378). Polymerisation of repeat O-units in *Salmonellae* is predominantly through the inner membrane bound Wzx/Wzy-dependant pathway (379). Wzx is the O-unit flippase, and Wzy the polymerase. *S.* Typhimurium Wzy is predicted to have 11 transmembrane domains, and produce an $\alpha_{1\text{-}2}$ O-antigen repeat linkage present in *Salmonella enterica* serogroup B (referred to as Wzy_B) , which produces ~16-30 repeats of the O-unit tetramer, a long multipurpose surface molecule commonly as a receptor for phage. From sequence analysis there appears to be a complex history of *wzy* genes within only a few serogroups (Figure I.7) (380). This includes deletion of the ancestral *wzy* from the O-unit biosynthetic

gene cluster (Figure I.7 (A)), subsequent insertion of a different *wzy* associated with an IS-617 transposase and remnants of various IS elements having re-introduced the proposed ancestral gene in some serogroups (Figure I.7 (A) C1). *S.* Typhimurium encodes a *wzy* gene not associated with the O-antigen biosynthetic gene cluster, consistent with it having been inserted through HGT.



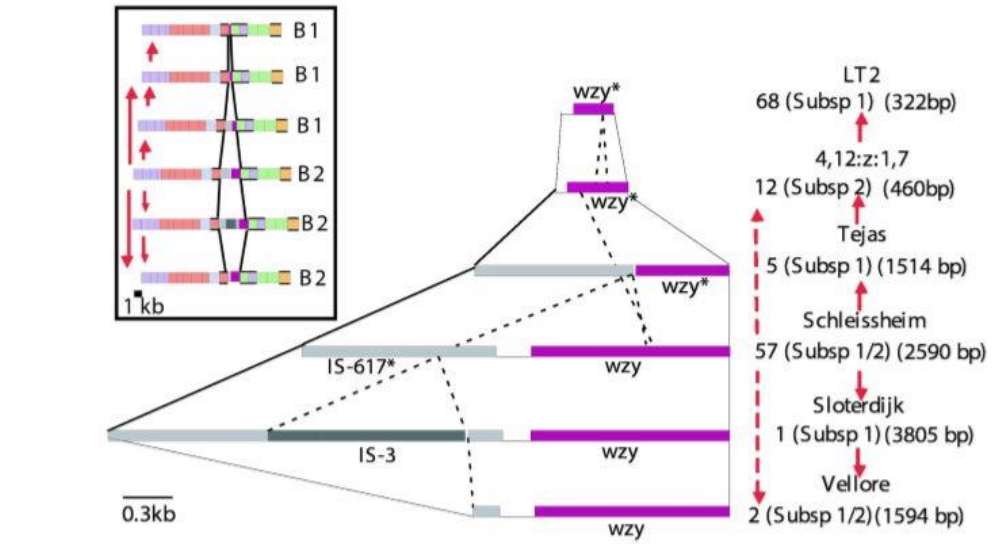


Figure I.6 O-antigen gene cluster and *wzy* diversity of *Salmonella* serogroups . (A) Diversity in O-antigen gene clusters of galactose-initiated O-antigens of 8 serogroups. Purple denotes genes involved in biosynthesis of rhamnose, and red genes involved in biosynthesis of abeqouse (with *abe* gene present) or paretose and tylevose, dependant on presence of *prt* and *tyv* respectively. *Wzy* genes are shown in purple, and those which contain sequence identity to the proposed ancestral serogroup B *wzy* shown in block purple, but *wzy* genes without significant sequence similarity shown with a purple stripe and white box. (B) Possible evolution of the serogroup B ancestral *wzy* locus from Reeves et al., (2013)(380).

O-antigen length can be altered through various means. BTP-1 prophage of ST313 strain D23580 encodes GtrC which is a glycosyltransferase that alters O-antigen length, protecting against further BTP-1 infection (381). Phages Gifsy-1 and Gifsy-2 use OmpC as a receptor and mutations in its coding region inhibit their adsorption. Introduction of the receptor to *E.coli* conferred ability to adsorb Gifsy phage (382). Natural variations in nucleotide content can inhibit adsorption, such as frameshift mutations in LPS polysaccharide biosynthesis genes, which can alter the length or content of LPS chains, possibly removing a motif that was previously a phage receptor. The O-antigen is subject to phase variation through DNA-methylation of OpvA and OpvB (383). This causes a trade-off between virulence and bacteriophage resistance: the 'long' phase of the O-antigen determines macrophage differentiation but also allows phage entry, and the 'short' phase attenuates *Salmonella* but protects against phage adsorption (375, 384). This trade-off occurs in many forms due to the important functions of LPS; LPS is an outer-membrane anchored polysaccharide found in most gram-negative bacteria and a crucial component of gram-negative bacterial pathogens (385). The main functions of LPS include membrane protection from chemical attack such as bile acids, contributing to the structural integrity of the bacteria, increasing the negative charge of the cell for stable membrane structure, biofilm formation and adhesion (386). LPS has an instrumental role during human and animal infection through innate immune

activation as a potent endotoxin which causes toxic shock like symptoms. It is a toll-like receptor 4 activator, as well as a ligand for intracellular inflammatory caspases (387).

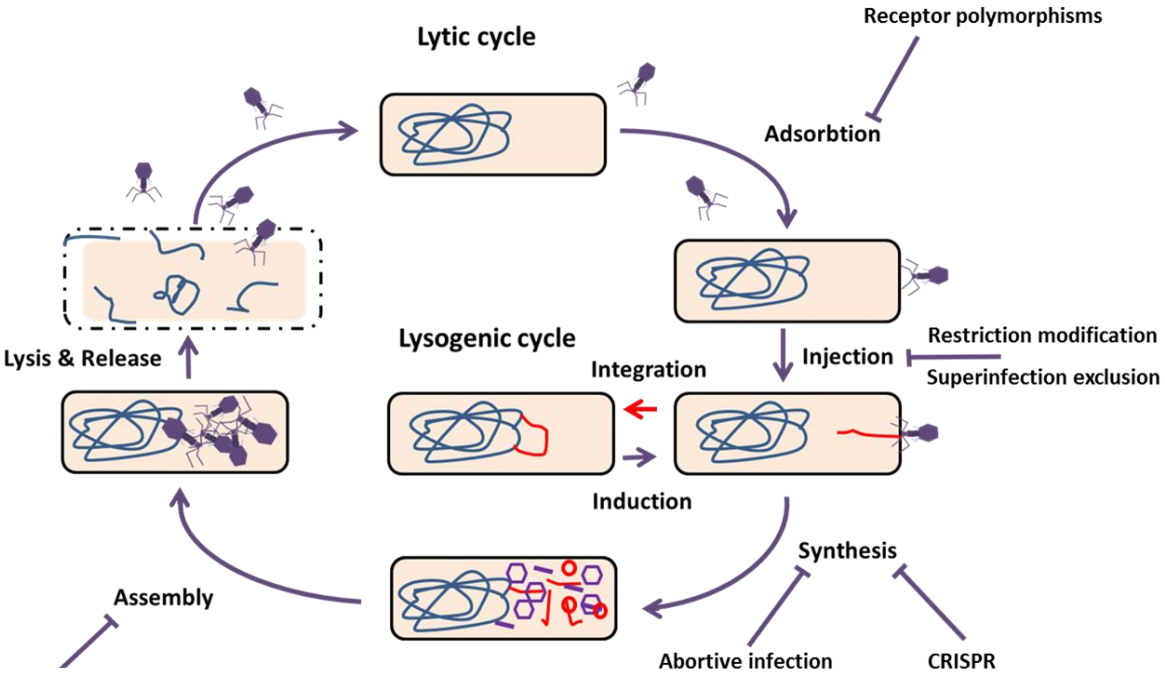


Figure I.7 Bacteriophage life cycle showing lytic and lysogenic phases and stages affected by bacterial phage resistance mechanisms. Phage initiate infection through adsorption to cell surface molecules (receptors), which can be inhibited through receptor polymorphisms such as SNP based mutations in receptor processing proteins, or length alteration from super infecting phage proteins. Phage DNA is injected via injectosome-like mechanism, which can be cleaved by inner membrane bound superinfection exclusion nucleases, or restriction endonucleases, albeit dependant on the modification of the phage genome. An adaptive mechanism encoded on the CRISPR-cas locus involves formation of a 'spacer', which is 24-47 bp homologous phage DNA sequence integrated into the host genome. Its transcript used as guide for endonucleases to cleave synthesised phage genomes. Abortive infection involves recognition of phage infection by the host and altruistic self-inhibited metabolic activity to prevent further action from the infecting phage.

I.4.6.2 Abortive infection

Abortive infection systems (Abi) provide resistance through the abortion of phage infection leading to the altruistic death of the infected cell (388). This is typically done by targeting a step of phage multiplication such as replication, transcription, or translation. These systems have been studied for over 50 years, but the modes of action are not completely understood due to their complexity and knowledge gaps in phage biology. A well characterized system of this type is the Rex system found in phage λ-lysogenic *E.coli* (388). A widespread TA system has been discovered which acts as an abortive infection system, and similarly there is evidence that TA systems can both inhibit and trigger Abi systems, suggesting complicated interplay between newly resided TA encoding prophages and the existing chromosomal TA systems (389, 390). There are no well characterized Abi systems in *Salmonella* to date.

I.4.6.3 Restriction modification

Restriction modification systems (RMs) are predicted to be present within all bacterial phyla, and include 4 types (type I-IV), dependant on subunit composition, sequence recognition, cleavage position, cofactor requirements and substrate specificity (391). These systems cleave phage DNA after entry by recognition of certain short sequences, and determine self and non-self-DNA through DNA modification (392). Usually modification is undertaken on nucleotide bases such as N6-methyl-adenine, N4-methyl-cytosine, C5-methyl-cytosine and 7-deazaguanine (393, 394). Due to variable DNA modification, RM systems are implicated in epigenetic gene expression in bacteria (395). The extent of DNA modification for anti-phage defence was shown to extend past base modification and include insertion of sulphur atoms substituting non-bridging oxygen atoms in the DNA backbone, producing phosphorothioate moieties in stereospecific and sequence specific residues (396). These modifications also have self/non self-sensing cognate restriction protein complexes for cleaving phosphorothioate-lacking foreign DNA (396, 397).

A 2016 study characterized the RMs of 221 *S. enterica* genomes, within 97 different serovars (398). They identified 113 putative RMs including 58 type I, 23 type II, two type 3 and 30 type IV. All genomes except one contained two to seven RMs, with one strain containing just one. All strain WGSs contained a type III RM, one of which was shared by 198/221 genomes.

I.4.6.4 Clustered, regularly interspaced, short palindromic repeats and associated genes

CRISPR systems are adaptive immune systems which convey immunity to phage and plasmids (399). They are characterized by three major features: a set of CRISPR associated (cas) genes, a leader sequence, and a CRISPR array. They are present in ~45 % of bacterial genomes (400), including *Salmonellae* (401). The CRISPR array contains 24-47bp direct repeat sequences separated by 21-72bp sequences termed 'spacers' which are typically derived from plasmids and phage (402, 403). There are six main types of CRISPR-cas system, each with different cas proteins and variation in processes (404). Due to the age of these molecular systems they are prototypical examples of the evolutionary arms race which results from strong selection pressures and counter selection, playing a role in determining bacterial pathogen evolution and population structures (405-407). Phage encoded anti-CRISPR mechanisms are abundant, for example including prevention of cleavage of invading phage DNA through protection by a proteinaceous 'nucleus' encapsulating the genomic DNA of some jumbo phages (408-410).

Salmonellae have two CRISPR loci, CRISPR1 and CRISPR2 and eight *cas* genes located upstream of CRISPR1 (411). CRISPR1 has the type one signature gene involved in DNA cleavage *cas3*, as well as universal *cas1* and *cas2* involved in spacer acquisition, and genes typical of type 1-E systems, *cse1*, *cse2*, *cas7*, *cas5* and *cas6*. Shariat et al. (2015)(368) reported high nucleotide identity across the cas operon among isolates of a single serovar and in some cases extending across serovars. This was consistent with the locus having ceased adaptive events and no longer being immunogenic. However, deleterious CRISPR loci should undergo degradation over time, suggesting another role for CRISPR loci in *Salmonella*. Accordingly, CRISPR loci and genes have been shown to possess other diverse functions (412).

I.4.6.5 Bacteriophage Exclusion

Bacteriophage exclusion (BrEX) refers to a cluster of genes associated with a methyltransferase encoding gene(*pglX*) that were identified in ~10 % of all sequenced microbial genomes and appear divided into six subtypes (413). The original description was of a cassette in *Bacillus subtilis* containing a putative Lon-like protease (*brxL*), an alkaline phosphatase domain protein (*pglZ*), an RNA-binding domain protein (*brxA*), an ATPase domain protein (*brxC*), a DNA methylase (*pglX*) and a protein of

unknown function (*brxB*). The cassette conferred significant immunity to bacteriophage when compared with loss-of-function mutants (414). Loss-of-function mutants showed no phage DNA degradation suggesting a novel mechanism of inhibiting phage replication. It has been elucidated that these genes exist in NTS (414).

I.5 Antagonistic co-evolution between phage and pathogenic bacteria

Integrative elements play a role in establishing complex traits in one evolutionary event, enabling rapid niche adaptation. However, phage and pathogen population structures are hypothesised to be constantly influencing each other due to the abundance of environmental phage, and their predation of bacteria (415). The significance of this effect, and which phage-defence mechanisms are employed by pathogens such as *S.* Typhimurium to counter phage predation, is not well understood. This question may be pertinent in *S.* Typhimurium due to high sequence identity between isolates, but significant host variation in lineages. A question raised by the diverse host-ranges exhibited by *S.* Typhimurium is: how do host adapted *S.* Typhimurium resist predation from new phages introduced to their niches? and; how do successful, broad-host-range *S.* Typhimurium evolve to resist phage predation from a probable wide range of ecological niches? The idea that understanding phage-pathogen interactions aids in understanding pathogen evolution, and therefore may be exploited for therapies, has been discussed in review for *Pseudomonas spp* (416). An important point raised in this review is that phages have been discovered which affect nearly all key cellular processes; transcription, translation, motility, cell division, RNA degradation, metabolism, CRISPR-based adaptive immunity, cytoskeletal functions, DNA replication, as well as conflicting effects on pathogenicity (416).

Due to reduced cost sequencing technologies and WGS we have an increased arsenal with which to discover and understand these interactions (417, 418). A recent study utilised WGS to investigate this effect from an evolutionary perspective in shiga-toxigenic *E.coli* (STEC O157:H7)(419). The study concluded that sub-lineages of this pathogen were stable in phage resistance profiles, and that occasional deviations from the standard profile do not outcompete the stable lineage, unless conveying a selective advantage. There was one example of increased phage resistance, associated with a mutation in *ompC* which encodes outer membrane protein OmpC - a receptor for T7-like phages – and accordingly the

mutation provided resistance to these phages. The toxin gene-encoding phage was associated with resistance to T4 group phages and hypothesised to allow improved persistence in the ruminant gut, a niche they are supposedly abundant in. An important question raised by small subpopulations of varying resistance is: why and how this is occurring? If these mutations are deleterious, why are the clones not rapidly outcompeted prior to infection, allowed to progress through the food production chain, cause patient infection, and subsequently be isolated? In the case of phage resistance within lineages; is this associated with divergence of two equally isolated clones, or does one clone exhibit increased sampling and positive selection? Importantly for *S.* Typhimurium, how do lineage phage resistance profiles stabilise within a host, or exclude phages from a broad range of hosts?

I.6 Hypothesis and aims of the study

This study was driven by the hypothesis that horizontal gene transfer, including bacteriophage predation, contribute to the genome evolution and therefore survival of *S.* Typhimurium, and that different lineages with varying host ranges and ecological niches will have different evolutionary requriements and phage sensitivities. The questions this study aimed to address were: i) what genomic features contribute to emergence and continued expansion of *S.* Typhimurium clones, such as the emergent, successful, broad-host-range monophasic *S.* Typhimurium ST34 clone?; ii) how does SGI-4 contribute to emergence of monophasic *S.* Typhimurium ST34, and what are its characteristics?; iii) does bacteriophage predation shape the genome evolution of *S.* Typhimurium? Do *S.* Typhimurium from different niches have different requriements, genomic elements, evolution, and phage sensitivities? and; iv) which mechanisms are employed by *S.* Typhimurium lineages to resist phage predation?

Chapter II: Materials and Methods

II.1 Bacterial culture

Bacterial strains or constructed mutants were stored at -80^{0}C. They were routinely grown overnight for 16-18 h at 37^{0}C atmospheric conditions in Luria Bertani (LB) broth, unless specified. If anaerobic conditions were specified this was undertaken in a variable atmosphere incubator, with 85 % N_2, 10 % CO_2, and 5 % H_2.

LB culture media was made as follows:

Reagent	Amount to add for ~1 l of broth
H_2O	950 ml
Tryptone	10 g
NaCl	10 g
Yeast extract	5 g
Stir to dissolve	20 minutes
H_2O	up to 1l
Agar for plates at 1.5 % (if required)	15 g
Sterilise by autoclave	121^{0}C, 15 psi, 20 minutes

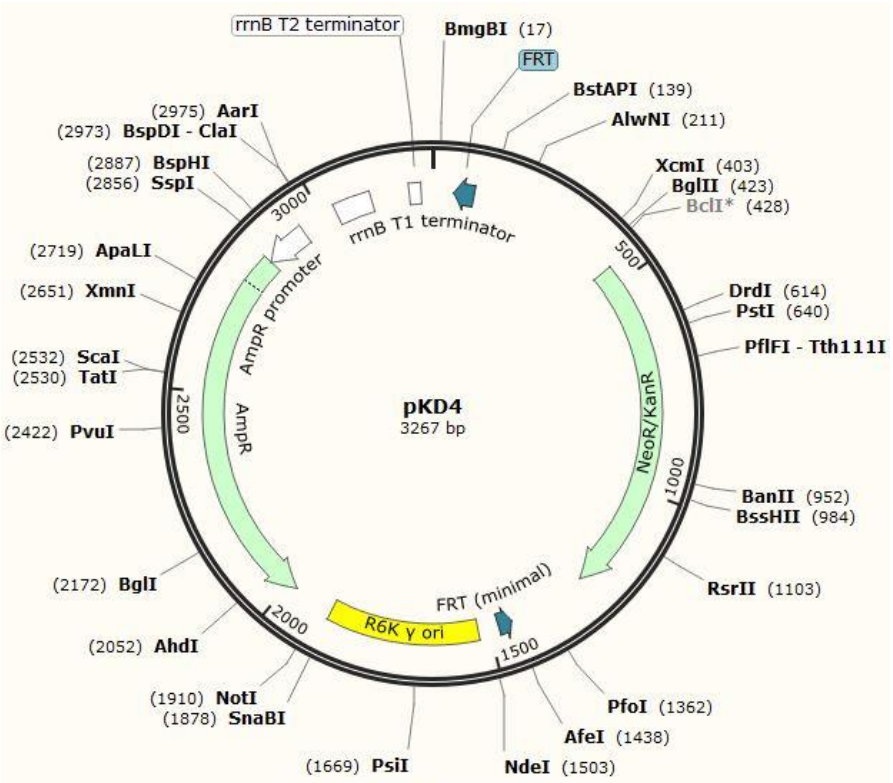
Table II.1 Ingredients and instructions for making LB broth or LB agar plates

II.2 Bacterial mutant construction

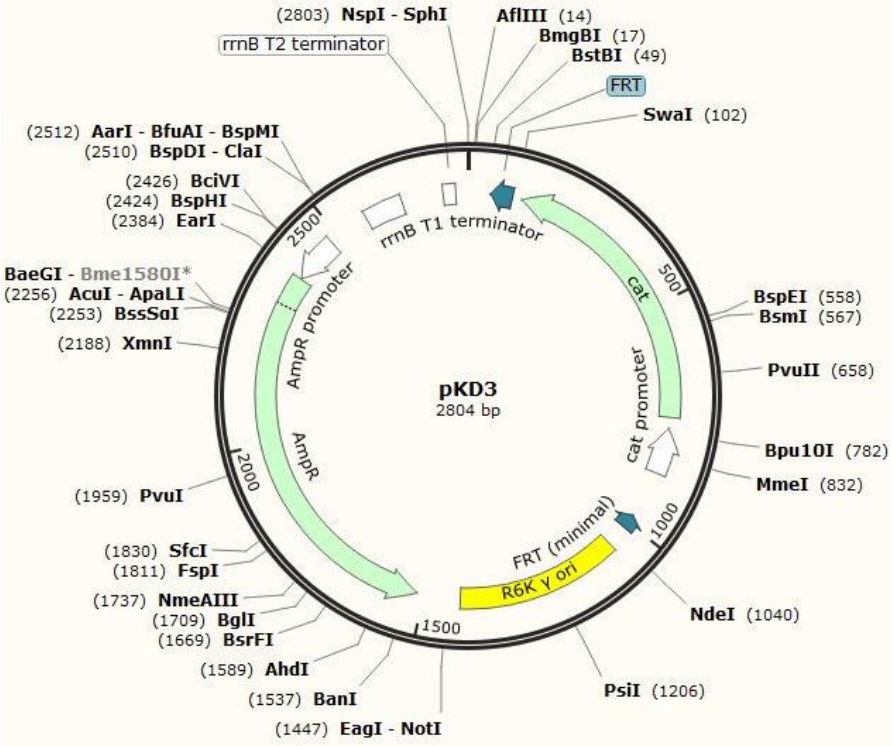
Bacterial mutants were constructed using the method described by Datsenko & Wanner (2002)(420), with substitution of recombination-based genetic engineering plasmids. Antibiotic resistance cassette containing plasmids pKD3 (chloramphenicol resistance via *cat*) and pKD4 (kanamycin resistance via *aphII*) (Figure II.1 (A) and (B)) were used as templates with primers of 50 bp flanking either side of the genomic locus being subject to resistance cassette insertion, with attached primers for polymerising the resistance cassette. The bacterial wild type or mutant which was to be subject to insertion of the cassette was electroporated with pSIM18 (containing hygromycin resistance cassette *hph*) (421). Plasmid

pSIM18 encodes phage recombination machinery genes *exo, beta,* and *gam*, from bacteriophage Lambda (Figure II.1 (C)). Gam inhibits the host RecBCD activity, therefore protecting the DNA substrate for recombination by inhibiting double strand break repair. Exo degrades nucleotides from each DNA end in a 5' to 3' direction, creating double-stranded DNA with 3' single-stranded DNA tails. Beta binds these 3' overhangs to protect and anneal them to complementary target sequences. The plasmids have temperature-controlled promoter regions, allowing plasmid reproduction and division into daughter host cells at >30^{0}C, inhibiting plasmid replication at 37^{0}C, and expressing Exo, Beta, and Gam at 42^{0}C. Electrocompetent cells were preferred over chemically competent cells due to evidence of higher transformation efficiency (422). Electrocompetent cells were made through harvesting 5 ml of overnight culture of a strain or mutant before growth to mid-log phase with 200 rpm shaking in atmospheric conditions at appropriate temperatures: 37^{0}C for non-pSIM18 containing strains or 30^{0}C for strains with pSIM18. Cells were then washed 5 times with ultra-pure water through centrifugation at 13,000 g, removal of supernatant, and resuspension in sterile ultra-pure water before adding DNA construct or plasmid. Room temperature was favoured due to evidence of greater electroporation efficiency than standard practise of keeping cells, cuvettes, and water on ice (423). Electroporation was carried out in a Bio-Rad® electroporator using settings generalised for *E. coli* in a 2 mm gap electroporation cuvette.

A


rrnB T2 terminator
BmgBI (17)
FRT
BstAPI (139)
AlwNI (211)
XcmI (403)
BglII (423)
BclI* (428)
DrdI (614)
PstI (640)
PflFI - Tth111I
BanII (952)
BssHII (984)
RsrII (1103)
PfoI (1362)
AfeI (1438)
NdeI (1503)
PsiI (1669)
NotI (1878)
SnaBI (1910)
AhdI (2052)
BglI (2172)
PvuI (2422)
ScaI (2530)
TatI (2532)
XmnI (2651)
ApaLI (2719)
BspDI - ClaI (2973)
AarI (2975)
BspHI (2856)
SspI (2887)
rrnB T1 terminator
AmpR promoter
AmpR
NeoR/KanR
FRT (minimal)
R6K γ ori
pKD4
3267 bp
500
1000
1500
2000
2500
3000

B


rrnB T2 terminator
NspI - SphI (2803)
AflIII (14)
BmgBI (17)
BstBI (49)
FRT
SwaI (102)
BspEI (558)
BsmI (567)
PvuII (658)
Bpu10I (782)
MmeI (832)
NdeI (1040)
PsiI (1206)
EagI - NotI (1447)
BanI (1537)
AhdI (1589)
BsrFI (1669)
BglI (1709)
NmeAIII (1737)
FspI (1811)
SfcI (1830)
PvuI (1959)
BaeGI - Bme1580I* (2256)
AcuI - ApaLI (2253)
BssSαI (2188)
XmnI
EarI (2384)
BciVI (2424)
BspHI (2426)
AarI - BfuAI - BspMI (2510)
BspDI - ClaI (2512)
rrnB T1 terminator
AmpR promoter
AmpR
cat
cat promoter
FRT (minimal)
R6K γ ori
pKD3
2804 bp
500
1000
1500
2000
2500

C

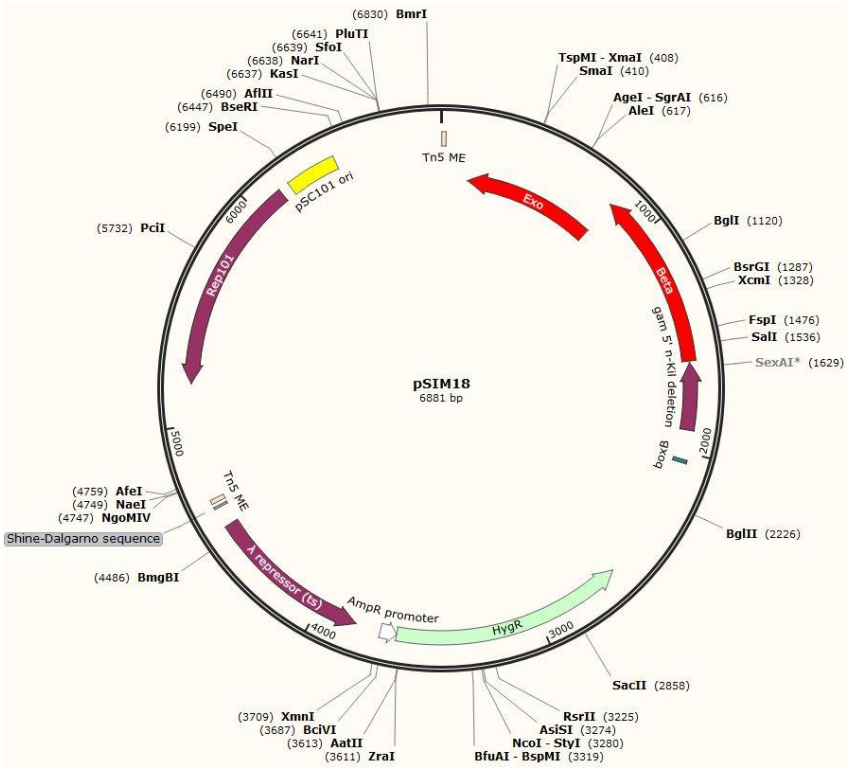


Figure II.1 Genetic diagrams of plasmids used in this study. (A) Plasmid pKD4, utilised as a template for polymerising an aminoglycoside phosphotransferase II encoding cassette (*aphII*) with 50 bp flanking regions for homologous recombination into genomic regions of interest. (B) Plasmid pKD3, utilised in the same manner as (A) but for polymerising a cassette with chloramphenicol acetyl transferase (*cat*). (C) Recombination-based genetic engineering (recombineering) plasmid pSIM18. The restriction enzyme sites, gene orientation, and genetic content of plasmids are displayed.

II.3 Phage-based experiments

II.3.1 Phage typing

Phage typing for in-house strains and mutants was carried out as described by Public Health England's phage typing protocol for *Salmonellae* & *Shigella flexneri* using typing phages 8, 10, 18, 20, 29, and 32. This is an adapted protocol first described by Anderson et al., (1977)(165). Stocks of bacteria to be typed were streaked out onto LB agar plates to form single colonies. A single colony was taken of the bacteria to be tested and incubated in 4 ml of LB broth with static, atmospheric conditions at 37^{0}C for 2 h. An LB agar plate was flooded with the culture and left to dry for >1 h and residual liquid removed carefully with a pipette. Once completely dried, 10 µl of each phage suspension at recommended titre

dilution (RTD) was spotted onto the plate and incubated for 16 h. Plaque morphology was identified visually, with reference to the phage typing scheme (Appendix I).

II.3.2 Plaque assay to determine phage titre

To determine the titre of phage within stocks provided by Public Health England and establish various starting concentrations for experiments, the top agar overlay plaque assay method was utilised (424). A single colony of the strain to be tested was harvested from an LB plate and grown for 4 h in LB broth with 200 rpm shaking in atmospheric conditions. A volume of 200 µl culture was added to 5 ml of pre-warmed, 40 ^{0}C, 0.5 % agar, then 90 µl of phage suspension at appropriate dilution was also added to the bacteria-agar mix at concentrations of neat, 10^{-1},10^{-2},10^{-3},10^{-4},10^{-5},10^{-6}. Any plates that had sufficiently visual singular plaques were counted, and the phage titre concluded as the mean average from all plates counted, with virus titre expressed as plaque forming units per ml (PFU/ml).

II.3.3 Liquid culture phage challenge assay

To assess whether mTmII was providing phage resistance to monophasic *S.* Typhimurium ST34 strains typed DT193 (Chapter V.9) the *cat* gene was introduced into an intergenic region toward the left-hand end of lysogenic mTmII in reference monophasic ST34 strain S04698-09 as described above. To isolate mTmII, *aphII* was introduced into an intergenic region of the chromosome of a phageless mutant of *S.* Typhimurium 4/74 (425). The donor S04698-09 and phageless 4/74 were grown in mixed culture, and LB agar plates supplemented with 30 µg/ml chloramphenicol and 50 µg/ml kanamycin to select mTmII lysogenic transductants of phageless 4/74. Prophage mTmII was then introduced into DT120 strains L00745-07 and L00979-07 through mixed culture with the phageless 4/74 strain with lysogenic mTmII, and mTmII lysogenic DT120 strains selected using LB agar plates supplemented with 30 µg/ml chloramphenicol and 5 µg/ml of tetracycline. Tetracycline was used to select for DT120 monophasic ST34 strains as they naturally encode tetracycline resistance locus *tetACR*, while strain 4/74 does not. Typing phages 8, 18, and 29 were used to challenge two DT120 strains L00745-07 and L00979-07, DT193 strains S04332-09 and S04689-09, and the same DT120 strains but with lysogenic mTmII (antibiotic resistance cassette insertions, phage transduction, and liquid culture assays for this experiment were undertaken by Luke Acton). Growth assessment was undertaken with a starting MOI

of 1 in atmospheric conditions at 37^{0}C with low shaking for 24 h in a Bioscreen© C MGR, reading the optical density at 600 nm every 15 minutes. The resulting data was subject to area under the curve analysis using the definite integral from 5 to 15 h with R package Plotrix (426). The area under curves for phages 8 and 18 were then grouped by phage type or mTmII presence, and the Mann-Whitney-Wilcoxon test was used to assess statistical significance.

Reference DT8 strain L01157-10, reference DT30 strain S03645-11, and L01157-10:*wzy*- were subject to liquid culture challenge with typing phage preparations 8, 10, 18, 20, 29, and 32 (Figure VI.7). This was undertaken as described above with one biological replicate and two technical replicates (assay undertaken by Luke Acton).

II.4 Bacterial resistance to phage preparation quantification

A measure of resistance was required to quantify a bacterial strain's potential ability to resist a specific set of phage preparations to infer differences between clades, sub-clades or within lineages. In this study we sought to utilise an already well curated and established set of bacteriophage interactions which has been collected over the years and maintained through widespread use of the Anderson Phage Typing Scheme (165). The more recently extended Anderson typing phages were excluded from calculations (Additional phages 1, 2, 3, 10, 10var2, 10var3, and 18) for resistance ability, as data was only available for a handful of phage types (DT193, DT193a, DT194, DT195, DT208, U302, U310, and U311). However, this data was kept within some Figures for interest and discussion when applicable. That these phages were added to the scheme for sub typing-isolates displaying resistance to all other typing phage was of possible significance when considering differences in some clades. Phage typing was carried out at PHE and the Animal and Plant Health Agency (APHA, UK) according to their method of *S.* Typhimurium phage typing, except where specified.

The resistance index (*Ri*) of a strain based on its phage type from the Anderson phage typing scheme was established such that:

$$Ri = 1 - \left(\frac{1}{p}\right) \times \left(\sum_{0}^{p} s_i = s_{0+1} + s_{0+2} \cdots s_{p-1} + s_p\right)$$

Ri refers to the resulting number between 0 and 1 that indicates a quantified level of resistance (1 being completely sensitive, and 0 being resistant to all phage tested), *p* represents the number of phages being tested for a given bacterial isolate, *s* denotes the outcome between each phage preparation and bacterial isolate of either resistance (0) or lysis (1), and *Si* refers to the summation of each strain's reaction outcomes.

II.5 Whole genome sequence data analysis

WGS data was generated using *S.* Typhimurium strains from the Animal and Plant Health protection Agency (APHA), and Public Health England (PHE), United Kingdom. Short-read, paired-end sequence data was produced by Illumina® Mi-Seq®, and long-read sequence data was generated through Pacific Bioscience® Single Molecule Real Time (SMRT®) sequencing technology, the reference strains for this study first described in Bawn et al, (2020) (39) (Appendix VIII). Draft genome assemblies were generated using De Bruijn graph-based SPAdes v-3.10.1 (427), with Kmer lengths of 31, 41, and, 51, before annotation using Prokka v-1.11(427-429). Comparative genetic diagrams were constructed using R package genoplotR (430).

II.5.1 Single nucleotide polymorphism variant calling and core genome identification

SNPs were identified in WGSs by aligning reads using BWA-MEM (431), variant calling with Freebayes (432) and SNP filtering using vcflib/vcftools (433), combined as a pipeline using Snippy-4.3.6 (434). The reference for SNP calling for the core genome alignment used to reconstruct phylogenetic trees in Chapter III and Chapter IV was *S.* Typhimurium SL1344. In Chapter V the reference for all SNP calling was S04698-09, and in Chapter VI it was DT8 reference L01157-10. Core genomes were identified using the snippy-core function of Snippy-4.3.6 (434).

II.5.2 Phylogenetic reconstruction

Many methods exist for estimating organism relatedness via nucleotide or protein sequence through phylogenetic reconstruction. Studies have shown that inference using time-free, maximum-likelihood (ML) methods for phylogenetic reconstruction produces qualitatively similar phylogenies in their bootstrap distribution than a Bayesian posterior distribution, and the output tree topology can be closer to the true topology (435). Therefore, ML based approaches were preferred for phylogenetic reconstruction from core genome SNPs. Of ML approaches, two are superior in computation speed or accuracy with tree topology and ML score – FastTree and RAxML respectively (436, 437). Due to favouring accuracy of the data over computation speed, RAxML was employed to construct most phylogenetic trees in this study (438), using the general time-reversable model of nucleotide substitution, with gamma distribution for amongst-site rate variation (GTR-gamma) with 1000 bootstraps (439). GTR was utilised as it was postulated to be the most general neutral, finite-site, independent, and time-reversable model possible, superior to other models via log-likelihood based model tests (440), although disputed (441). Using the gamma distribution for amongst site rate variation allowed for better comparison between phylogenetic trees. IQ tree was used to determine if different models of nucleotide substitution constructed phylogenies with significantly improved log-likelihoods, or fundamentally different topology (442). For the data in this study almost all phylogenetic trees were constructed using RAxML with the GTR and gamma model of amongst site rate variation. One exception is the phylogenetic tree of mTmII-like phages, which was constructed using proteomic, nearest-neighbour construction as determined by VIP-tree (Figure V.5) (443). Most phylogenetic trees presented with similar topologies whether GTR or model-optimisation was used. This scrutiny was required as GTR can generate nuisance parameters if not enough information exists to calculate every substitution rate skewing the topology and the phylogenetic trees in this study were mostly constructed from closely related isolates, which due to the similar genetic content may have not had sufficient information for complex model parameters (439).

One-hundred and thirty-three strains representing the diversity within *S.* Typhimurium, as outlined by Bawn et al., (39) were used to construct the phylogenetic tree in Chapter III.3, and 112 of the 120 that had corresponding phage typing data used to construct the phylogenetic tree in Chapter IV, Figure

IV.2, and Figure IV.11. One-thousand-four-hundred and fourteen strains from one year of PHE infections were used to construct a tree to assess cladal boundaries and sample phage types from within these clades for overrepresentation assessment in Figure IV.1. Seven-hundred-and-sixty-three monophasic *S.* Typhimurium ST34 strains collected from human infections between 2014-2015 were used to construct the phylogenetic tree from Chapter V, as described above. The phylogenetic trees in Figure V.8 and Figure V.9 were generated using 373 strains encompassing the global diversity of monophasic *S.* Typhimurium ST34. Two-hundred and ninety-two strains, mostly from ducks but also other sources, were used to construct the phylogenetic tree from Figure VI.1. One-hundred and ninety-six strains including 162 DT8 and 34 DT30 strains were used to construct the phylogenetic trees from Figures VI.3, VI.4, VI.6 and Appendix III.

The WGS of 4 phages and 2 prophages were used to construct the neighbour-joining amino-acid-based phylogenetic relationship in Figure V.5 as determined by VipTree (443).

II.5.3 Identifying and purging recombination

Recombination was detected within bacterial WGSs by identifying locations which contained >3 SNPs in a row predicted using Gubbins-2.3.4 (444), with a core genome constructed using the snippy-core function of Snippy-4.3.6 (434). Phylogenetic trees purged of recombination were produced by using the output of RAxML or IQ tree as starting trees along with the core genome alignments, and the resulting phylogenetic trees used for downstream analyses (438, 445).

II.5.4 Defining clade boundaries and sampling phage types within phylogenetic trees

The R package rhierBAPS was used to assign clade boundaries using a depth of 2 and estimated number of populations of 30 (Figure IV.1) (446). The resulting clades were used to group strains for intercladal phage type comparisons. Phage types were randomly sampled 1000 times from each second order clade using base R, and frequency values (F) determined by dividing the number of times a phage type was sampled (s) by the number of random samples taken (n = 1000) (F=s/n). The frequency value provided a measure of how densely represented a majority phage type is within each second order clade.

II.5.5 Genomic element sequence extraction and characterisation

The sequences of SGI-4, mTmII, and mTmV were identified by aligning monophasic *S.* Typhimurium ST34 reference strain S04698-09 with SL1344. The insertions were viewed and nucleotide sequences extracted using Artemis Comparison Tool (ACT) (447). Each coding sequence's translated amino acid sequence was used for searching the NCBI database using BLASTp (448), and profile hidden Markov model searches undertaken using the pFam-11.1 database (449), implemented with Hmmer-3 (450). Gene annotation was based on the consensus protein coding annotation from these searches, or protein functional domains used as annotation if a domain was present with no further characterisation.

II.5.6 SGI-4-like element searches and characterisation

Potential SGI-4-like elements (SLEs) were identified from WGS assemblies in the NCBI sequence database (accessed June 2018) by aligning sequences with SGI-4 excluding the *ars*, *sil*, and *pco* loci, using discontiguous megaBLAST. This identified nucleotide sequences from *Edwardsiella ictaluri* (accession CP001600, 326500..380800) *Erwinia tracheiphila* (accession CP013970, 1856073..1916737), *Enterobacter cloacae* (accession CP012162, 4040884..4152906), *Enterobacter hormaechei* (accession CP010376, 3932000..4028800), *Enterobacter hormaechei* (accession CP012165, 395200..536000), *Pluralibacter gergoviae* (accession CP009450, 1604785..1778603) and *Salmonella* Cubana (accession CP006055, 4214040..4311200). Phylogenetic reconstruction of putative SLEs, was undertaken through sequence alignment using ClustalW-2.1 (451) and a maximum likelihood tree was constructed from aligned full nucleotide sequences using RAxML as described above (438).

II.5.7 Identifying phages similar to mTmII

VipTree was used to assess which phage sequences were closely related to mTmII and mTmV (Chapter V.5). Phage sequences closely related to mTmII and mTmV were: SfI, SfII, SfIV, and SfV (accession numbers NC_027339.1, NC_021857.1, NC_022749.1, and NC_003444.1 respectively). Phage open reading frames (ORFs) were annotated as described in II.5.5.

II.5.8 Whole genome alignment for identifying large-scale genomic differences

Long read WGS data was utilised to identify large-scale rearrangements and recombination in reference isolates. Prophage regions were identified and labelled using PHASTER (452). Sequence similarity plots were generated using BLASTn (470), and Artemis comparison tool (ACT) (471), and multiACT cartoon generated as described by Bawn et al, (39, 453, 454). (Initial analysis and figures in Chapter IV.6 and Chapter VI.1 (C) constructed by Matthew Bawn, data reanalysed by this study). This was undertaken using 12 reference *S.* Typhimurium strains for Chapter IV.6 (strain details in Appendix VIII), and undertaken with reference DT8 strain L01157-10, reference DT30 strain S03645-11, reference *S.* Typhimurium strain SL1344, and a further DT8 strain S04527-10 for Chapter VI.1 (C).

II.5.9 Ancestral state reconstruction

II.5.9.1 Continuous trait ancestral state reconstruction

Continuous trait ancestral state reconstruction for Figure IV.5 was undertaken using Brownian-motion-based estimates using both a univariate evolutionary rate, and multivariate evolution rates. A distribution for univariate continuous trait ancestral history estimation was initially estimated and plotted using the R package Phytools (455). Brownian motion describes the trajectory of an object through space (456). This is done by determining a sequence composed of summation of random sampling from a normal distribution to determine the point in space of an object from one point to the next. For ancestral history through 'Brownian evolution', the sequence starts at the tips, using the indicated continuous trait (*Ri* in the case of Figure IV.5). Each point toward the root is calculated by drawing randomly from a normal distribution dependant on the surrounding values. The resulting sequence of points estimated through extrapolating backwards from the tips to the root can then be plotted across the phylogenetic tree branches. Univariate transition rate indicates that all points on the branches have an equal rate of transition. Multi-variate transition rates allow for different rates of evolution in different branches. Both univariate and multivariate estimations were conducted, and the significance determined through the likelihood ratio test (LRT). Multivariate Brownian evolution was determined using R package mvMORPH (457). Ninety five percent confidence intervals were calculated using the square root of the variance for each point's distribution as determined by Phytools (455).

II.5.9.2 Discrete trait ancestral state reconstruction

Discrete trait ancestral state reconstruction was undertaken with maximum likelihood calculation, maximum posterior probability estimates, and probabilistic distribution-based methods, the latter distributions estimated using a Markov-Chain with Monte-Carlo sampling (MCMC) approach. These were preferred due to literature evidence that these methods are superior to parsimony methods such as ACCTRAN or DELTRAN that produce bias at ambiguous nodes dependant on whether starting calculations at the root or the tips (458). Maximum-likelihood based states at each hypothetical ancestral node were estimated with Ancestral character estimation (ACE) from R package ape (459) and pastML (460), using maximum posterior probabilities (MPP). The transition rate matrices (Q) were estimated from the data, and models allowing for different rates of transition were used. MCMC approach was conducted with discrete character mapping using posterior sampled maps from SIMMAP (461), and plotting the results as probability density across branches of fixed maximum-likelihood phylogenetic trees, generated as previously described. One thousand sampled stochastic character maps were constructed after a burn-in period of 1000 iterations for Q, followed by 100,000 Markov chain steps and sampling for the posterior every 100 generations using the pre-computed distribution of Q (462). A burn-in period was favoured to enable a starting position of the Markov chain closer to the equilibrium distribution, enabling analyses to be conducted on a dataset of sampled posterior values with appropriate distribution range. Competing statistical models using either equal rates of transition or allowing different values for Q were compared. This was undertaken using the likelihood ratio test (LRT), computed via an in house generated R script, requiring the more complex model to reside in the right-hand most 5 % of a χ^2 distribution with 1 degree of freedom to be considered significant (Appendix IV). For MCMC calculated distributions, this was conducted using mean log-likelihoods of sampled estimated histories from MCMC analysis using both an equal rate Q, and a pre-computed distribution of values of different rates for Q with the starting value estimated from the data. Resulting data was interpreted and viewed using R package phytools (462), iTOL (463), and pastML (460).

For Figure V.2, to gain additional evidence that the phenotype of the root node was DT120 based on the outcome of ML estimation, a probabilistic Bayesian approach to ancestral history estimation was undertaken. This was conducted by labelling tips of the phylogeny as either “DT120” or “Other” to

assess the probability of this phage type occurring at nodes toward the root branches and assessing the probability that it is the root node phenotype. To rule out the common ancestor of monophasic *S.* Typhimurium ST34 being DT193 and counter the possibility that ancestral reconstruction methods may bias root nodes toward the most frequent phenotype, tips were labelled as either "DT193" or "Other", and MCMC analysis carried out as described above.

Ancestral reconstruction analyses were subject to permutation tests to identify if the proposed resulting character state at a node was due to the phylogenetic topology, and not a skewed result from sampling bias, or overrepresentation of a phenotype. Due to computational limitations 20 permutations were undertaken for each dataset. Tip labels were randomly assigned to the tree through the R command 'sample', and a starting tree generated using maximum likelihood-based function ACE, with a subsequent tree sampled every 100 iterations of MCMC, 100 times, resulting in a distribution of 100 trees with ancestral states for each permutation. To assess whether the permutated data was from the same distribution as that estimated from actual tip data, pairwise, Mann-Whitney-Wilcoxon tests were performed.

II.5.10 Identifying presence and main clade acquisition time of mTmII

The presence of mTmII in isolates from one year of UK human infections and a global isolate collection of monophasic *S.* Typhimurium ST34 was determined by extracting the sequence of mTmII from S04698-09, identifying open reading frames (ORFs) using Prokka as previously described. WGS reads were then searched for each gene of mTmII through mapping and generating local assemblies using ARIBA-2.3 (464). To identify the acquisition time of mTmII, R package Bactdating was used (465). A phylogenetic tree of a global collection of monophasic *S.* Typhimurium ST34 was analysed for temporal signal before MCMC analysis undertaken for 1 million iterations to determine the temporal structure of the tree, including 95 % confidence intervals. Fifty permutation tests were undertaken to assess if the temporal structure of the tree was dependant on the real data. The Mann-Whitney-Wilcoxon test was used to determine if the dates from each node of permutations were significantly different to the dates from nodes generated from the real data. A frequency table of the date estimate for the root

node for each permutation was constructed, and the frequency density calculated and plotted using ggpubr (466).

II.5.11 Bacterial genome-wide association

Genome wide association (GWAS) is a statistical-based method where traditionally SNPs within a genome are associated with a particular phenotype. GWAS was conducted to discern potential genetic polymorphisms associated with differences in phage sensitivity. In Chapter V this was undertaken with the WGS assemblies of 605 DT193 strains and 125 DT120 strains, and in Chapter VI undertaken with 34 DT30 strains and 162 DT8 strains. Methods of bacterial GWAS used were DNA of length K (Kmer) based analysis, where Kmers were identified from draft genome assemblies of each strain using frequency-based string mining algorithm FSM-lite-1.0 (467). Kmer based analyses are effective for bacterial GWAS, as using DNA Kmers accounts for SNPs and recombination, but SNP association methods limit association of phenotypes to SNPs (468). The population structure was estimated using genome-hash-based Mash which generated a three dimensional distance matrix (469). Subsequently, a mixed-linear-model approach was used for testing Kmer significance implemented with Sequence Element Enrichment (SEER) (470). This was initially conducted with no significance filtering, but subsequently a Kmer filtering significance level was established by keeping the top 1 % of Kmers in the range of $-\log_{10}$ p-values, resulting in an LRT p-value threshold for Kmer significance of $p<1\text{x}10^{-3}$ for Chapter VI, and $p<1\text{x}10^{-7}$ for Chapter V. Secondary bacterial GWAS to give greater certainty to the results in Chapter VI was conducted using a Kmer and De Bruijn graph approach with dbGWAS (471).

II.5.12 Assessing read depth of *wzy* locus

A 6kb locus encompassing *wzy* and flanking genes was extracted from reference sequence L01157-10, and short read Illlumina® HiSeq sequence data was mapped to this region for 196 DT8 complex isolates and 120 reference *S.* Typhimurium isolates using Bowtie-2.2.9 (472), without filtering secondary mapped reads to ensure that any reads mapping to the region would be included. Bedtools-2.26.0 (473) was used to extract the read depth per nucleotide. These were split into 250 bp bins and the mean average of each bin used as heatmap raw data for gheatmap of R package ggtree (474). The read data was normalised by dividing the raw value for 250 bp sections by the average read depth over the chromosome,

using *thrS* as a guide for the read depth of a single copy gene. This enabled optimal analysis and visualisation of each *wzy* region in the context of each sequencing run.

II.5.13 Quantifying clonal expansion using pairwise patristic distances

In order to quantify clonality and infer a measurement for clonal expansion the pairwise core-genome SNP distance between tips (patristic distance) of a core genome SNP constructed phylogeny (the co-phenetic matrix) was extracted using R package Ape (475). The pairwise SNP distance of each tip's closest relative was extracted, using this data to measure differences between clades. A clade with high patristic distance indicates that many SNPs have occurred between the isolates, increasing the branch lengths between them. A clade with low patristic distance will have fewer SNPs between them, indicating more frequent sampling of a lineage, consistent with increased clonal expansion. A limitation here is repeated isolating from the same source, which may indicate isolating the same clone many times, lowering patristic distance, and increasing assumed clonality. This method potentially produces more meaningful results when samples have been isolated from various sources, yet retain low patristic distance inferring high clonality, a trait indicative of positive selection.

II.6 Polymerase chain reaction experiments

II.6.1 Polymerase chain reaction

Polymerase chain reaction (PCR) was routinely undertaken to check the lengths of a variety of polymerised products, PCR experiments where set up using GoTaq® G2 as polymerase, with primers designed to within 2^0C of paired melting temperatures. Temperatures 4-5^0C below primer melting points were used as annealing temperatures. Primer design tool Primer3 was used to determine effective primers for a site to undergo polymerisation (476). If PCR products were to be used for sequence-sensitive downstream analysis, proof-reading enzyme Q5 was used to amplify DNA without polymorphisms that standard polymerases can introduce.

II.6.2 Quantitative polymerase chain reaction

Quantitative polymerase chain reaction (qPCR) was used to observe genotype changes when WT DT8 L01157-10 and L01157-10:*wzy*- were challenged with and without typing phage 10. The SYBR

green fluorescent reporter was used for measuring DNA quantity at each PCR cycle (Lightcycler 480 SYBR Green Master Mix, Roche Diagnostics), supplemented with high ROX reference dye (Life Technologies Ltd). All qPCR was conducted using a StepOne™ Plus Real-Time PCR System, with 1 minute at 95^{0}C for initial denaturization of genomic DNA followed by 40 cycles of 10 s at 95^{0}C, and 20 s at 60^{0}C for combined annealing and extension.

Standard curves were generated for each of the PCR reactions to identify a suitable range of template concentrations for qPCR reactions. This was done through generating a nested PCR product using primer pairs to polymerise: a 154 bp fragment within housekeeping gene *rpoD*; a 207 bp fragment within O-antigen polymerase *wzy,* and; a 214 bp fragment crossing the attB site when the *wzy-wjx* locus has been excised or deleted. Initially, a PCR was undertaken to obtain purified product templates by polymerising the products using gDNA of L01157-10 or L01157-10:*wzy*- as templates. Subsequently, the purified PCR products were diluted to 1x10^{-6} and quantified using a Thermo Scientific Nanodrop™ 2000 and Invitrogen Quibit™. Standard curves were generated from qPCR results by diluting the resulting products 7 times in 1x10^{-1} serial dilutions, polymerising each dilution, recording the cycle at which the threshold was reached, and plotting this with respective copy number, calculated by using the molecular mass of each product and estimating that one base pair weighs 650 Daltons.

To quantify genotypic changes when WT DT8 L01157-10 and L01157-10:*wzy*- were challenged with and without typing phage 10, first strains were grown overnight at 37^{0}C in shaking atmospheric conditions. Next, bacterial culture and phage lysate were adjusted to a multiplicity of infection (MOI) of 1 (~3.5x10^4 virions and bacteria). The phage-bacteria culture was then incubated at 37^{0}C in atmospheric conditions with shaking for 5 s once an hour, and 400 μl aliquots taken at 2 hourly intervals for 8 h before being frozen at -20^{0}C with 100 μ of glycerol. A final aliquot was collected 24 h post inoculum. DNA was extracted using a Promega Maxwell® and normalised to similar concentrations - within the range observed to be accurate for qPCR with these primer pairs through standard curve calculation as previously described. This experiment was repeated 3 times. Data was analysed using a script constructed as part of this study, a self-programmed $\Delta\Delta$ Ct method (Appendix V).

II.7 Island transfer experiments

Selectable markers conferring resistance to specified antibiotics above innate levels were inserted into islands to construct donor and recipient strains to check for genetic element transferability. Recipients were marked with an antibiotic resistance cassette different to those in donor cells, agar plates supplemented with corresponding antibiotics used for enumerating donors and recipients, and agar plates with both antibiotics used for dual selection used to quantify transfer as a proportion of donor or recipient cells. Both donor and recipients were quantified with appropriate single selection plates to check for any growth disturbances. These experiments were utilised to test transfer capability of SGI-4 from monophasic ST34 S04698-09 and *wzy-wjx* from DT8 L01157-10 due to WGS and literature evidence indicating they may be mobile elements. A further step was taken to introduce a mutation conferring nalidixic acid resistance to recipient strains of L01157-10:*aphII, wzy-* for transfer in Chapter VI.6

II.7.1 Assessing transfer of SGI-4

For a selectable marker indicating the presence of SGI-4, a strain of monophasic *S.* Typhimurium S04698-09 in which the *bar* gene on SGI-4 was replaced by the *cat* gene (S04698-09 SGI-4 Δ *bar*::*cat*) was constructed, conferring resistance to chloramphenicol. To provide a selectable marker for the recipient strain, a strain of *S.* Typhimurium SL1344 was constructed in which the *copA* gene was deleted and replaced by an *aphII* gene, conferring resistance to kanamycin (primer table can be viewed in Appendix III). Donors and recipients were cultured as previously described. The OD_{600} nm of each culture was adjusted to 0.1 with fresh LB broth and 2.5 ml of each added to a 50 ml tube and incubated statically for 18 h at 37^{0}C in aerobic or anaerobic atmosphere, and in the presence or absence of 0.5 mg/l mitomycin C. The number of CFUs per ml of donors and recipients were quantified by culturing serial dilutions on LB agar supplemented with 0.03 mg/l chloramphenicol or 0.05 mg/l kanamycin, respectively. The presence of SGI-4 Δ *bar*::*cat* in recipient strains was quantified by serial dilution on LB agar supplemented with chloramphenicol and kanamycin. Transfer frequency was defined as the number of recipients containing SGI-4 Δ *bar*::*cat* as a proportion of donor cells in the culture. To determine whether transconjugant recipient strains contained SGI-4 in the same chromosomal location as the

donor, the predicted right junction was amplified using primers that annealed on either side of the right junction of SGI-4 by PCR. This experiment was repeated 3 times.

II.7.2 Assessing SGI-4 circularisation

To capture SGI-4 in the process of circularisation 0.5 µg/ml mitomycin C was added to broth containing OD_{600} adjusted S04698-09 and SL1344, and 2 µL aliquots taken at 2, 4, 6, & 8 h. To detect any possible circularisation of SGI-4 after excision, outward facing primers that annealed at the left and right junction of SGI-4 were used for PCR amplification, and nested PCR performed to increase yield. The resulting sequence of the amplicons was determined by Sanger sequencing, using four sequencing primers designed to cross the whole 2Kb product in four reactions (Eurofins sequencing service) before aligning of sequencing results to the genome of strain S04698-09. This process was repeated twice.

II.7.3 Assessing transfer of *wzy* locus

To test the hypothesis that the *wzy* locus was transferable between a mixed genotype population of a single strain, a chloramphenicol selectable marker (*cat*) was introduced using recombination as previously described into the *wzy-wjx* locus of *S.* Typhimurium L01157-10 and was tested to ensure this did not interfere with the region deletion through phage selection and PCR. This gene knock-in was preferred over direct allelic replacement and deletion to conserve the nucleotide sequence of *wzy* for any possible sequence recognition or binding that may have be required for deletion or transfer, and to be able to test return of the phenotype via *wzy* reversion if transferable. Co-culture was set up for 24 h in static conditions with L01157-10 Δ *wjx*::*cat* and L01157-10::*wzy*-, *aphII, nalR,* a selected mutant lacking the *wzy* locus selected as colonies growing in phage plaques after challenge with typing phage 10. This mutant also carried a kanamycin cassette introduced as a knock-in upstream of *iciA,* and a hypothesised mutation in *gyrA* selected through challenge with high CFU of cells grown on 30 µ g/ml nalidixic acid. This double selection for the recipient was required to ensure that the *wzy:cat* locus was transferring to the recipient, and not the *aphII* cassette to the donor. Transfer experiments were repeated 3 times without mitomycin C and 3 times with mitomycin C to address the hypothesis that phage transduction may be involved in transfer the locus.

II.8 Determining minimum inhibitory concentrations

Strains or mutants to be tested were grown overnight for 16 h at 37^{0}C with shaking in either atmospheric conditions or anaerobic conditions (10 % CO_2, 10 % H_2, and 80 % N_2), using a Whitely A95 Anaerobic Workstation (Don Whitley scientific) and corresponding heavy metal compound ($CuSO_4$, $AgNO_3$, or H_3AsO_4) added to 50 ml of LB broth and the pH adjusted to 7.4. Concentrations of $CuSO_4$ tested included: 0, 1, 2, 3, 4, 5, 6 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 and 20 mM, concentrations of $AgNO_3$ tested included: 0, 0.3, 0.5, 0.7, 0.9, 1.2, 1.4, 1.7, 1.9, 2.1, 2.4, 2.6, and 2.8 mM, and concentrations of H_3AsO_4 tested included 0, 0.5, 1, 1.5, 2, 2.5, 3, 3.4, 4, 4.4, 4.9, 5.4, and 6 mM. Each heavy metal concentration tested was in a volume of 250 μl LB broth accounting for the subsequent addition of 50 μl of bacterial culture and solutions were added to a 96 well plate. Plates were left for 24 h in the desired condition for acclimatisation before inoculation with the strains or mutants to be tested, which was done by adjusting the OD_{600}/ml to 0.3 and adding 50 μl to each well, resulting in a starting OD of 0.05. These were grown for 24 h in atmospheric or anaerobic conditions as stated above. The concentration at which the OD_{600} was below 0.2 was considered the minimum inhibitory concentration (MIC). Copper sulphate MICs were repeated 3 times, silver nitrate and arsenic acid MICs undertaken once.

II.9 Statistical analyses

All data was analysed using base R or R packages Stats, pysch, ggpubr, and kSamples (466, 477-479). Summary statistics were determined for any datasets being tested for significant differences using describeBy (480), and data distributions tested for normality using normal theoretical quantile-quantile plots and correlating this with the observed data. Frequency plots of data points were subsequently generated for visualising distributions and the Shapiro-Wilks and Kolmogorov-Smirnov tests used to reject a null hypothesis of normally distributed data (481). For one or two sample comparisons requiring non-parametric tests, the Mann-Whitney-Wilcoxon and Kruskal-Wallace tests were used to determine significance, typically for α levels of 0.05 and 0.01, depending on the dataset (482-484). A combination of tests was used, and if the null hypothesis was rejected at alpha levels of <0.01 for all tests no further query into the data was undertaken. All dataset distributions in this study were significantly different to

a normal distribution. As previously described, competing models of nucleotide substitution for use in phylogenetic reconstruction and ancestral state reconstruction were scrutinised using log-likelihood values for the LRT test, as well as the Bayesian information criterion (BIC) as implemented in ModelFinder as part of the IQTree pipeline (485-487). Linear regression models were generated using base R, and R^2 values for regression data calculated programmatically as demonstrated (Appendix IV).

Chapter III: *Salmonella* Genomic Island 4 of Monophasic *S.* Typhimurium ST34

In this study novel aspects of genome evolution that may explain the emergence of *S.* Typhimurium ST34 were investigated. Recently a novel genomic island termed SGI-4 was discovered that was specifically associated with *S.* Typhimurium ST34 (184). SGI-4 was discovered in a retrospective study of epidemic isolates of monophasic *S.* Typhimurium ST34 from the UK and proposed to be an MGE. SGI-4 was shown to encode heavy metal resistance genes as well as candidate genes involved in mobilisation and transfer. Parsimony-based ancestral reconstruction suggested the element was introduced into a common ancestor of the monophasic *S.* Typhimurium strains analysed and lost at low frequency. Early ancestral insertion and minor subsequent losses from the monophasic ST34 population is consistent with SGI-4 being an MGE and countering selection pressures toward the initial stages of clonal expansion. It was proposed that copper resistance genes encoded on SGI-4 may have been beneficial for circulation in pigs due to use of copper as an antimicrobial and growth promoter in pig farming (213). This study hypothesised that the acquisition of SGI-4 was beneficial to the success of the monophasic *S.* Typhimurium ST34 clade due conveying enhanced resistance to heavy metals. The aim of this study was to characterize this novel genomic island by addressing the following questions: i) what type of genomic element is SGI-4, and what is its genetic content?; ii) how is SGI-4 related to other MGEs?; iii) does SGI-4 or its encoded heavy metal resistance genes exist within other lineages of *S.* Typhimurium?; iv) are the observed copper efflux encoding genes phylogenetically distinct from those found chromosomally in *E. coli*?; v) is SGI-4 self-transferable?, and; vi) are the heavy metal resistance genes on SGI-4 functional?

III.1 SGI-4 is an integrative conjugative element encoding genes for self-transfer and heavy metal resistance

To characterise SGI-4 and predict its coding capacity the nucleotide sequence was determined by defining the site of insertion. The insertion site was identified through aligning the complete and closed

WGS of DT104 reference strain NCTC13384 and monophasic *S.* Typhimurium ST34 reference strain S04698-09 (39). A gene model was predicted using Prokka and each coding sequence was annotated by searching protein families through pFam-11.1 and the NCBI non-redundant amino acid sequence database using BLASTp for each predicted amino acid sequence (Figure III.1 (B))(449, 488).

SGI-4 was inserted adjacent to a phenylalanine tRNA encoding locus and consisted of 80,741 bp with 87 open reading frames (ORFs) that were absent from DT104 strain NCTC13384. The DNA excision, DNA processing, and conjugative transfer mechanisms encoded on SGI-4 were those that previously have been shown to enable transfer of mobile genetic elements (MGEs) such as ICE, also known as conjugative self-transmissible integrative (CONSTIN) elements (489), integrative mobilizable elements (IMEs), and cis-mobilizable elements (CIMEs) (490). Several ORFs encoded on SGI-4 exhibited sequence similarity to DNA processing enzymes. Two ORFs had sequence identity to genes encoding machinery for excision and integration; *xerC* and *xerD*. No recombination directionality factor (RDF) was observed, although, these are known to exhibit little sequence conservation (491). ORF86 had sequence similarity to *traI* that encodes a relaxase protein (492). Relaxase proteins bind dsDNA at the origin of transfer (oriT) and induce a single strand nick. In SGI-4 this may occur in conjunction with putative DNA helicase encoded by a *uvrB* (ORF85). DNA unwinding may by facilitated by these proteins together with a putative topB protein (ORF13) (493). The left side of SGI-4 encoded *parA* and *parB* orthologues (ORFs 1 and 5) which translate to putative chromosome partitioning proteins (494). Several genes were encoded across SGI-4 that have been shown to be components of a conjugative T4SS: ORF26 encoded putative type IV coupling protein TraD, which initiates conjugal transfer (495); ORF51 encoded a putative inner membrane protein, TraG, for T4SS stabilisation (496); ORF53 encoded a TraU orthologue involved in pilus assembly (497); ORF19 encoded a PilL F-type pilus protein (497), and; ORF45 encoded an F-pilus assembly protein (497). These ORFS were in separate operons (ORFs 14-27, 37-46, 48-53), along with genes encoding proteins with no significant similarity to those in searched databases, but frequently associated with MGEs (Figure IV.1). No putative *oriT* sequence, required for relaxase binding, was identified in SGI-4 using oriTfinder (498). However, two pairs of inverted repeats that may be associated with this function were present, each with 80 % sequence identity within their

repeat sections. One was present in ORF19 and the other in an intergenic region between ORF64 and ORF65.

Arsenic resistance encoding genes were present at one locus in SGI-4. These included: a putative arsenical efflux pump (*arsAB*) present at ORFs 30 and 31, respectively (499); a gene encoding ArsC, which enzymatically catalyses reduction of arsenate(V) to arsenite(III) present at ORF32 (499); an arsenite(III) metallochaperone ArsD encoding gene at ORF29; a second copy of *arsA* (ORF34) and *arsD* (ORF35), and; regulatory genes *arsR* (ORF28) and *sigL* (ORF36) (500) (501). This locus also had a gene with similarity to N-acetyltransferase from Enterobacterales species that may have a secondary function conferring resistance to the herbicide glufosinate, as observed for N-acetylytransferase gene (*bar*) from *Streptomyces* (ORF34)(502, 503). Genes with similarity to those conferring resistance to copper and silver, interchangeably annotated as both *sil* and *cus* in the literature, encode a resistance nodulation division (RND) efflux pump, regulation genes, and molecular sponge (*sil/cusABFCRSE)* were present at ORFs 72, 71, 70, 69, 68, 67, and 66 respectively (504, 505). *S.* Typhimurium strains commonly encode two P-type ATPase efflux proteins, CopA and GolT, and copper and zinc binding protein CueP (506). Addition of an extra efflux pump may be enhancing copper resistance. SGI-4 also encodes genes similar to those for a putative P-type silver efflux pump encoding gene (*silP*, ORF74), and a second copy of heavy metal molecular sponge encoding gene (*silE* , ORF73)(507). Genes *silP* and *silE*, along with the reportedly dual function RND copper and silver efflux pump, were hypothesised to be enhancing resistance to silver (505, 508). The RND efflux pump is thought to pump copper and silver ions from the cytoplasm directly to the external milieu (504). Adjacent to the *sil/cus* genes were genes encoding a copper detoxification system (*pcoGEABCDRSE*, ORFs 76-84). Similar genes are present on the chromosome of some *E.coli* strains and have been shown to encode periplasmic copper handling and detoxification proteins, detoxifying cuprous ions (Cu^{2+}) to the less toxic cupric ion (Cu^{1+})(509, 510).

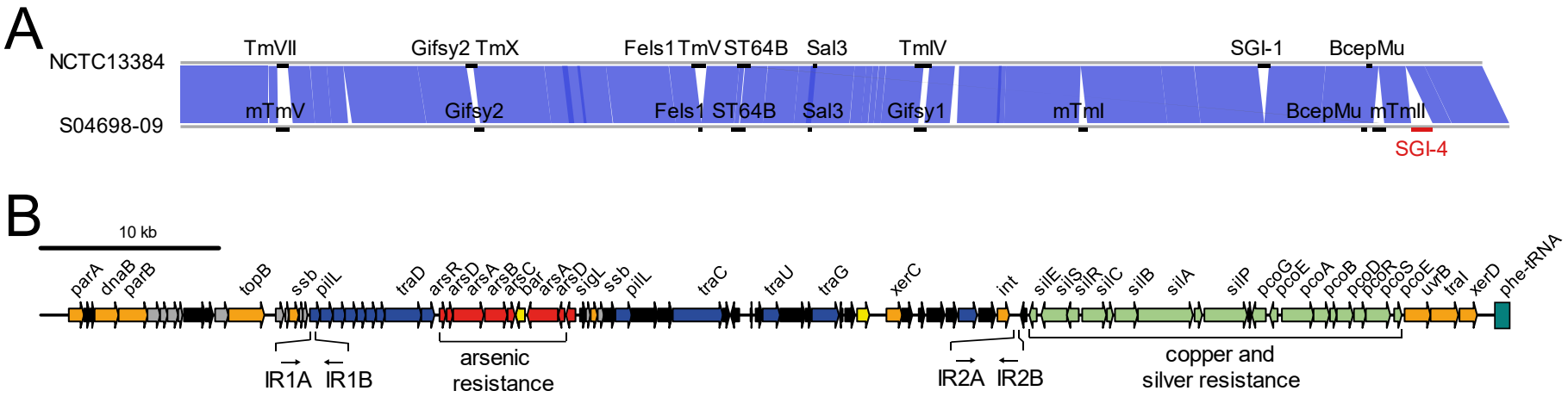


Figure III.1 Chromosomal insertion site and genetic organisation of SGI-4. (A) Chromosome alignment of NCTC13384 and S04698-09 showing differences in large insertions including SGI-4 (red horizontal line), and prophages (black horizontal lines). (B) Gene model of SGI-4 with predicted open reading frames (arrows) colour coded for predicted function as well as inverted repeats, and the hypothetical insertion site. Genes are colour coded for proteins involved in: integration, excision and DNA processing (orange); a hypothetical novel type IV secretion system (blue), arsenic resistance (red); copper and silver resistance (green), hypothesised phenylalanine tRNA encoding insertion site (dark turquoise), and genes of functional interest (yellow), including *bar* (yellow, left), a gene hypothetically encoding resistance to the herbicide glufosinate, and a probable DNA adenine methyltransferase (yellow, right). Sequences encoding hypothetical proteins commonly associated with ICEs are shown (grey), as well as other hypothetical proteins (black).

III.2 SGI-4 is a member of a candidate novel family of ICE

To investigate whether closely related elements to SGI-4 were present in other bacteria, a comprehensive search of the NCBI nucleotide database was undertaken. This search was conducted with genes from SGI-4 predicted to be involved in transfer. Nucleotide sequence similarity of transfer genes was conducted because inclusion of the putative copper and arsenic resistance loci skewed searches due to their widespread distribution on other mobile genetic elements (*sil*) and Enterobacterales chromosomal copper resistance loci (*cus*). Alignment of nucleotide sequences in the NCBI database to SGI-4 transfer genes using BLASTn resulted in identification of 7 candidate chromosomal regions not previously described as MGEs. These MGEs were named SGI-4-like elements (SLEs), and, due to transfer gene synteny and nucleotide sequence identity including genes for a possible T4SS and integration, are proposed to comprise a new family of ICEs. All bacteria containing SLEs were members of the order Enterobacterales, suggesting order specific transfer of SLEs. Each MGE had >74 % sequence identity in shared genetic content, mostly within transfer genes. Phylogenetic reconstruction indicated that the elements shared a common ancestor. Two clades were identified (Figure III.2). One clade contained SGI-4 and two closely related MGEs, one from a strain of *Edwardsiella ictaluri* which causes septicaemia in catfish, and the other *Erwinia tracheiphila* which causes bacterial wilt in the plant family Cucurbitaceae. The second clade contained five SLEs, all of which encode heavy metal resistance genes. The synteny of transfer genes was identical in each SLE, although frequently interrupted by genes inserted between the transfer gene modules. The only other SLE from the genus *Salmonella* was from an isolate of *S. enterica* serovar Cubana, which encoded the arsenic efflux pump common to other SLEs, copper and silver resistance genes, cadmium resistance genes, antibiotic resistance genes, and several IS elements with moron genes.

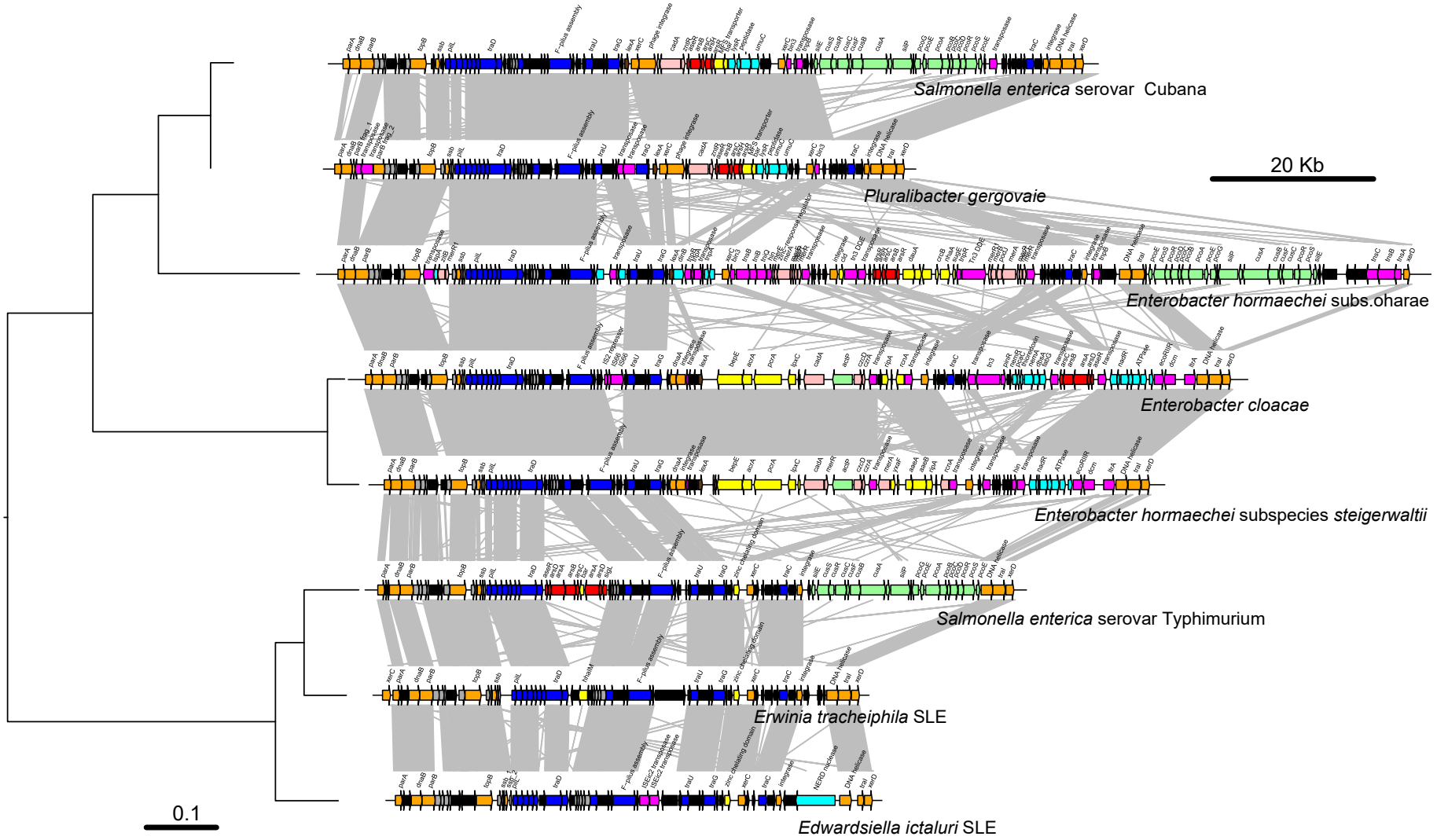

Salmonella enterica serovar Cubana
20 Kb
Pluralibacter gergovaie
Enterobacter hormaechei subs.oharae
Enterobacter cloacae
Enterobacter hormaechei subspecies steigerwaltii
Salmonella enterica serovar Typhimurium
Erwinia tracheiphila SLE
Edwardsiella ictaluri SLE
0.1

Figure III.2 Evolutionary relationship and comparison of genetic content of the SGI-4-like element (SLE) family of integrative conjugative elements. Filled arrows represent open reading frames showing the size and orientation of each gene. Genes are grouped by predicted function: integration, excision, and DNA processing (orange); hypothetical type IV secretion system (T4SS) (dark blue); copper and silver resistance genes (green); arsenic resistance genes (red); cadmium, cobalt, mercury and zinc resistance genes (pink); transposon and insertion sequence associated genes (purple); antibiotic resistance, herbicide resistance, and efflux pumps (yellow); genes of various functions hypothetically inserted via transposons (bright blue), repressor gene *lexA* (brown); hypothetical genes frequently associated with MGES (grey), and; other hypothetical genes (black). The phylogenetic tree scale bar shows substitutions per site in corresponding branch length. SLE labels indicate the host organism for each SLE.

III.3 SGI-4 and its genetic content is specific to the monophasic *S.* Typhimurium lineage

To investigate whether other lineages of *S.* Typhimurium contained SGI-4 or SGI-4 encoded genes, presence of SGI-4 genes in the short-read sequence of 120 representative strains of *S.* Typhimurium were identified by mapping and local sequence assembly using ARIBA (464). To place strains containing SGI-4 in a phylogenetic context, a maximum-likelihood tree of 120 strains was reconstructed based on sequence variation in core genome sequences (Figure III.3). SGI-4 transfer genes were only present within the monophasic *S.* Typhimurium lineage. A single strain (3193-1995) contained the copper resistance encoding loci (*sil/cusABFCRSE* and *pcoGEABCDRSE*) but not the transfer genes (Figure III.3). Two strains that shared a recent common ancestor with the monophasic *S.* Typhimurium ST34 lineage had the copper efflux pump encoding genes (*sil/cusABFCRSE*) but not the copper detoxification machinery (*pcoGEABCDRSE*) or transfer genes. In each case it is likely that these genes were present on a different mobile genetic element.

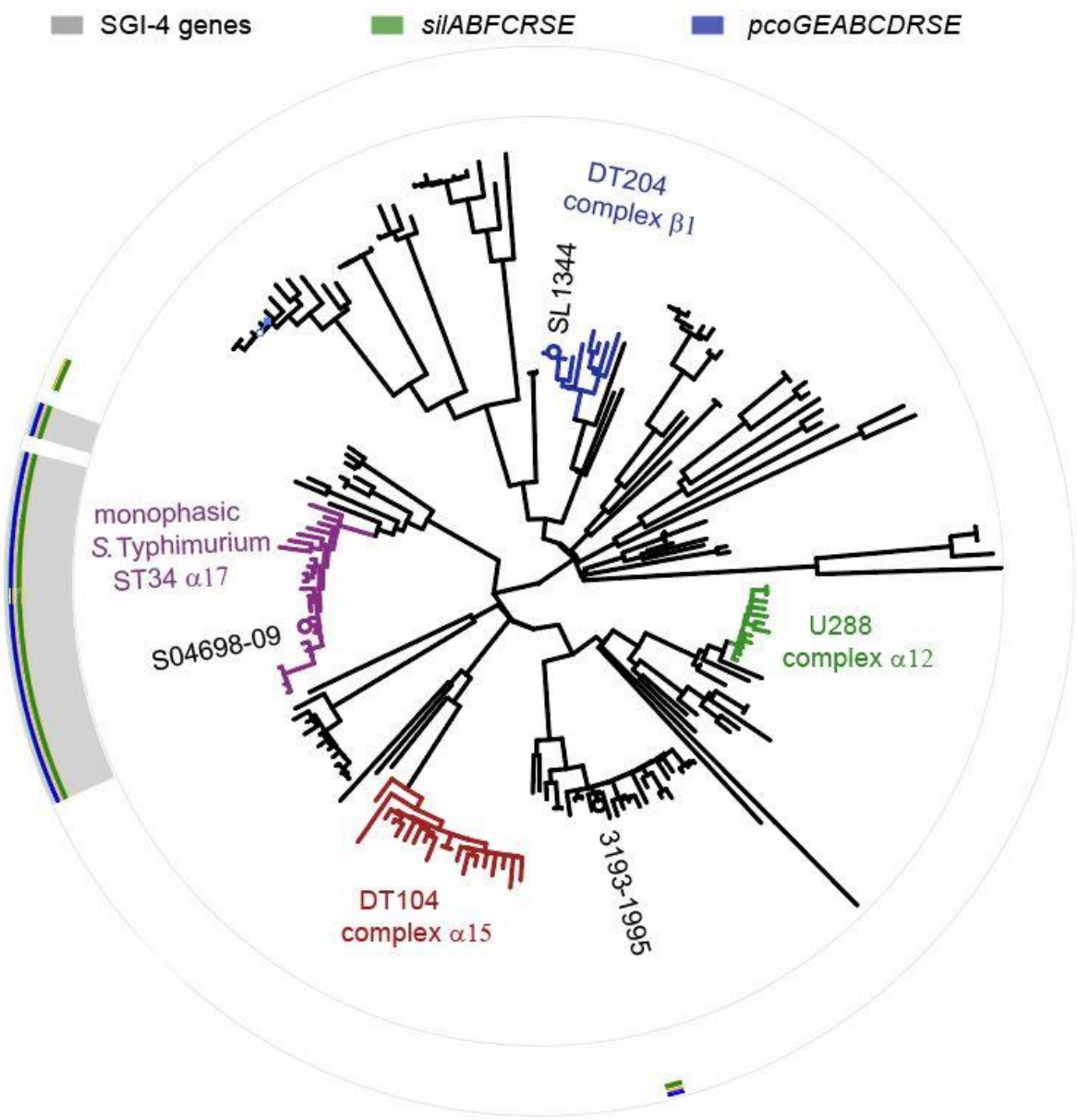


Figure III.3 Presence of SGI-4 genes within population representative isolates of *S.* Typhimurium. Phylogenetic reconstruction of 120 population representative isolates was undertaken using 17,823 core genome SNP sites. Four clades with distinct host ranges are labelled: broad-host DT204 majority type lineage β1, porcine adapted U288 majority type lineage α12, broad-host DT104 majority type lineage α15, and monophasic ST34 lineage α17. The ring around the tree displays SGI-4 transfer genes (grey), copper efflux pump encoding genes (green), and copper detoxification encoding machinery (blue). Figure and data generated by Matthew Bawn, data analysed by the study author.

III.4 The SGI-4 encoded copper efflux system is phylogenetically distinct from that found chromosomally in *Escherichia coli*

A locus composed of 18 ORFS on SGI-4 had 14 genes which exhibited sequence similarity to those hypothesised to encode a dual function, copper and silver ion, RND-family efflux pump and its regulation. Genes from the efflux pump encoding locus are interchangeably called *cus* or *sil* in the literature without reference to ancestry (*cusCFBA* and *silRSE silCBAP*)(507, 508, 511-513). The phylogenetic relationship was determined for *cus* and *sil* genes from 8 plasmids, 6 chromosomal MGEs and 2 from uncharacterised locations. Through aligning the RND-family efflux pump genes of SGI-4 with sequences from previously characterized homologues, it was determined that *cusRSCFBA* genes from the *E. coli* chromosome formed a distinct outgroup from a closely-related cluster that included the SGI-4 genes. This is consistent with the SGI-4 efflux pump encoding genes evolving from a common ancestor closely related to the *silRSE silCBAP* on pMG101 (Figure III.4)(513).

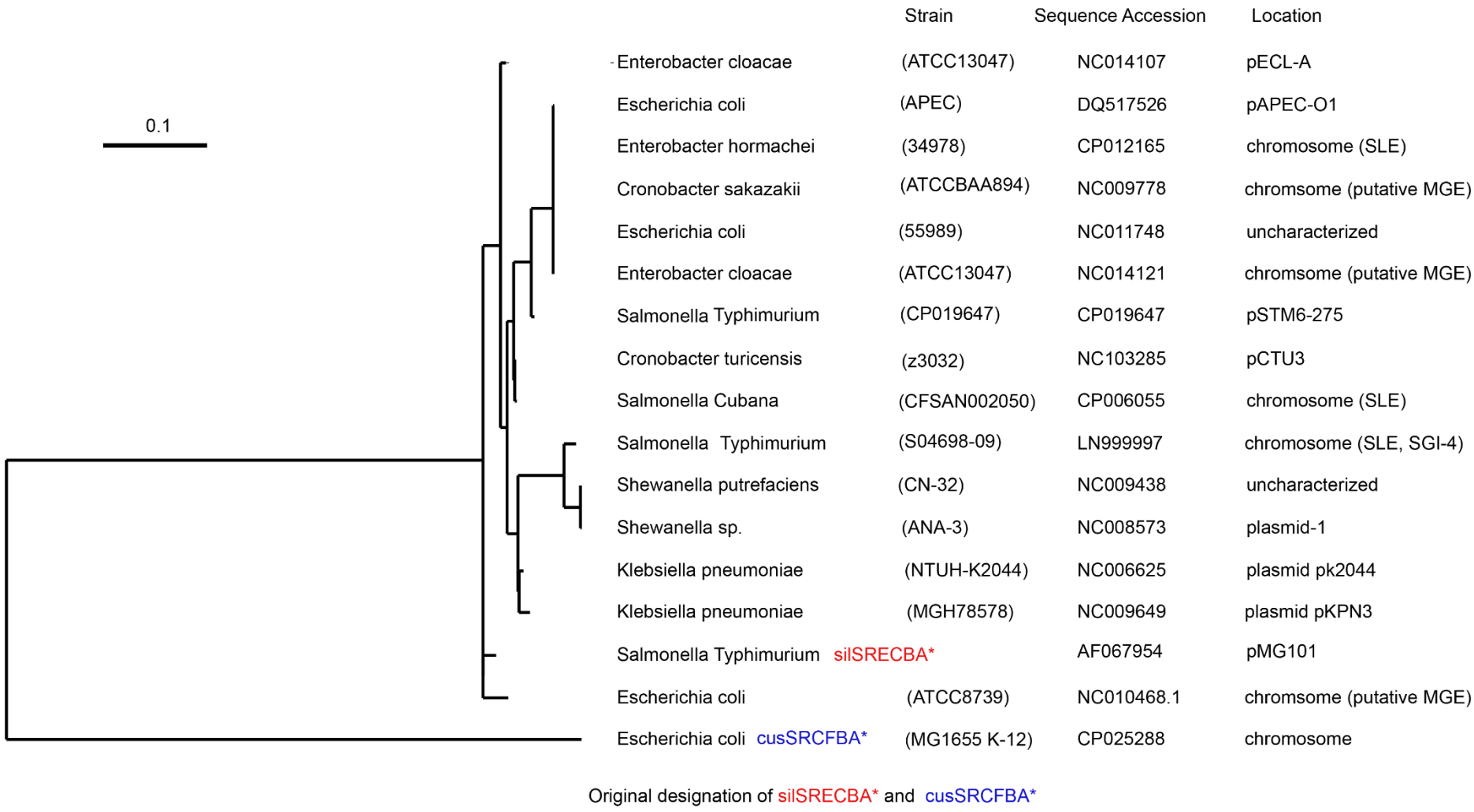


Figure III.4 Phylogenetic relationship of copper and silver efflux pump encoding *cus* and *sil* genes from Enterobacterales mobile genetic elements and the chromosome of *Escherichia coli* K-12. Maximum-likelihood phylogenetic tree was constructed using alignments of each

silSRECBA or *cusSRECBA* locus, extracted from the nucleotide sequence of the displayed accession number. The tree scale bar shows SNPs per SNP site in a corresponding branch length.

III.5 SGI-4 is self-transferable using mechanisms with characteristics typical of ICE

To address whether SGI-4 was transferable between strains of *S.* Typhimurium antibiotic selectable markers were introduced into SGI-4 in a donor strain and a second antibiotic selectable marker was introduced into a strain lacking SGI-4. A *cat* gene conferring resistance to chloramphenicol was inserted into the *bar* gene of the arsenic resistance locus of S04698-09 by allelic exchange (S04698-09 SGI-4::*cat*) enabling selection of cells with SGI-4 by culture on media supplemented with chloramphenicol. To construct a recipient *S.* Typhimurium strain, an *aphII* gene conferring resistance to kanamycin was inserted into the *copA* gene of reference *S.* Typhimurium SL1344. Mixed culture and subsequent single and dual antibiotic selection were then used to assess SGI-4 transferability by determining CFUs of recipients (kanamycin resistant), donors (chloramphenicol resistant) and transconjugants (kanamycin and chloramphenicol resistant) (Figure III.5). SGI-4 was transferred to recipient cells at low frequency without stress ($5x10^{-9}$ CFU per donor cell, Figure III.5 (E)), which was significantly increased when the mixed culture was grown with oxygen tension ($p<0.05$) or mitomycin C ($5x10^{-6}$ CFU per donor cell, $p<0.01$) and a further enhanced transfer observed if challenged with both ($5x10^{-5}$ CFU per donor cell, $p<0.01$). To confirm the insertion site and presence of SGI-4 ~2Kb of the right-hand end of SGI-4 was polymerised and viewed on a gel (Figure III.5 (C) and (D)). A faint band was observed for using primers A and C with SL1344 gDNA and was hypothesised as spill-over from another well; presence of SGI-4 would present with a band intensity as seen for the S04698-09 and SL1344:SGI-4. For assessment of a hypothetical circular SGI-4 intermediate, common to most ICE life-cycles, a nested PCR product was generated using primers facing outwards of each end of SGI-4 and the resulting PCR product sequenced by Sanger sequencing (Figure III.5 (B)).

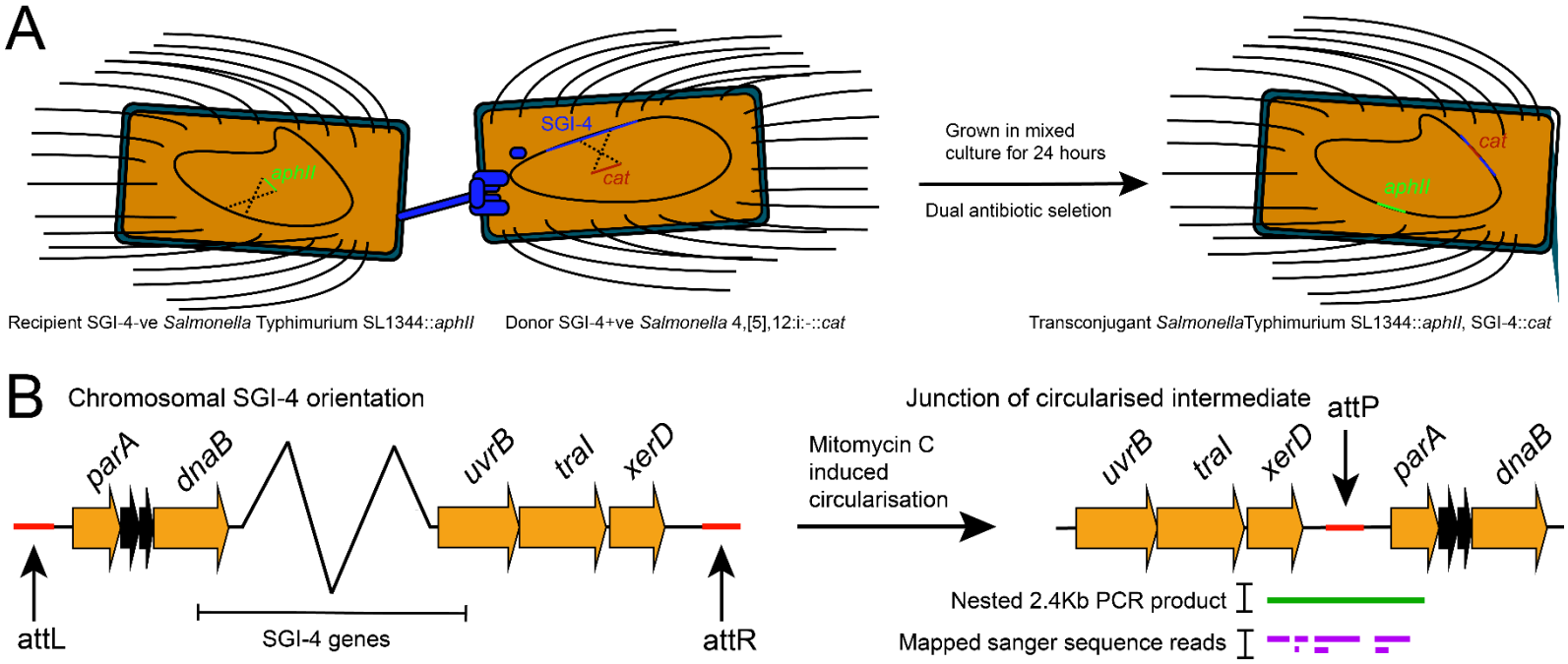

A
aphII
SGI-4
cat
Grown in mixed culture for 24 hours
Dual antibiotic seletion
cat
aphII
Recipient SGI-4-ve Salmonella Typhimurium SL1344::aphII
Donor SGI-4+ve Salmonella 4,[5],12:i:-::cat
Transconjugant SalmonellaTyphimurium SL1344::aphII, SGI-4::cat
B
Chromosomal SGI-4 orientation
parA
dnaB
uvrB
traI
xerD
attL
SGI-4 genes
attR
Junction of circularised intermediate
Mitomycin C induced circularisation
attP
Nested 2.4Kb PCR product
Mapped sanger sequence reads
C
SO4698-09:SGI-4
SL1344
SL1344:SGI-4
Phe-tRNA
yjdC
ethD
hmrR
(4742300bp)
(4746000bp)
(4585200bp)
(4587940bp)
(4664920bp)
(4668620bp)
2.1kb (1)
2.4kb (2)
No amplification
D
1 Kb Ladder
S04698-09
SL1344::SGI-4
SL1344
16S (+ve control)
10
3
2
0.5
E
Transfer per donor cell (log10)
Mitomycin C (μg/ml)
Aerobic
Anaerobic
***
*

Figure III.5 Determination of SGI-4 transfer between *S.* Typhimurium donor and recipient strains. (A) Hypothetical interaction between donor and recipient, showing the insertion locations of antibiotic resistance cassettes: *aphII* enabling kanamycin resistance (green), and *cat* enabling resistance to chloramphenicol (red). (B) Summary of experimentation to induce circularisation of SGI-4 using transfer inducing DNA intercalating molecule mitomycin C. Genomic arrangement of SGI-4 in the chromosome, flanked by DNA repeat sites attL and attR (left), and the predicted junction that would be generated following excision and circularisation (right). The amplicons using PCR primers to amplify 2.4Kb across the hypothetical circular junctions, and purple lines indicate the resulting reads when this PCR product was sequenced using four internal primers. (C) Diagram of the regions for amplification to detect presence of SGI-4 in recipient transconjugants. (C, primer A) The forward primer for polymerisation of both presence of SGI-4 and determining the same insertion site. (C, primer B) Location of reverse primer for determining SGI-4 presence. (C, primer C) Location of reverse primers for polymerising outside of SGI-4, within the hypothetical insertion site (Phe-tRNA). (D) Agarose gel separated bands of the products from PCR reaction 1 and 2 from (C). (E) Quantity of transconjugants per donor cell in aerobic and anaerobic conditions, with and without DNA damage inducing mitomycin C.

III.6 SGI-4 confers enhanced resistance to copper, silver, and arsenic

Presence of a coding sequence for a gene discovered through WGS does not mean that the gene encodes for a functional protein. This is particularly true for complex systems which have been transferred between bacteria, as each gene is required to fit in well with the host regulatory networks, interact with host transcription factors, and require effective interactions with host proteins (514). The ability of MGE encoded proteins to interact more efficiently with the host proteins than existing genes governs the success and spread of MGEs through a phenomenon known as bacterial genetic dominance (194). To investigate if the encoded heavy metal resistance genes on SGI-4 were functional, minimum inhibitory concentrations for their survival in liquid culture with heavy metal compounds was assessed. Three heavy metal compounds were tested: copper sulphate ($CuSO_4$), silver nitrate ($AgNO_3$), and arsenic acid (H_3AsO_4). The metal ions in liquid culture were therefore Cu^{2+}, Ag^{+}, and AsO_4^{3-}. The three

heavy metal compounds were utilised to test the hypotheses that the dual function copper and silver efflux pump was functional, as well as the arsenic efflux pump. Strains which harboured SGI-4 as determined through WGS were grouped and tested against strains from lineages which did not. SGI-4 positive strains exhibited significantly increased copper, silver, and arsenic compound resistance in both atmospheric and anaerobic conditions. This is consistent with SGI-4 conferring resistance to these heavy metals.

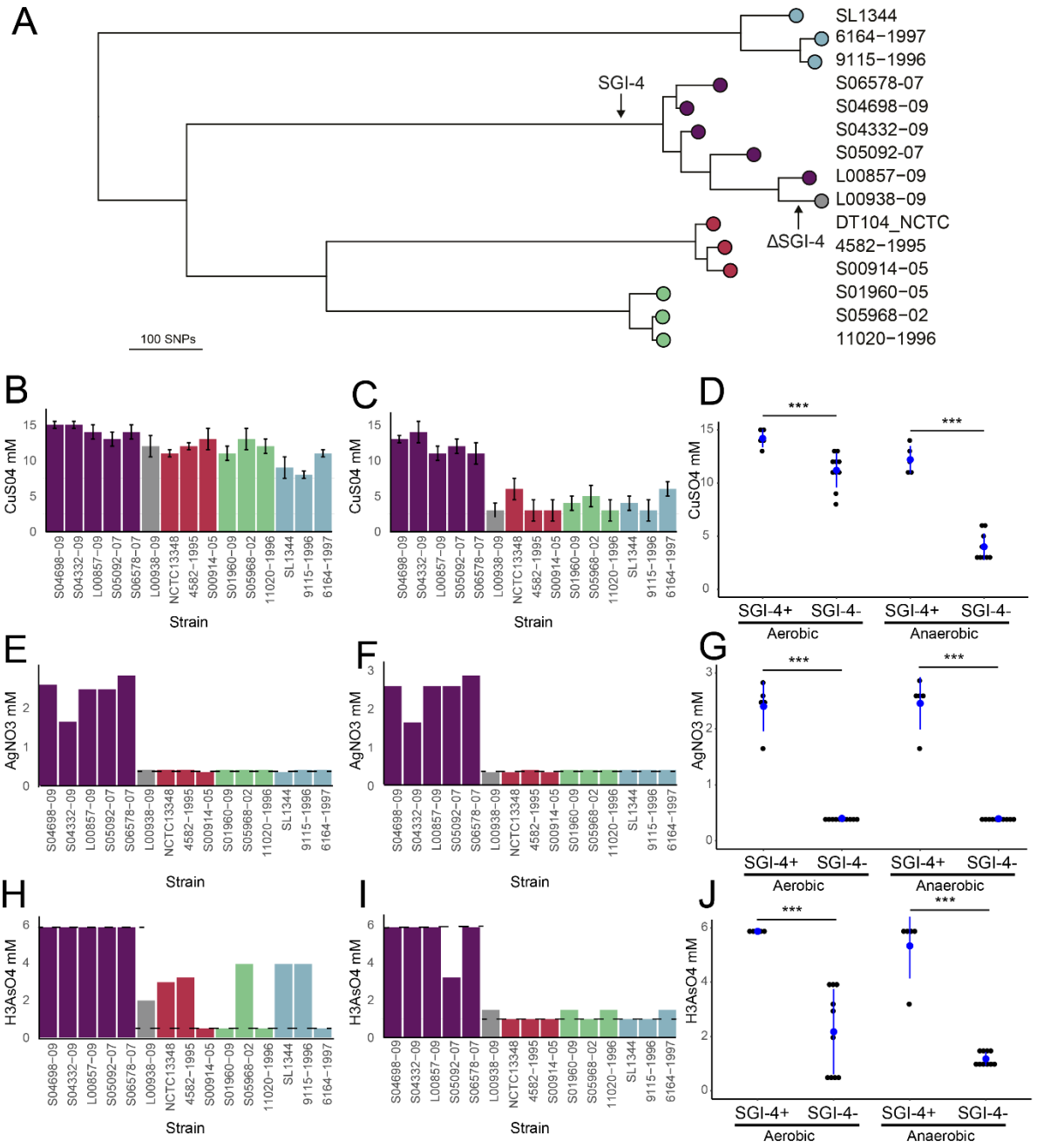


Figure III.6 Minimum inhibitory concentrations of *S.* Typhimurium strains and monophasic *S.* Typhimurium ST34 strains challenged with different concentrations of heavy metal compounds. (A) Phylogenetic relationship of the strains tested (n=15). Blue indicates strains from the broad-host-range DT204 lineage, purple indicates strains from the broad-host-range monophasic ST34 lineage, red indicates strains from the broad-host-range DT104 lineage α15, and green indicates strains from pig adapted U288 lineage. (B) Copper sulphate minimum inhibitory concentrations (MICs) tested in atmospheric conditions with lineage colour coded strains as in (A). (C) Copper sulphate MICs in anaerobic conditions. (D) Copper sulphate MIC comparisons when undertaken in either atmospheric or

anaerobic conditions with strains grouped by SGI-4 presence or absence. (E) Silver nitrate MICs in atmospheric conditions. (F) Silver nitrate MICs in anaerobic conditions. (G) Silver nitrate MIC comparisons when undertaken in either atmospheric or anaerobic conditions with strains grouped by SGI-4 presence or absence. (H) Arsenic acid MICs in atmospheric conditions. (I) Arsenic acid MICs in anaerobic conditions. (J) Arsenic acid MIC comparisons when undertaken in either atmospheric or anaerobic conditions with strains grouped by SGI-4 presence or absence. Horizontal dashed lines indicate the limit of detection. Copper sulphate MICs were undertaken by Priscilla Branchu. Mean values of (D), (G), and (J) are shown (blue dots), as well as the standard error of the mean (blue vertical lines).

III.7 Chapter III Discussion

SGI-4 was discovered to encode putative copper, silver, and arsenic resistance loci, as well as genes for integration, excision, chromosome partitioning, and a probable T4SS. This led to the hypotheses that it was transferable among *S.* Typhimurium strains, and that monophasic *S.* Typhimurium lineage strains would have significantly increased copper, silver, and arsenic resistance. Transfer was observed at low frequency, but was significantly increased in anaerobic conditions and further increased with DNA SOS response inducing mitomycin C. This is consistent with known regulators of ICE transfer, and consistent with the 'selfish island' hypothesis - that ICEs will escape a potentially dying or damaged host cells by upregulating their own excision and transfer through encoding regulatory proteins that interact with the host DNA SOS response machinery (515, 516). Genes encoding copper resistance included an RND-type copper efflux pump, of which similar encoding genes have been shown to translate to machinery that effectively pumps copper from the cytoplasm to the external milieu (504). The presence of chromosome encoded *cus* genes in *E. coli* but absence from *S. enterica* is consistent with the region having been lost during the evolution of the *Salmonella* genus. A remnant of *cus* genes also exists in *S. enterica* strains, supporting this hypothesis (517). The *cus* genes of *E. coli* were phylogenetically distinct from those on MGEs which were first termed *sil* (513). Copper resistance was observed in both atmospheric and anaerobic conditions. The level of cell death in strains without SGI-4 was much lower in the presence of copper in anaerobic conditions compared with aerobic conditions, consistent with

previous reports in the literature that the efflux pump genes are expressed in the anaerobic environment (518). Specifically, the *cus* efflux pump is thought to protect iron-sulphur clusters of dehydratases from anaerobic copper toxicity (519, 520). Despite its toxicity, copper is a biologically important heavy metal that forms the active site of numerous enzymes (521, 522). For example, SodCI is a copper and zinc dependant enzyme that enables *S.* Typhimurium macrophage survival by quelling the harsh redox potential of the rogue electron of superoxides through catalysing superoxide dismutation into oxygen and hydrogen peroxide (523). The copper efflux pump has also been shown to pump silver ions, consistent with it contributing to the observed silver resistance. Other genes for silver resistance are encoded on SGI-4; a silver specific, inner membrane, P-type, ATP dependant efflux pump encoded by *silP*, and silver molecular sponge *silE* (507). SilP is thought to pump silver ions to the periplasm, and SilE thought to act as a periplasmic molecular sponge (524). *Salmonellae* typically encode silver resistance on plasmids or MGEs, as there are no known specific chromosomal silver resistance determinants, possibly explaining the observed aerobic silver resistance in SGI-4 positive strains. SGI-4 negative *S.* Typhimurium encodes GolT, CopA, CueO and CueP, all of which have copper resistance activity in aerobic conditions, consistent with the comparatively higher aerobic copper sulphate MICs than silver nitrate MICs for strains without SGI-4 (518, 525). Silver toxicity is thought to occur through silver binding to the membrane and causing proton leakage through ion channel destabilization (505, 526), although the exact mechanism is disputed (526). The arsenic encoding genes have sequence similarity to those which have previously been shown to encode: arsenate to arsenite reductase catalysing protein (ArsC) (527, 528); arsenical P-type ATP dependant arsenical efflux pump (ArsAB) (529); arsenite metallochaperone (ArsD) (530); and; arsenic locus repressor protein (ArsR) (531). Presence of these genes is consistent with them functioning to provide the observed arsenate resistance. The upper limit of detection for the MICs in this study was 6 mM as this was the highest concentration the solution provided would allow and the lower limit 0.5 mM. A study investigating the arsenic resistance in strains of monophasic *S.* Typhimurium ST34 isolated in Japan discovered that SGI-4 encoding strains had an upper limit of detection of 64 mM and strains without SGI-4 had a lower limit of 0.25mM (532). This is consistent with the arsenic efflux and detoxification system having a considerable effect on arsenic resistance.

Seven previously unidentified sequences with similarity to SGI-4 were identified and a new family of ICEs termed SLEs was proposed. SLEs were frequently associated with heavy metal resistance genes. The only other *Salmonella* SLE host was that of an isolate of *Salmonella enterica* serovar Cubana (*S.* Cubana). *S.* Cubana has been implicated in outbreaks of disease in Swedish pigs, consistent with the hypothesis that SLEs and SGI-4 are countering selection pressures generated through pig husbandry practises due to the use of copper in pig farming in Europe (533).

Chapter IV: Bacteriophage Sensitivity Within the *S.* Typhimurium Population

The selection pressures that drive emergence of different *S.* Typhimurium clones remains an open question. This is especially true for successful, broad-host-range clones that appear to replace previously dominant clones with similar characteristics. How *S.* Typhimurium persists in the face of predation from the abundance of environmental phage is not well understood, but may play an important role in succession. To investigate the changes in phage sensitivity along lineages, and whether this might be impacting population structure and genomic evolution, the Anderson phage typing scheme data from PHE was utilised as a phage sensitivity phenotypic matrix (165). The Anderson phage typing scheme currently uses 30 phage preparations for typing. If resistance to all initial 30 is observed an additional 8 phage preparations are used. Phage preparations are well curated and historically stable, therefore providing an opportunity to investigate variation and evolution of phage resistance within the bacterial populations.

The central hypothesis tested in this chapter was that phage resistance varies in the population structure of *S.* Typhimurium and this is related to adaptation to circulation in distinct host populations and clonal expansion of epidemic strains. A phenomenon of phage differences being related to presence of *Vibrio cholera* strains of altered phage sensitivity in water and this coinciding with seasonal outbreaks of cholera has been reported, suggesting phage play a role in shaping other bacterial pathogens (572). Here, phage types were used to assess phage sensitivity differences in the *S.* Typhimurium population, an aspect not eluded to in the previous study (39). Phage type variation in the *S.* Typhimurium population was assessed through identifying the majority type in second order clades that represent clonal groups with distinct epidemiology and host range, in a collection of 1,407 strain's WGS from one year of human infections. To understand how differences in phage type correspond to potential phage resistance differences a quantifiable measure of phage resistance was developed, the resistance index (*Ri*). To address the question of whether *S.* Typhimurium has variable phage resistances to account for

potential predation from new phages, *Ri* changes and resistance patterns were analysed within each lineage. Evidence that phage play a role in shaping dominant livestock-associated lineages of *S.* Typhimurium was investigated through analysis of *Ri* differences of first order clades of *S.* Typhimurium, α and β. Clades α and β contain lineages with distinct host ranges. Clade α comprises epidemic, pandemic, and livestock-associated lineages, and clade β predominantly host adapted lineages (Bawn et al, 2020), with the exception of the DT204 lineage, which was widely disseminated (169, 534). Association of phage sensitivity with succession and clonal expansion of epidemic strains was tested by comparing *Ri* with the four consecutive, most dominant clones from 1965 to present. To address whether there are clade specific large genomic rearrangements, and if these were associated with differences in phage resistance, the prophage and genomic island flux in the population was investigated by comparative genomics of complete and closed genome sequence of one representative strain from each of the 11 third order clades. Furthermore, to address the question of what the population's ancestral phage sensitivity was, ancestral *Ri* was estimated through continuous data ancestral state estimation.

IV.1 DT193 that is characterised by resistance to all standard phage preparations is the most frequent majority phage type

To gain understanding of the phage type variation in *S.* Typhimurium and identify clade majority phage types, the most frequent phage types within biologically relevant, distinct lineages were assessed (Figure IV.1, below). A phylogenetic tree of 1,407 strains from one year of human infections in the UK was reconstructed and second order clades defined through hierarchical clustering with minor manual curation (Figure IV.1 9A). The second order clades contained clusters of strains and phage types associated with clades that corresponded with well characterised epidemic clonal groups, including the DT104 and the monophasic S. Typhimurium ST34 epidemics (184, 535, 536). The phage types of each isolate within 38 second order clade were then randomly sampled 1000 times. Sampling phage types in this manner identified each clade's majority phage type while adjusting for the different size of clades. The frequency of each majority phage type within its respective clade was also determined (Figure IV.1 (B)). Twenty-four majority phage types were determined within 38 distinct second order clades. Second order clades represented biologically important lineages with distinct host variation (Figure IV.1 (A)).

The most frequently sampled phage type from within clades was DT193 (Figure IV.1 (C)). Second to DT193 was DT135, which was most frequently sampled 3 times. All other frequent phage types were the most frequent in one or two clades. DT193 strains are resistant to all 30 standard phage preparations in the Anderson scheme. That DT193 is the most frequent majority phage type from multiple, distinct second order clades is consistent with phage resistance being selected for many times throughout the *S.* Typhimurium population. DT135 was the second most frequent cladal majority type, being most frequent in three clades. DT135 is lysed by all typing phages except three.

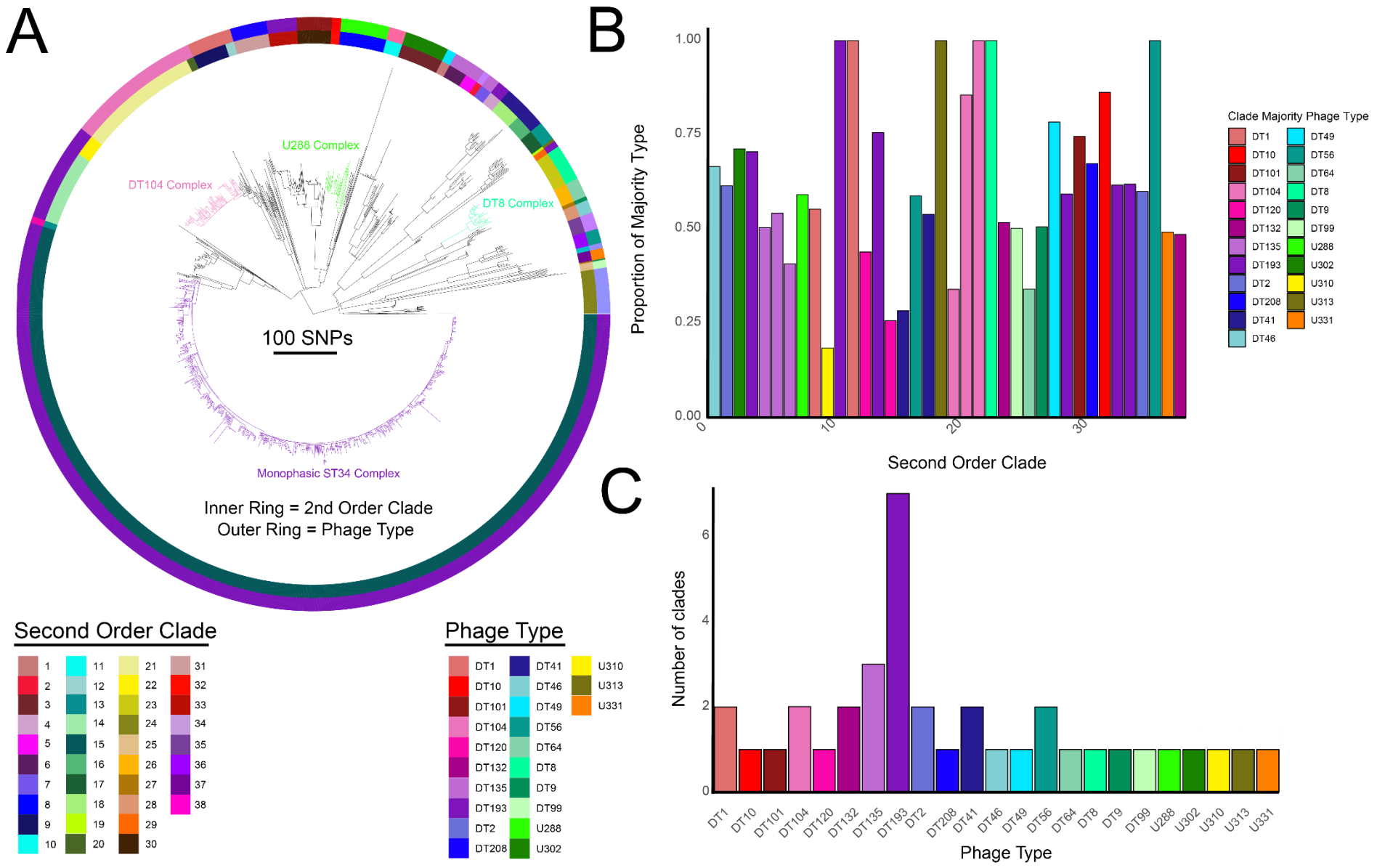

A
U288 Complex
DT104 Complex
DT8 Complex
100 SNPs
Monophasic ST34 Complex
Inner Ring = 2nd Order Clade
Outer Ring = Phage Type
Second Order Clade
1 2 3 4 5 6 7 8 9 10
11 12 13 14 15 16 17 18 19 20
21 22 23 24 25 26 27 28 29 30
31 32 33 34 35 36 37 38
Phage Type
DT1 DT10 DT101 DT104 DT120 DT132 DT135 DT193 DT2 DT208
DT41 DT46 DT49 DT56 DT64 DT8 DT9 DT99 U288 U302
U310 U313 U331
B
Proportion of Majority Type
1.00
0.75
0.50
0.25
0.00
0
10
20
30
Second Order Clade
Clade Majority Phage Type
DT1 DT10 DT101 DT104 DT120 DT132 DT135 DT193 DT2 DT208 DT41 DT46
DT49 DT56 DT64 DT8 DT9 DT99 U288 U302 U310 U313 U331
C
Number of clades
6
4
2
0
DT1 DT10 DT101 DT104 DT120 DT132 DT135 DT193 DT2 DT208 DT41 DT46 DT49 DT56 DT64 DT8 DT9 DT99 U288 U302 U310 U313 U331
Phage Type

Figure IV.1 Major phage types within second order clades from one year of human gastrointestinal infections, United Kingdom, 2014-2015. (A) Maximum-likelihood phylogenetic tree constructed using 16,681 core genome SNP sites from the whole genome sequence of 1,407 *S.* Typhimurium isolates. (A) Second order clade designations as determined by hierBAPS (inner ring), with corresponding colours in the legend. (A, outer ring) Cladal majority phage type as determined through sampling 1000 times from a list of the phage types of each isolate within the clade. (B) Proportion of strains in each clade with respective majority phage type (1-38). The colour of each bar indicates majority phage type as in (A). (C) Number of clades for which each phage type was the majority type within the 38 clades with colours as in (A). Four lineages have been highlighted on their respective branches. Clades corresponding to the DT104 complex (pink) and monophasic ST34 complex (purple) that are two dominant epidemic groups in the UK since 1990, and a duck associated DT8 complex (bright turquoise) and pig-associated U288 complex (bright green) are indicated.

IV.2 Resistance index quantifies phage resistance using the data from the Anderson phage typing scheme

Since phage type is defined by sensitivity to lysis by a series of 30 phage preparations, the phage type contains information about the level of sensitivity to certain phages. A value termed Resistance index (*Ri*) for an strain's phage resistance was calculated by transforming the proportion of phage preparations to which each strain of a phage type was resistant, such that;

Ri represents a strain's phage resistance index determined from the Anderson typing scheme, where *p*

$$Ri = 1 - \left(\frac{1}{p}\right) \times \left(\sum_{0}^{p} s_i = s_{0+1} + s_{0+2} \ldots s_{p-1} + s_p\right)$$

represents the total number of phage preparations, and *Si* is the summation of outcomes of either resistance (1) or lysis (0).

IV.3 Phage resistance is variable among *S.* Typhimurium lineages circulating within distinct hosts

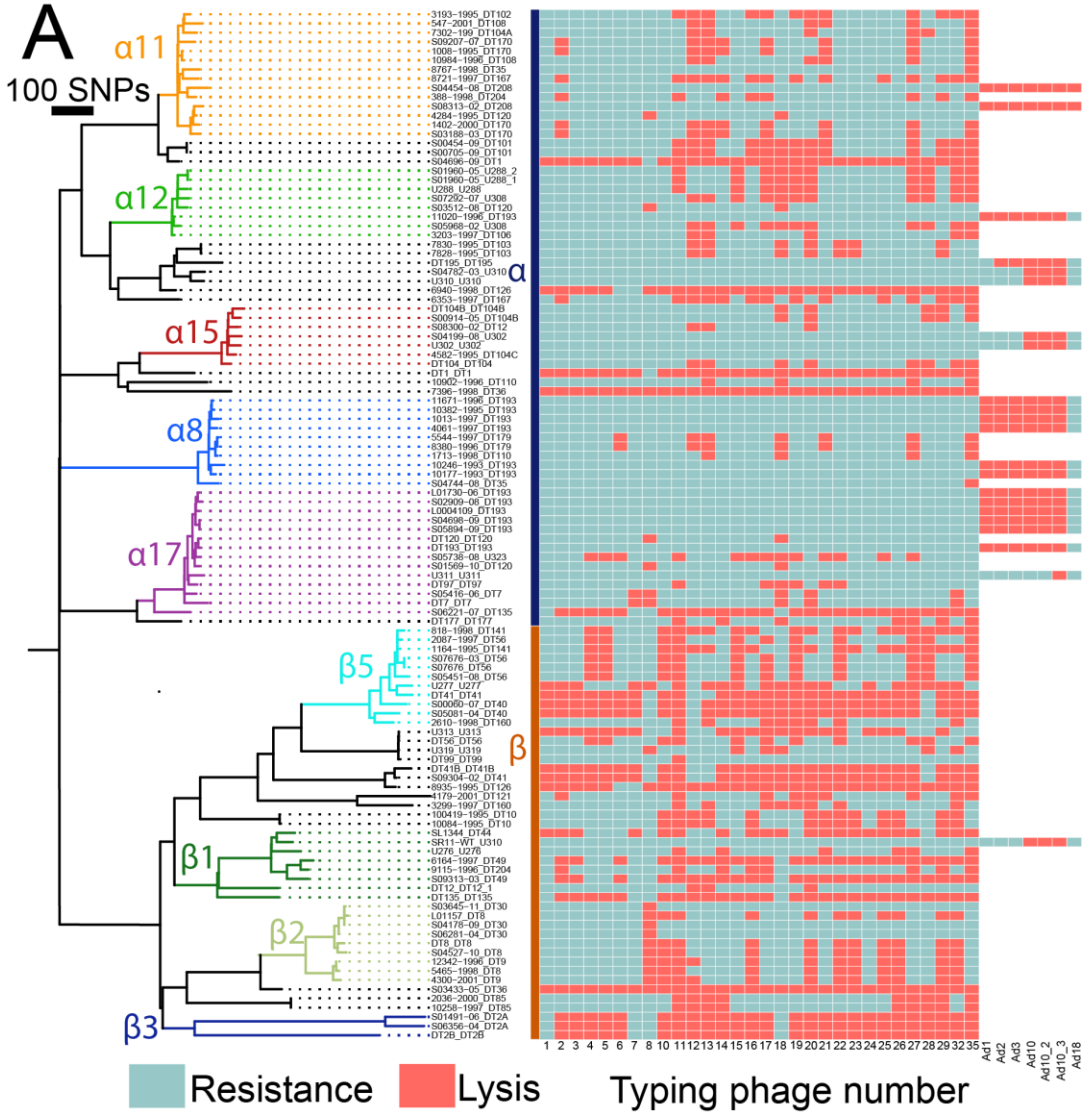
To address whether lineages that circulate within distinct hosts require varying phage resistance, the lysis pattern of 112 isolates representing the population diversity of *S.* Typhimurium in the UK were

assessed for differences in *Ri.* The population structure of this strain collection was defined in a previous study (39). Two first order clades were defined in the population, and designated α and β. These two clades had distinct characteristics. Strains from clade α were characteristically from agricultural and epidemic lineages, while clade β contained strains mostly associated with wild avian reservoirs. Each third order clade was associated with distinct hosts and host ranges, for example, clade α contained two lineages highly associated with pigs (DT170 lineage α11 and U288 lineage α12) and 2 broad-host-range epidemic lineages (DT104 lineage α15 and monophasic ST34 lineage α17)(169, 184, 202, 204, 221{Petrovska, 2016 #6085, 535, 537-541). Clade β contained at least three lineages found mostly within specific hosts, one host associated (duck associated DT8 lineage β2), and two host restricted (passerine bird restricted DT56 lineage β5 and pigeon adapted DT2, lineage β3)(77, 185, 542, 543). Not all third order clades were analysed in this study as some defined clades had no phage typing data (human iNTS ST313, lineage β7) and others were not defined in the literature, making any association of phage resistance to distinct hosts or host ranges difficult.

A maximum likelihood phylogenetic tree was reconstructed using a similar approach to Bawn et al., (39), but excluding strains with no phage typing data. Next, Anderson phage typing scheme phage preparation lysis profiles were associated with each isolate (Figure IV.2 (A))(39). The *Ri* for each clade was calculated and analysed for differences between lineages (Figure IV.2 (B)). Clades α11 and α12 are associated with pigs (39, 169, 537, 538), and displayed high mean cladal *Ri* (0.76 and 0.775, respectively). Clades β3 and β5 had comparatively lower mean *Ri* values that those from clade α (0.133, and 0.3833, respectively), except for duck adapted β2 which had a reduced *Ri* (0.7333) but had an *Ri* value closer to those from clades α11 and α12 (0.76 and 0.775 respectively). These represent host adapted lineages which circulate mostly in pigeons, passerine birds and ducks respectively (77, 185, 187, 542, 543). β2 has a majority type of DT8, and a significant change of 10 phage preparations resisted in the clade's second most common phage type, DT30.

Clades β1, α17, and α15 are composed of strains with a broad host range and strains of these lineages were successively the most frequently isolated from 1975 to the present day, with successive clonal expansion and replacement every ~15 years (107, 169, 184, 204, 219, 539, 544). Clade β1 has the largest

Ri range, consistent with the clade having a wide phage resistance range. Both α17 and α15 display high cladal mean *Ri* (0.929 and 0.864 respectively). Strains from these lineages have been isolated from a wide range of livestock species.


A
100 SNPs
α11
α12
α15
α8
α17
β5
β1
β2
β3
α
β
Resistance
Lysis
Typing phage number

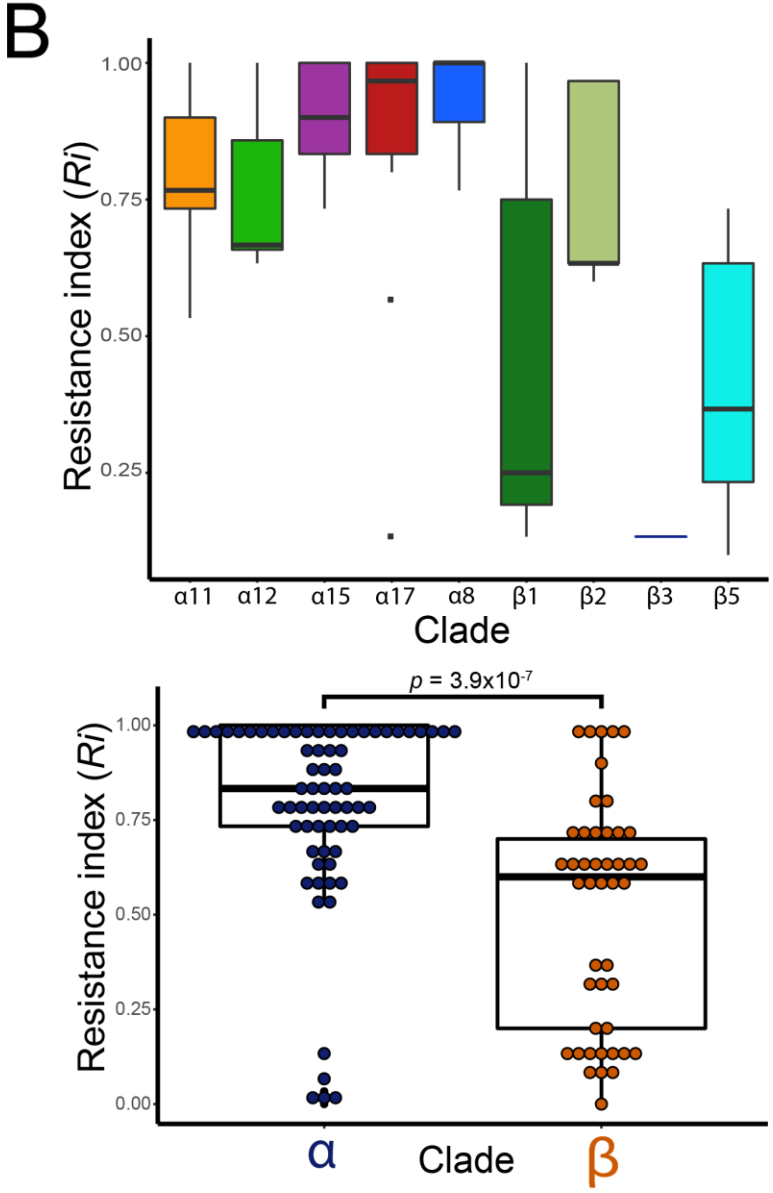

B
Resistance index (Ri)
1.00
0.75
0.50
0.25
α11
α12
α15
α17
α8
β1
β2
β3
β5
Clade
p = 3.9x10-7
0.00
α
β

Figure IV.2 Variation in phage resistance in representative strains of *S.* Typhimurium. (A) Maximum-likelihood tree based on 17,823 recombination-purged core genome SNPs from 112 isolates of *S.* Typhimurium from various animal sources. Each coloured clade represents an *S.* Typhimurium lineage. Sensitivity to phage preparations in the Anderson typing scheme are indicated in the heatmap as denoted in the key. (B) Phage resistance index (*Ri*) of *S.* Typhimurium population representative strains are displayed with range (whiskers), interquartile range (hinges), and, median (black horizontal line) for each clade. Bars are colour coordinated with lineages in (A). (C) *Ri* of *S.* Typhimurium population grouped by first order clades.

IV.4 Livestock associated lineages have increased phage resistance over wild-avian-adapted lineages

The two first order clades (α and β) in the population represent groups of *S.* Typhimurium with distinct host ranges. Clade α is associated with livestock and broad-host-range lineages, and β mostly with wild avian species. To address whether there is any evidence of differences in phage selection pressure at this level of the population, the two clades were grouped and assessed for significant differences in *Ri* (Figure IV.1 (C)). Clade α had a significantly higher *Ri* than clade β, consistent with strains from clade α being under more significant phage selection pressure.

IV.5 Successive most frequently isolated pandemic and epidemic clones exhibit a stepwise increase in their phage resistance potential

The reasons for successive clonal expansion and replacement of dominant phage types of *S.* Typhimurium with a periodicity of approximately 15 years is not known (169). To investigate whether this was associated with phage sensitivity changes, the *Ri* of the majority phage type from each lineage was assessed against a timeline of isolation frequencies (Figure IV.3). There was a general trend of increase in *Ri* of each successive, dominant clone, with each exhibiting similar or increased phage resistance.

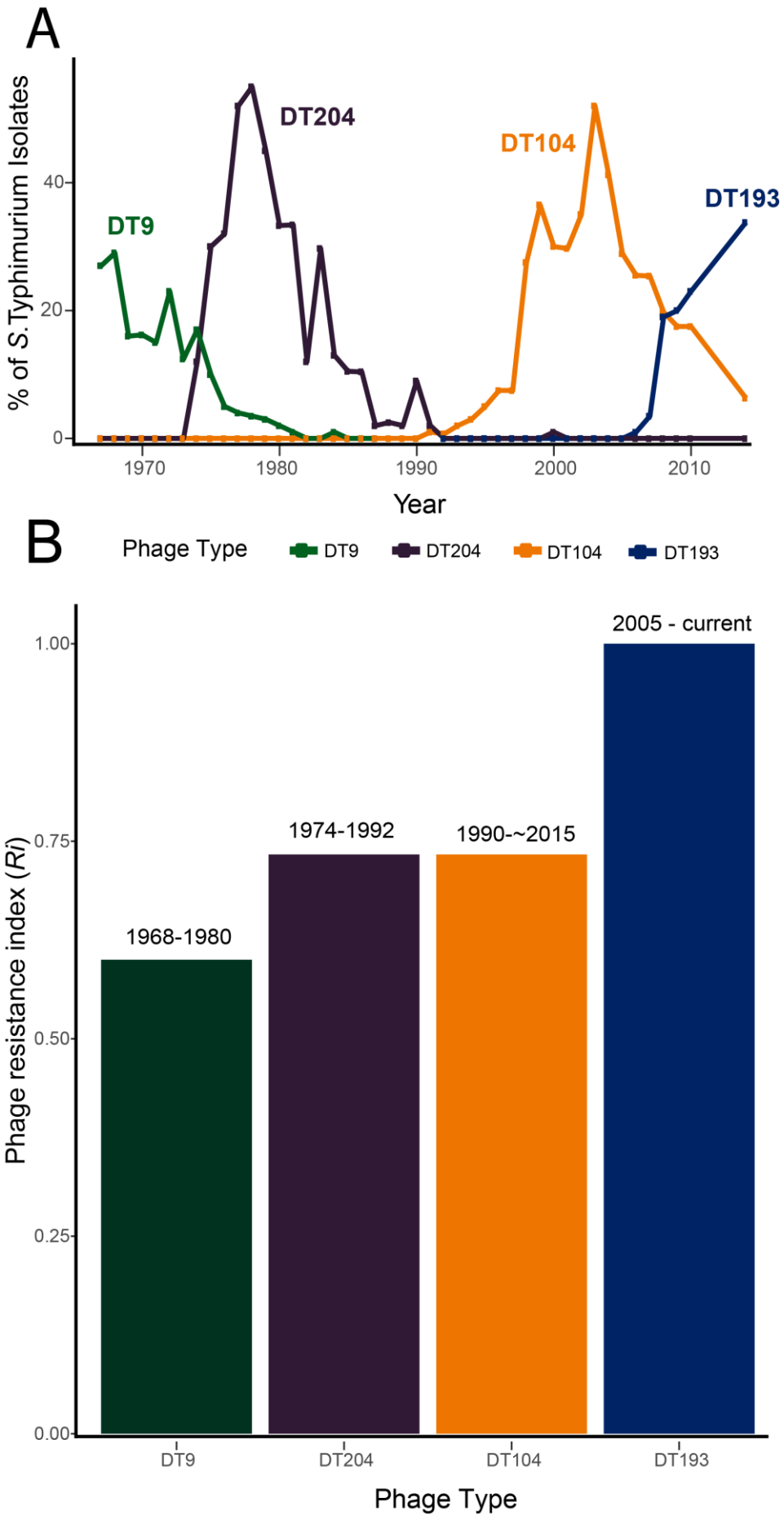

A
% of S.Typhimurium Isolates
40
20
0
DT9
DT204
DT104
DT193
1970
1980
1990
2000
2010
Year
B
Phage Type
DT9
DT204
DT104
DT193
1.00
0.75
0.50
0.25
0.00
Phage resistance index (Ri)
1968-1980
1974-1992
1990-~2015
2005 - current
DT9
DT204
DT104
DT193
Phage Type

Figure IV.3 Phage resistance of dominant clonal groups of *S.* Typhimurium. (A) Frequency of major epidemic clones identified by phage type from 1965-2018. (B) Phage resistance index (*Ri*) of the most common phage types from each lineage.

IV.6 Variation in large genomic features is evident in each third order lineage

Prophages frequently encode phage exclusion mechanisms such as membrane-bound superinfection exclusion proteins, toxin-antitoxin (TA) systems, mechanisms to evade host self- and non-self-identification through encoded DNA modification enzymes, and phage encoded CRISPR-cas, enzymes that modify phage receptors such as acetyl transferases, amongst others (545-548). These features of prophages contribute significantly to the phage resistance of bacteria. Clade α had significantly higher *Ri* over clade β, consistent with increased phage resistance. To investigate whether increased phage resistance was linked to prophage variability, 11 long read WGS were generated using Pacific Biosciences SMRT® technology (39). The long reads enabled assembly of complete and closed WGS, necessary to unambiguously determine prophages. The genomes were arranged into the same orientation of reference strain SL1344 and aligned using ACT (447). Prophage regions were determined by searching the PHASTER database using the PHASTER web search tool, which has proven to be robust for finding prophages (549). Recombination within prophages was inferred from the prophage sequence analysis undertaken by Bawn et al,. (39), and where recombination events occurred during evolution of the lineages assessed through associating the specific recombination events with the phylogenetic tree structure. A total of 83 prophages were discovered in 11 reference genomes encompassing 18 annotated prophage and 12 variably occupied insertion sites (Figure VI.4). Recombination occurred mostly within prophages Gifsy-1 and ST64B. All isolates harboured prophages Gifsy-1, Fels-1, Sal3, and a remnant of Bcep-Mu, the common presence of which is consistent with them shaping fitness of all S. Typhimurium lineages and being under positive selection. Reference strains from clade α had ten prophages not observed in clade β, and clade β had nine prophages not observed in clade α. Using a ratio of prophages observed only within one clade to number of strains from each clade accounts for the different number of reference strains from each clade (four strains from clade α, and

seven from clade β). Clade α has a ratio of strains to unique prophages of 1:2.5, and clade β 1:1.29. This is consistent with clade α having a higher prophage variation than clade β.

Two genomic islands were present in separate loci unoccupied by prophages in other strains investigated. These islands were SGI-1, the multidrug resistance IME common in pandemic DT104 (Figure IV.4, NCTC13384, green), suggesting that this insertion site was specific for this IME and not used by prophage, and SGI-4 (Figure IV.4, S04698-09, green) of monophasic *S.* Typhimurium ST34 described previously in Chapter III of this study (184, 206).

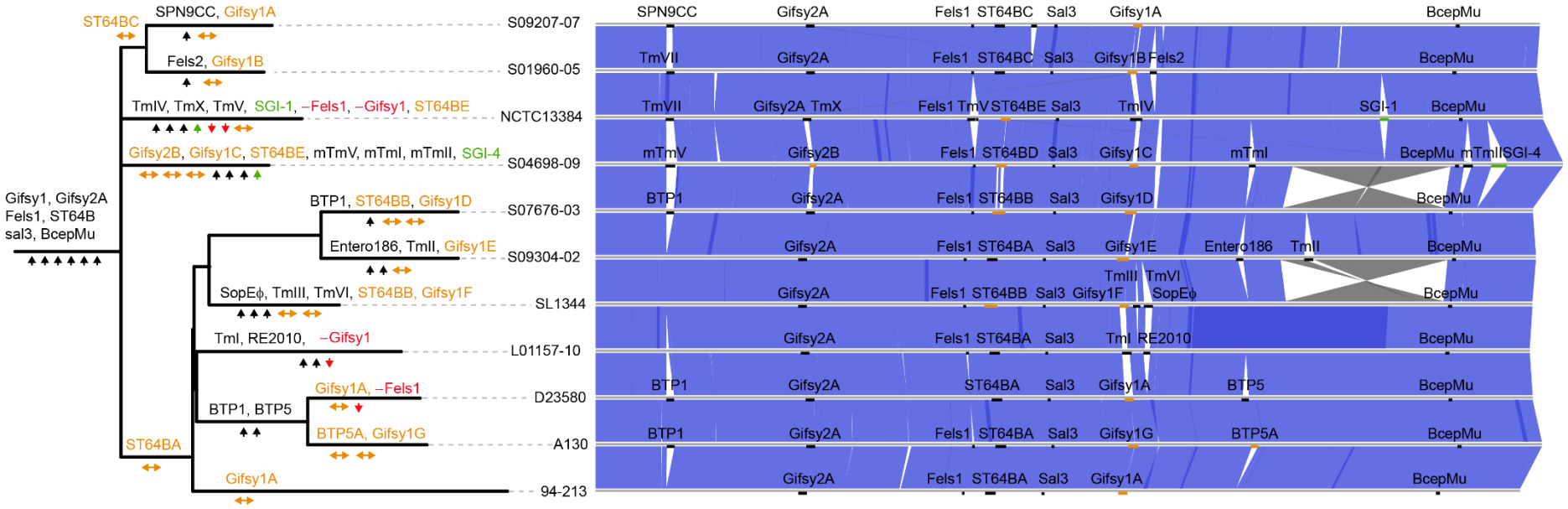


Figure IV.4. Chromosome sequence alignments of representative strains of *S.* Typhimurium. Large genome rearrangements and MGE flux in 11 reference strains representing each clade of the population structure of *S.* Typhimurium. Each grey horizontal line represents each genome, and vertical lines represent regions of >90 % sequence identity (blue) or reverse complement (grey). Maximum-likelihood tree (left), indicating gain of a prophage (black arrows), recombination occurring within a prophage sequence (orange), loss of a prophage (red) and insertion of genomic islands (green). S09207-07 is representative of DT170 majority, agricultural lineage α 11; S01960-09 of pig adapted, U288 majority type lineage α 12; NCTC13384 of pandemic, DT104 majority type lineage α 15; S04698-09 of pandemic, monophasic ST34 lineage α 17; S07676-03 of passerine bird adapted, DT56 majority type lineage β 5; S09304-02 of an avian adapted lineage with majority type DT41; SL1344 of epidemic, DT204 majority type lineage β 1; L01157-10 of duck associated, DT8 majority type lineage β 2, D23580 of human invasive ST313 lineage 2, A130 of ST313 lineage 1, and 94-213 of pigeon restricted, majority type DT2 lineage β 3.

Adapted from Bawn et al., (39), notably through re-analysing the data and adding an estimation of recombination events within prophages at locations in the phylogenetic tree.

IV.7 Ancestral history estimates are difficult to discern at this level of the population

To investigate the trajectory of changes in phage resistance during the evolution of *S.* Typhimurium, a hypothetical estimate of the *Ri* of the common ancestor was calculated using continuous trait ancestral history estimation in 112 representative strains (Figure IV.5). Ancestral state reconstruction methods typically include discrete trait analyses via parsimony, maximum likelihood including maximum posterior probability estimates, or continuous-time Markov chain models with Monte Carlo sampling (MCMC) (550). *Ri* is a continuous trait. To identify if this trait can infer any meaningful hypotheses about the common ancestor of *S.* Typhimurium, continuous trait ancestral state reconstruction was used. Brownian motion-based continuous trait ancestral estimation analyses are common and robust (550), with multi-variate Brownian motion models removing the assumption of neutral-rate evolution across all branches (551). Continuous Brownian evolution is undertaken by drawing randomly from a normal distribution within a defined boundary based on the value of the previous point on the tree, here starting at the tip values (*Ri*) and drawing values iteratively form point to point backward to the root. A distinct aspect of these methods is that they are for exploring and generating hypotheses, not testing them (552). Extrapolation back toward the common ancestor of clade α and β using an assumption of neutral evolution suggested the ancestor had a mid-range *Ri* value (0.48) (Figure IV.5 (A) and (B)). Multi-variate Brownian motion analyses had no significant increase in log-likelihood as determined by the LRT ($p > 0.9$) but had a minor change in the calculated root node *Ri* (0.466). Ancestral reconstruction analyses estimate that the common ancestor of *S.* Typhimurium had a lower-mid range *Ri*, consistent with mid-level phage sensitivity. However, the 95 % confidence intervals for ancestral *Ri* values become either 0 or 1 when extrapolating ~50 SNPs back from each tip. This poor resolution for possible ancestral *Ri* values is likely due to the wide variation observed in *Ri* values in different lineages. Analysis of biologically distinct clades may provide better insight into phage resistance dynamics.

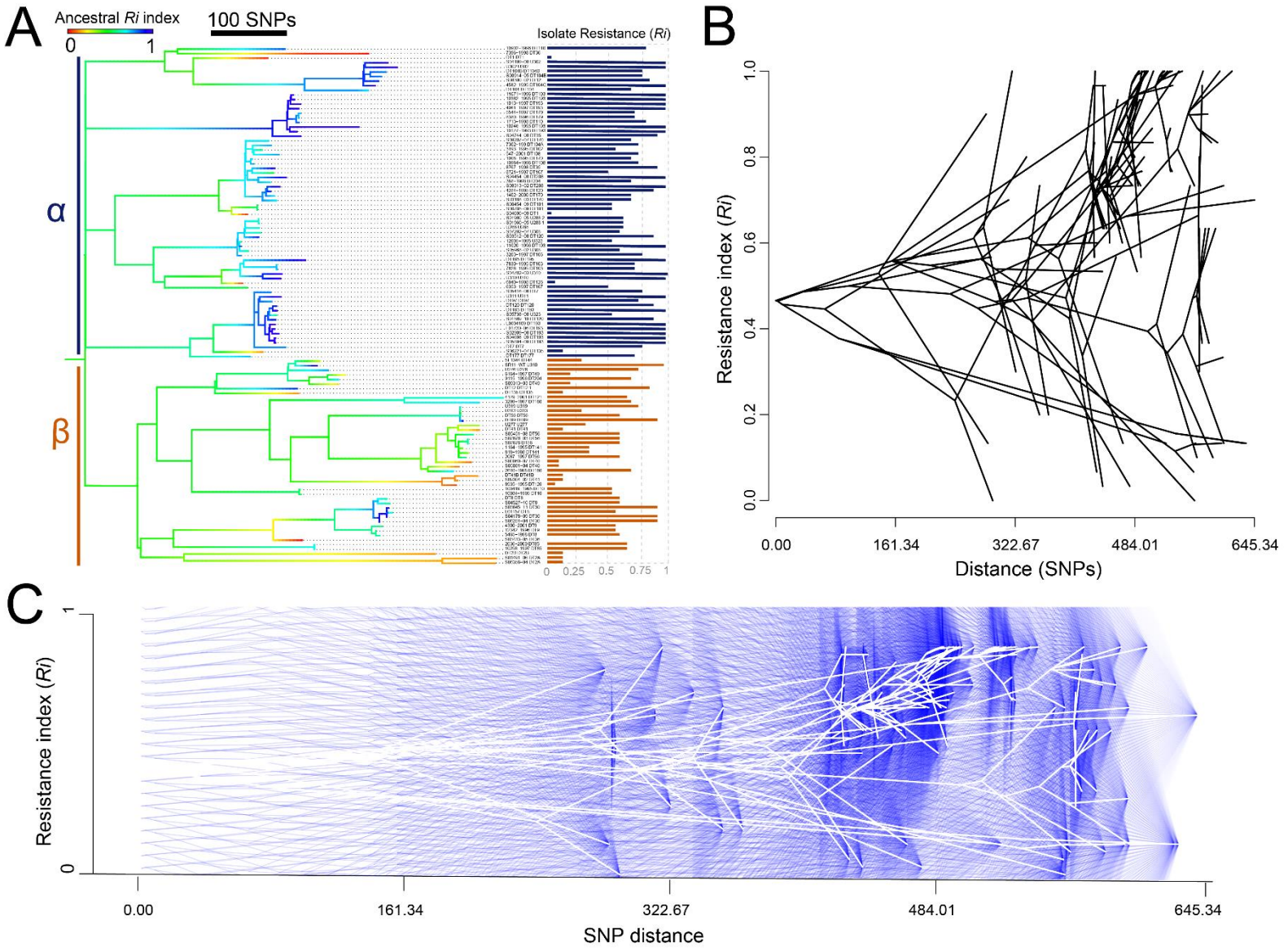

A
Ancestral Ri index
0
1
100 SNPs
Isolate Resistance (Ri)
α
β
0
0.25
0.5
0.75
1
B
Resistance index (Ri)
0.0
0.2
0.4
0.6
0.8
1.0
0.00
161.34
322.67
484.01
645.34
Distance (SNPs)
C
Resistance index (Ri)
1
0
0.00
161.34
322.67
484.01
645.34
SNP distance

Figure IV.5 Continuous ancestral state reconstruction based on phage-type-generated resistance index values. (A) Hypothetical ancestral phage sensitivities as determined by using univariate, Brownian-motion-based ancestral state reconstruction. Horizontal bars at each tip display each strain's *Ri* value for clade α (blue) and clade β (orange). (B) Ancestral history data from (A) in a phenogram, where each tip of the tree in (A) has a corresponding point at the end of each black line. (C) Ninety-five percent confidence intervals (blue lines) for sections of branches from the data in (B).

IV.8 Chapter IV discussion

An important caveat of using phage sensitivity data from the Anderson typing scheme to infer differences in strain phage sensitivity is that the question as to how predation by the 30 phage preparations in the Anderson phage typing scheme reflects predation by extant phages in the host environment is unknown. In fact, we know very little about the phages that prey on *Salmonella* during the evolutionary history of *S.* Typhimurium. Nonetheless, there is no evidence of changes in phage predation over the timescale of the evolution of this pathogen and no data on phage sensitivity of *S.* Typhimurium to these phage preparations represents the most extensive such dataset for any bacterium. What we do know is that they are 30 different phage preparations, of multiple phage (172). This study investigated phage sensitivity of collections of genotypically diverse strains of *S.* Typhimurium. DT193, a phage type encompassing resistance to all the standard phage preparations, was the most frequently observed phage type within the population. DT193 was the majority type in seven lineages, four more than any other phage type. This suggests that phage pressure from the abundance of phages in the biosphere is a significant selection for *S.* Typhimurium lineages and a resistant phenotype is selected for many times throughout the population.

To quantify phage sensitivity with respect to phage preparations of the Anderson typing scheme the phage resistance index was developed (*Ri*). Extrapolating differences in phage resistance from phage typing data generates various conclusions. One question addressed was whether *S.* Typhimurium circulating in distinct hosts require variable or increasing phage sensitivity to account for the threat of new phages in their respective niches. All lineages had variable phage resistance profiles and sensitivity, frequently with an increase in *Ri*, except for the pigeon adapted DT2 complex of β 3. The two porcine

associated lineages (α11 and α12) had high *Ri* suggesting significant evolutionary selection pressure from phages. Clades of broad-host-range lineages also had high phage resistance. Clade β1 comprises strains from a broad-host-range lineage that was most frequently isolated ~1975-1990 (169, 248). Clade β1 had the largest range of *Ri* values observed. Unlike the other 2 broad-host-range lineages this clade sits within mostly wild avian adapted clade β. This association with clade β is consistent with the lineage possibly originating from an avian host. This is corroborated with the observed phage resistance. Within an avian host the lineage is hypothesised to have a restricted biome and require less phage resistance. However, after transmission to cattle the effective biome size increases, which would hypothetically require resistance to phage predation for persistence. Becoming broad-host would extend the effective biome size more so, possibly accounting for the strains observed with the highest *Ri* values. Duck associated β2 had comparatively higher *Ri* than host adapted lineages of clade β. This is consistent with duck reservoir associated *S.* Typhimurium requiring moderate phage resistance to persist in the lineage, and a rapid switch to phage resistant DT30, consistent with the idea that new phages may be more readily introduced into the niche than for the other two wild avian adapted lineages. Perhaps this is due to ducks being both water and air borne, unlike the other avian species from which strains in this study were isolated, potentially increasing their effective biome size and therefore increasing the diversity of phages that they are required to resist (Hall, 2021, personal communication). Avian restricted lineages from clade β had lower *Ri* values consistent with lower phage resistance. The much lower cladal mean *Ri* and variation of clades β3 and β5 is consistent with the biology of these lineages. This more concentrated lifestyle suggests that isolates from these lineages would need to resist phages within their respective biomes, but for host-restricted lineages the variation of phages encountered is potentially less than for agriculturally circulating lineages. Clade α had a significantly higher *Ri* than clade β. This was consistent with the prospect that evolutionary pressure to adapt and resist phages may be greater for broad-host-range and agriculturally circulating lineages than non-agricultural, host adapted and restricted lineages.

An increasing phage resistance trend was observed for successive dominantly isolated clones from 1965-2018. Whether this indicates that one clone can outcompete the other through a more proficient

ability to resist phage selection pressures, or that there is a limitation on the ability to resist all different phages within the biosphere causing the decline of numbers, is unknown. It is likely to be the latter, especially when considering the gap between the start of the decline of the DT204 lineage and the rise of the DT104 lineage. An ever-expanding *S.* Typhimurium clone will be generating selection pressure for phages to evolve the ability to lyse the clone. Since phages evolve faster than bacteria, the prospect that there is only a certain limit of phage defence that *S.* Typhimurium can evolve in a certain time-frame forms a likely scenario. In response to evolutionary pressure from a new *S.* Typhimurium clone in their niche, phages will be evolving and hypothetically gain lysis ability. This also raises a question for further investigation - do we see an increase in phage resistances within these lineages during their evolution?

The observed variation in prophages suggests that phages are fundamental to the success of lineages with distinct host ranges. This is likely due to genetic content which convey novel functions, such as the insertion of SopE encoding prophages in widely disseminated DT204 majority type lineage β1 (Figure IV.4, SL1344, SopEθ) and monophasic ST34 lineage α17 (Figure IV.4 S04698-09, mTmV). Novel prophages also require protection from other phages and frequently carry phage exclusion mechanisms. The difference in prophage repertoires will hypothetically be contributing to resistance to phage predation, and therefore phage type variation, corroborated by the presence of lineage specific prophages. The link between encoded prophages and phage resistance suggests that selection pressure from phages, as well as the ability to allow prophage variation through introduction of new prophages, are fundamental to the success of *S.* Typhimurium clones with distinct host-ranges. Reference strain 94-213 of the pigeon restricted DT2 lineage β3 appears to have less variation in prophage than all other strains analysed (Figure IV.4, bottom). Strain 94-213 encodes prophages which are observed within most strains of the other lineages, but no unique prophages. Parsimoniously, this is consistent with the lineage being more representative of the ancestral type *S.* Typhimurium, although undetermined.

Ancestral history estimation was conducted on the population representative isolates, with a probabilistic hypothesis of the ancestral *Ri* being lower-mid range, consistent with the common ancestor having a lower-mid phage resistance ability. However, this analysis was subject to considerable variation

when producing a distribution of possible ancestral histories, making the confidence intervals approach 1 or 0 before any ancestral nodes. This analysis suggests that a more in-depth investigation into isolates within a reduced temporal and SNP range would shed more light on the variation in phage resistance within specific lineages.

Two areas of interest were identified from these data for further investigation; genomic differences resulting in phage type variation of monophasic *S.* Typhimurium ST34, and differences resulting in the ability of DT8 duck associated strains to potentially resist phage predation, as observed through switches to DT30 within closely related strains.

Chapter V: Prophage mTmII Confers Phage Resistance and Acquisition was Associated with Continued Expansion of Monophasic *S.* Typhimurium ST34

Distinct *S.* Typhimurium clones appear in successive waves of dominant phage types, lasting for 10-15 years, with four main clones recorded since initial *Salmonella* surveillance and intra-serovar stratification through phage typing in the 1960s (169). The selection pressures that drive emergence of dominant phage types have not been well characterised. The role of phage predation in shaping bacterial populations is emerging as an area of research interest, in part driven by the availability of WGS (419). In Chapter IV the *S.* Typhimurium population phage sensitivities were determined to be variable. Ancestral history estimates displayed high confidence intervals when extrapolating >50 SNPs backward in evolution from each tip. It was reasoned that investigating distinct *S.* Typhimurium lineages within a range of 50 SNPs would shed more light into the phage sensitivity dynamics. Strains forming the monophasic ST34 lineage are <50 SNPs apart, and a current public and agricultural health concern. Routine WGS by public health agencies also presents an opportunity to investigate this lineage in unprecedented detail. Here, the population structure of monophasic *S.* Typhimurium ST34 strains in the UK were investigated for changes in sensitivity to phage predation using *Ri* derived from phage typing data to quantify phage resistance of strains (553). The central hypothesis was that the emergence of dominant, broad-host-range clones of *S.* Typhimurium is associated with increased resistance to phage predation to counter selection from phages within an extended biome. This Chapter describes how acquisition of a prophage, previously termed mTmII (39), is associated with an increase in phage resistance in the lineage and plays an important role in generating phage resistance.

V.1 Phage resistance is associated with increased clonal expansion in monophasic *S.* Typhimurium ST34

The monophasic ST34 lineage (clade α 17) is the most recently emerged broad-host-range lineage of *S.* Typhimurium and is currently dominant in the UK and other countries globally (168, 178, 184, 219, 222, 225, 536, 541). Due to recent emergence of the monophasic ST34 lineage and WGS becoming public health surveillance tool of choice, WGS data is abundant. This study aimed to investigate genomic elements associated with frequent sampling of closely related strains. Time adds an extra dimension when investigating the frequency of genomic regions and analysing closely related strains. Minimising time provides an understanding of whether particular lineages have clonally expanded, providing a 'snapshot'. To minimise temporal effects and assess which genomic features are associated with phage resistance and clonal expansion, 763 strains isolated in one year from human gastrointestinal infections in the UK were subject to maximum-likelihood phylogenetic reconstruction (Figure V.1 (A, right)). Phage typing scheme sensitivity data was then associated with each strain (Figure V.1 (A, left)). Nineteen phage types were dispersed throughout the phylogenetic tree, with most of the variation apparent toward the root node (Figure V.1 (A, tree label X and tree label Y)). The four most frequent phage types were U323, DT120, DT193, and U311. DT120 is frequent toward the root and appears in separate clades along with U323 (Figure V.1 (A, tree label X and tree label Y)). There is one clade of closely related strains mostly exhibiting U311 (Figure V.1 (A, tree label X)). The latter part of the tree encompasses a clade mostly of DT193, with 3 strains of other phage types. (Figure V.1 (A, tree label Z)). The phylogenetic tree structure is consistent with an increasing *Ri* as the lineage evolves, (U323 =0.567, DT120 = 0.967, DT193 = 1, U311 = 1, Figure V.1 (C)). To investigate how the most frequently observed phage types cluster on the phylogenetic tree, proximal strains of each phage type were summed, and the mean average determined (Figure V.1 (D)) U323 exhibiting strains cluster in fewer numbers (n=3-11). Strains phage typed DT120 have the second smallest clusters (n=3-31), U311 strains display 3 clusters larger than 10 strains, and DT193 has 3 clusters larger than 30 strains with clade Z having >200 strains. Larger clusters of phage types with increased *Ri* values is consistent with increased sampling of strains with increased phage resistance.

To investigate the relationship between phylogenetic relatedness and phage type, the *Ri* values of each strain were assessed against their pairwise core genome SNP distance (patristic distance) via linear

regression (Figure V.1 (B)). Reduced patristic distances in a group of isolates is consistent with a lineage being under positive selection, as we previously described (589). Linear regression of the relationship between *Ri* and patristic distance determined a weak negative correlation, consistent with phage resistance having an association with clonal expansion. However, linear model is hindered by many phage types having an *Ri* of 1 (Figure V.1 ((B)). To further investigate if repeated sampling of strains with increasing *Ri* values occurs, the 4 most frequently observed phage types were split into subtrees and patristic distance extracted (Figure V.1 (E)). Strains phage typed U311 had the highest *Ri* (1.2) but lowest mean patristic distance (3.67 SNPs). Strains phage typed DT193 had the second highest *Ri* value (1) and second lowest mean patristic distance (6.6 SNPs). Linear regression of the four majority phage types' *Ri* and patristic distances resulted in a greater negative constant than assessing all phage types together (Figure V.1 (E), all phage types = -1.18, four main types = -5.74). A negative correlation between *Ri* and patristic distance is consistent with an increase in phage resistance being associated with more frequent sampling of closely related isolates, possibly due to positive selection.

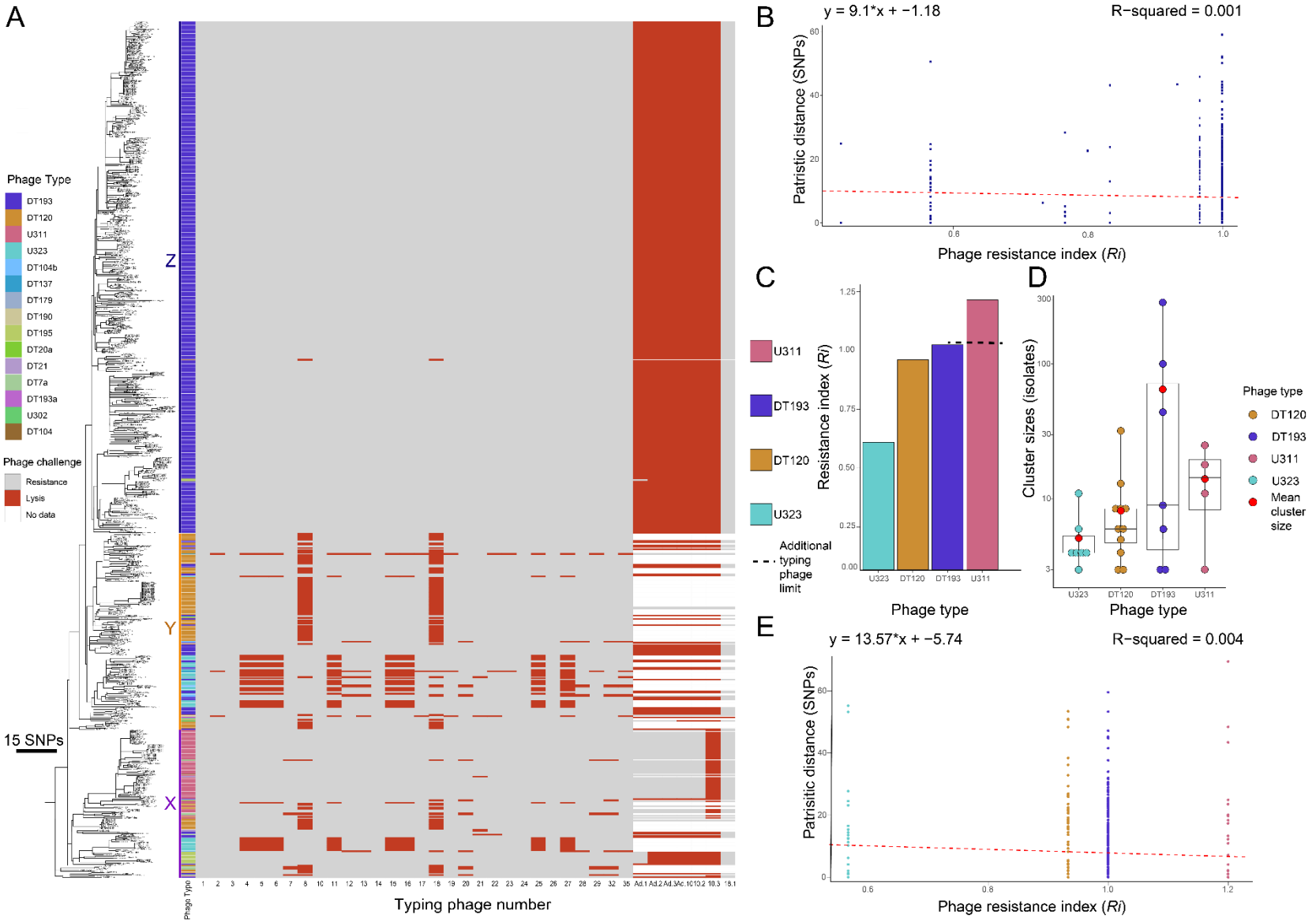

A
Phage Type
DT193
DT120
U311
U323
DT104b
DT137
DT179
DT190
DT195
DT20a
DT21
DT7a
DT193a
U302
DT104
Phage challenge
Resistance
Lysis
No data
Z
Y
X
15 SNPs
Typing phage number
B
y = 9.1*x + −1.18
R−squared = 0.001
Patristic distance (SNPs)
Phage resistance index (Ri)
C
U311
DT193
DT120
U323
Additional typing phage limit
Resistance index (Ri)
Phage type
D
Cluster sizes (isolates)
Phage type
DT120
DT193
U311
U323
Mean cluster size
E
y = 13.57*x + −5.74
R−squared = 0.004
Patrisitic distance (SNPs)
Phage resistance index (Ri)

Figure V.1 Phage sensitivity of monophasic *S.* Typhimurium ST34 strains from one year of human gastrointestinal infections, United Kingdom, 2014-2015. (A) Maximum-likelihood phylogenetic tree constructed using 6,403 core genome SNP sites from 763 monophasic *S.* Typhimurium strains isolated from one year of gastrointestinal infections, UK 2014-2015. Anderson typing scheme phage preparation sensitivity data is associated with each tip, and the tree split into three sections: X (purple vertical bar) is comprised of three separate clades; Y (orange vertical bar) is composed of 4 separate clades, and; Z (blue vertical bar) encompasses a single clade of strains phage typed DT193, except for two DT195 and one DT120 typed strains. (B) Linear regression of patristic distance against resistance index of each strain (*Ri*). (C) *Ri* of the four most frequently observed phage types. (D) Cluster sizes of the four most frequently observed phage types. (E) Pairwise, nearest neighbour, core genome SNP distances from subtrees of the four most frequently observed phage types plotted with corresponding *Ri* values for U323 (bright turquoise), DT120 (orange), DT193 (dark blue), and U311 (dark pink).

V.2 The ancestor of the ST34 epidemic clade that was hypothetically of phage type DT120 gave rise to multiple phage types during clonal expansion

Phage sensitivity ancestral states were reconstructed to assess phage resistance dynamics within the ST34 lineage and estimate the phage type of the most recent common ancestor (MRCA) of the monophasic ST34 lineage (Figure V.2). This was conducted using maximum posterior probabilities (MPP) from a joint probability distribution with transition rates estimated from the tip data, and Bayesian MCMC inference (455, 554, 555). MPP estimates suggested that DT120 was the phage type of the root node ancestor, possibly due to the frequency of isolates typed DT120 and their proximity to the root (Figure V.2 (A and B)). MPP estimates suggested that descendants increased or decreased their phage sensitivity in different lineages including 12 transitions to DT193 in this dataset. Due to large distribution-based probabilistic approaches such as stochastic maps generated with MCMC having a more robust estimate, the data was tested for possible root node ancestors using SIMMAP (455, 554). The resulting mean probability distribution for each point was then plotted on phylogenetic tree branches. DT120 had the greatest mean root node probability distribution of 47 % DT120 with 53 % probability distribution for any other phage type, consistent with an almost even chance that the MRCA

exhibited DT120 or another phage type. DT193 had a root node probability distribution of <5 %, consistent with the common ancestor of the monophasic ST34 clade having low probability of being DT193. These data are consistent with the hypothesis that phage resistance increased as the lineage evolved, and that DT193 was probably not the phage type of monophasic *S.* Typhimurium ST34 MRCA.

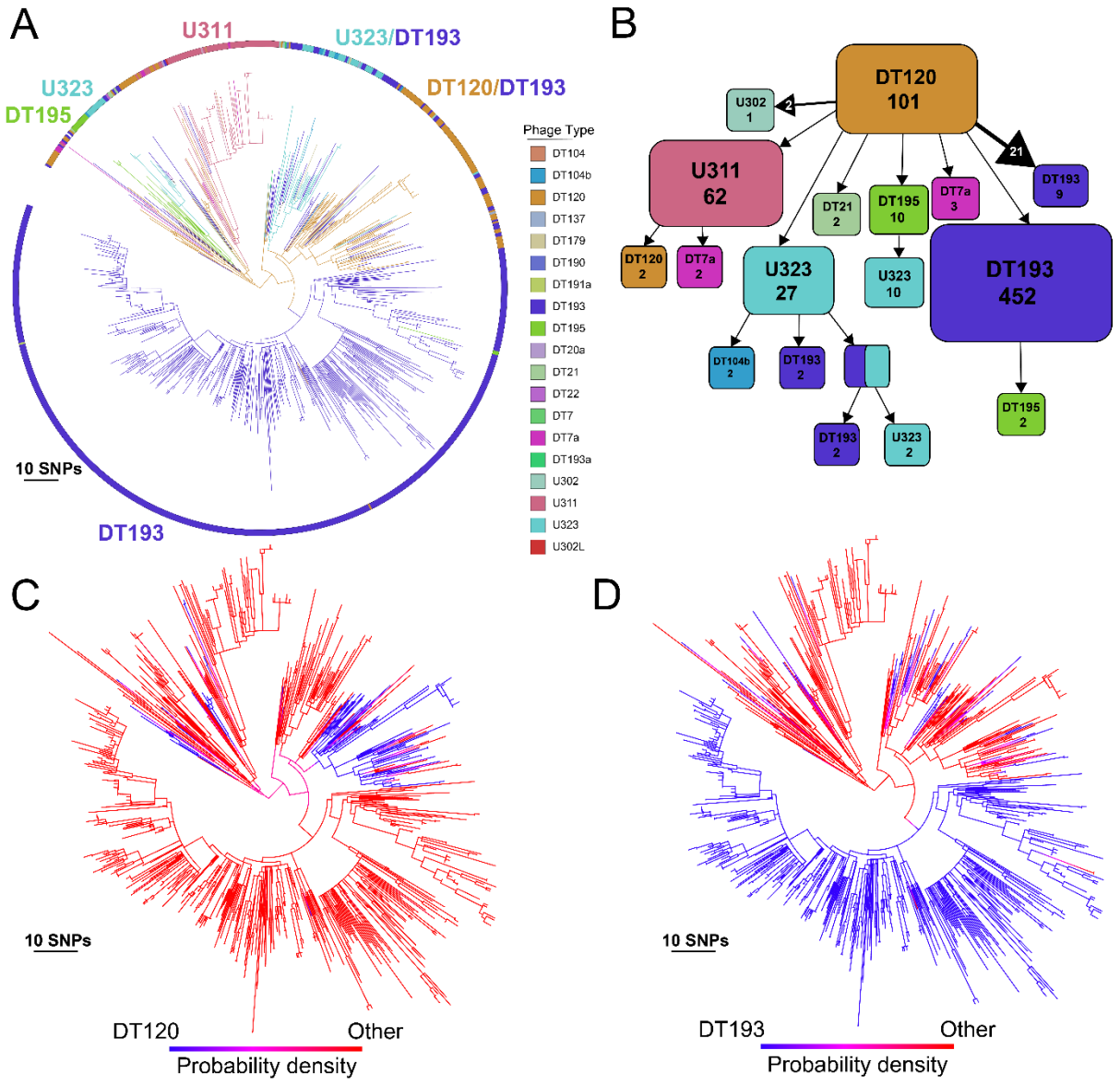


Figure V.2 Ancestral phage type history estimates of monophasic *S.* Typhimurium ST34 from one year of human gastrointestinal infections, United Kingdom, 2014-2015. (A) Marginal posterior-probability-based ancestral reconstruction using 19 phage types observed within the monophasic ST34 lineage α 17. The ring shows the phage type of each isolate and corresponding colours on branches show the estimate

of ancestral phage types. The majority phage type for distinct clades are labelled with their respective phage types as coloured in the key. (B) Collapsed-branch node graph using the same information as in (A) but with singular nodes removed. Arrows indicate how many nodes in (A) are between the two large nodes in (B) without state changes other than as indicated. The value associated with each node corresponds to how many tips can be tracked back to a certain node without changes of phage type. (C) Probabilistic Bayesian analysis labelling tips as either DT120 or 'other' to test the hypothesis from (A) and (B) that the root node phage type is DT120. Probability tending is toward DT120 is shown (blue, 75 %-100 %), as is probability tending toward any other phage type (red, 0-25 %), and values in between (pink). (D) Probabilistic Bayesian ancestral analysis labelling tips as either DT193 or 'other'. Probability tending toward DT193 is shown (blue, 75 %-100 %), as is probability tending toward any other phage type (red, 0-25 %), and values in between (pink).

V.3 Prophage mTmII is associated with DT193 phage typed strains

To investigate genomic regions that determine phage resistance in DT193, Kmer-based bacterial GWAS was undertaken with DT193 and DT120 strains. GWAS revealed a 43 Kb region highly associated with DT193 (LRT *p*-values from $1x10^{-5}$ to $1x10^{-19}$, Figure V.3). This region encompassed >99.9 % of the Kmers significantly associated with DT193 strains (~10,000 Kmers), with other genomic regions having reduced Kmer density, but high significance (LRT *p*-values from $1x10^{-7}$ to $1x10^{-19}$, Figure V.3 (A)). Loci with high significance but low Kmer density included: 2 copies of fimbrial usher protein encoding *fimD*; spermidine synthase encoding *speE*; carbon starvation protein encoding *cstA*; nitrate reductase alpha chain encoding *narZ*; putative aminopeptidase encoding *ysdC*; glucose 6-phosphate to 6-phosphogluconolactone oxidation catalyst encoding *zwf*; peptidase encoding *yfgC*, and; the left-hand end of the MDR transposon common to monophasic ST34 strains (Figure V.3 (D)). The alleles from regions excluding mTmII were plotted against a phylogenetic tree of the isolates used for GWAS (Figure V.3 (D)). Prophage mTmII was noticed to have synteny and sequence similarity with prophage mTmV. Prophage mTmV was recently shown to encode virulence factor SopE (536). Presence of mTmV containing *sopE* was significantly associated with a reduction of patristic distance.

Due to sequence similarity between mTmV and mTmII I wanted to ensure that the observed significant Kmers were associated specifically with mTmII (Figure V.3 (B and C)). Significant Kmers were mostly associated with regions specific to mTmII (Figure V.3 (C)). That most Kmers map to mTmII is consistent with mTmII imparting the change in phage resistance observed between DT120 ($Ri = 0.96$) and DT193 ($Ri = 1$).

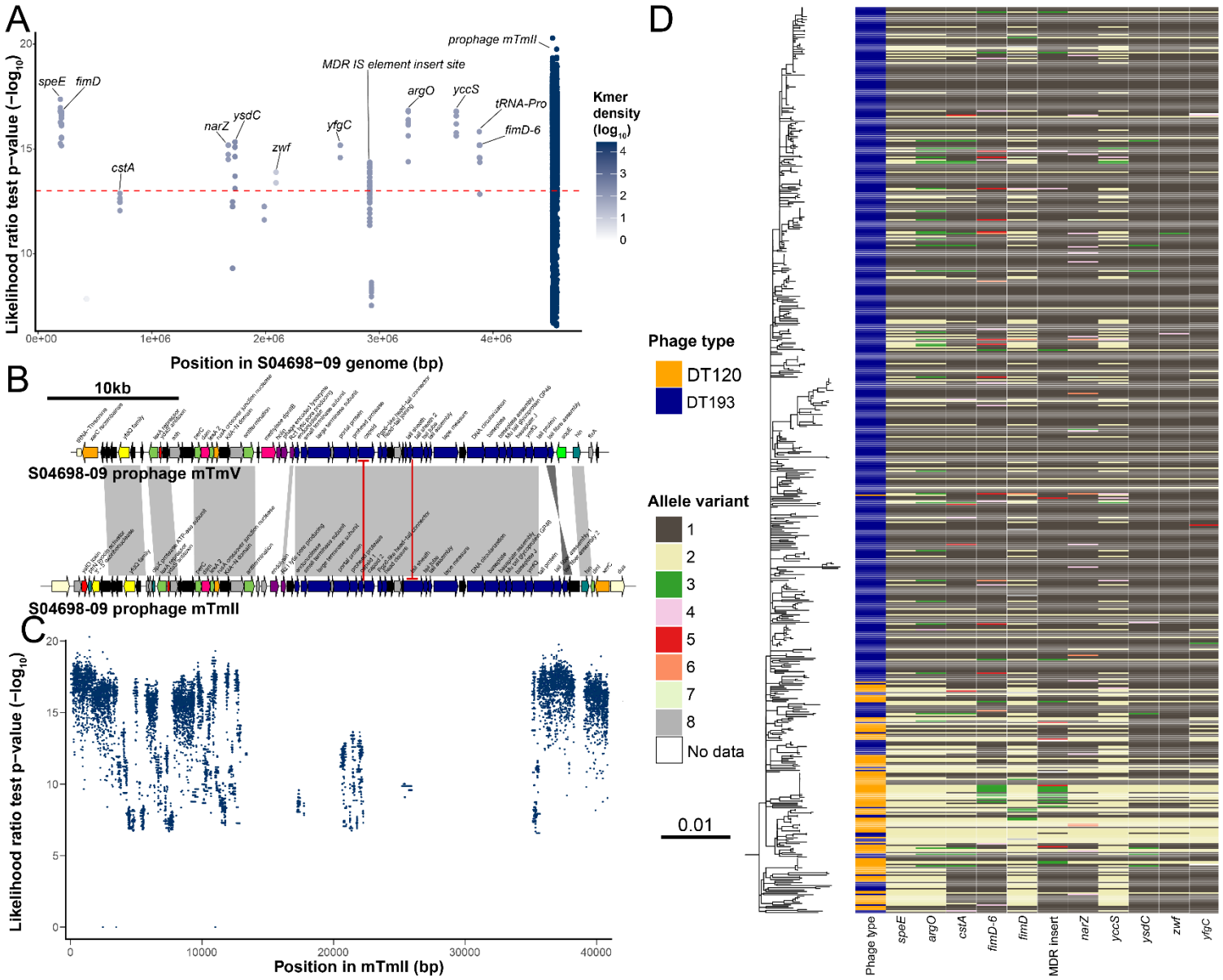

A
Likelihood ratio test p−value (−log10)
Position in S04698−09 genome (bp)
prophage mTmII
MDR IS element insert site
speE
fimD
cstA
narZ
ysdC
zwf
yfgC
argO
yccS
tRNA-Pro
fimD-6
Kmer density (log10)
B
10kb
S04698-09 prophage mTmV
S04698-09 prophage mTmII
C
Likelihood ratio test p−value (−log10)
Position in mTmII (bp)
D
Phage type
DT120
DT193
Allele variant
1
2
3
4
5
6
7
8
No data
0.01
Phage type
speE
argO
cstA
fimD-6
fimD
MDR insert
narZ
yccS
ysdC
zwf
yfgC

Figure V.3 Genome-wide association of DT193 and DT120 with whole genome sequence Kmers from strains of monophasic *S.* Typhimurium ST34. (A) Kmers significantly associated with DT193 mapped to the chromosome of monophasic ST34 DT193 reference strain S04698-09. The red dashed line indicates the inferred *p*-value threshold. (B) Genetic diagram showing monophasic ST34-associated prophage mTmV and DT193-associated prophage mTmII. Sequence with >75 % nucleotide sequence similarity (vertical grey lines), and genomic inversion (dark grey) are indicated. Predicted genes and open reading frames (arrows) encoding proteins involved in integration and excision (orange), toxin-antitoxin systems (red), repressor proteins (green), DNA modification encoding genes (pink), nucleases (yellow), cell lysis (purple), phage particle (dark blue), conserved hypothetical prophage associated (grey), other hypothetical proteins (black), *sopE* (bright green), and other well characterised genes not grouped by function (dark green). Vertical red lines indicate insertion of stop codons. (C) Location of significantly-associated Kmers mapping to mTmII, congruent with genetic locations of mTmII in (B). (D) Maximum-likelihood phylogenetic tree constructed using 7,268 core genome SNP sites from 605 monophasic *S.* Typhimurium strains phage typed DT193 or DT120. The tree is associated with alleles to which significant Kmers mapped from each strain. Unique alleles are displayed, with numbers corresponding to how frequent each gene's allele is within the population (1-8).

V.4 Prophage mTmII is correlated with transitions to phage type DT193 in the monophasic *S.* Typhimurium ST34 population

To identify which strains had mTmII within monophasic S. Typhimurium ST34 isolated from one year of human infections in the UK, the nucleotide coding sequence of mTmII was extracted from monophasic ST34 reference strain S04698-09, and a database of each ORF generated. ARIBA was used to map and construct local assemblies from strain WGS read files (478). Presence of a gene was defined as 90 % sequence similarity, and mTmII presence defined as any strain which had 57 of the 61 genes.

Four clades had mTmII introduced on separate occasions. Each introduction of mTmII was associated with a switch to DT193. However, not all DT193 phage typed strains had mTmII (33 of 493; 7.7 %), suggesting secondary mechanisms of phage defence being introduced into these strains that remain

elusive. These data are consistent with prophage mTmII conveying the phage resistant phage type DT193, but phage type DT193 also evolving due to other mechanisms.

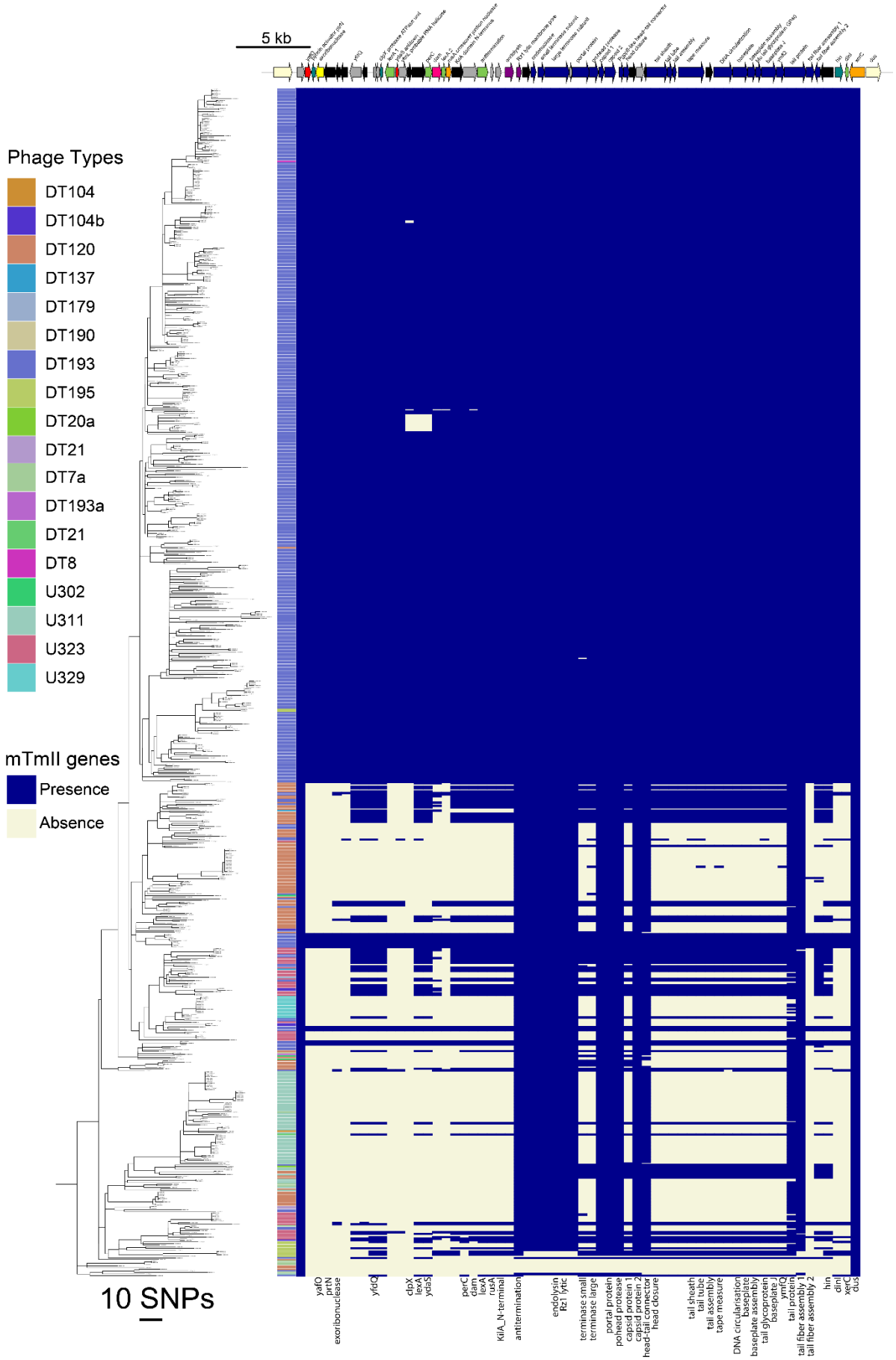

5 kb
Phage Types
DT104
DT104b
DT120
DT137
DT179
DT190
DT193
DT195
DT20a
DT21
DT7a
DT193a
DT21
DT8
U302
U311
U323
U329
mTmII genes
Presence
Absence
10 SNPs
yafO
prtN
exoribonuclease
yfdQ
clpX
lexA
ydaS
perC
dam
lexA
rusA
KilA_N-terminal
antitermination
endolysin
Rz1 lytic
terminase small
terminase large
portal protein
portal protein
capsid protein 1
capsid protein 2
head-tail connector
head closure
tail sheath
tail tube
tail assembly
tape measure
DNA circularisation
baseplate
baseplate assembly
tail glycoprotein
baseplate J
ymfQ
tail protein
tail fiber assembly 1
tail fiber assembly 2
hin
dinI
xerC
dUTPase

Figure V.4 Presence of prophage mTmII within a collection of monophasic *S.* Typhimurium ST34 isolated from one year of human infections, UK 2014-2015. Maximum likelihood phylogenetic relationship constructed using 6,403 core genome SNPs from 763 monophasic *S.* Typhimurium ST34 strains isolated from human infections, 2014-2015. The tree is associated with each strain's phage type (first column) and gene presence or absence based on local assemblies from read files mapping to each gene from prophage mTmII, as extracted from reference monophasic ST34 strain S04698-09 (columns 2-62).

V.5 Prophage mTmII is related to prophage mTmV and both share a common ancestor with *Shigella flexneri* phages

To investigate well characterized phages related to mTmII its nucleotide coding sequence was extracted from monophasic ST34 reference strain S04698-09 by identifying the insertion site as described above. The amino acid sequence of each ORF was then used as a query for neighbour joining phylogenetic reconstruction with characterised reference phages as implemented with VipTree (443). Next, Nucleotide sequences of closely related phages defined by VipTree were downloaded from the NCBI nucleotide repository. The ORFs from phage sequences were annotated manually through query of each ORFs translated protein sequence using the NCBI non-redundant amino acid coding sequence database.

The closest related phage to both mTmV and mTmII was a phage isolated from *Shigella flexneri*, SfV (328), suggesting that these phages have a host-range that is at least intergenera. Other closely related phages which share a common ancestor with mTmII include another prophage from S04698-09, mTmV, which was characterised in our recent study (536), and 3 further phages observed as prophages in *Shigella flexneri*; SfI (556), SfII (557), and SfIV (558)(Figure V.5). Genes encoded on mTmII and related phages that are similar include those for phage particle structure and assembly (Figure V.5, dark blue). Prophage mTmII has a lysis gene cluster not found in closely related phages (Figure V.5, dark purple). Prophage mTmII also encodes a pyocin transcriptional activator encoding gene, *ptrN*, and protease ATPase domain encoding *clpX*. The functions of these genes are difficult to discern or hypothesise. It is possible that they are involved in regulation in a cI, cII, and cro-like manner, as these genes appear missing from mTmII and would hypothetically be required for effective lysogeny

regulation. There are two genes encoding possible toxins: *ydaS* and *yafO* (559, 560). YdaS is labelled as an antitoxin in the pFam database, and despite evidence that it functions as a toxin (561), this is disputed (562). YafO encodes a type II toxin typically found with its cognate antitoxin, YafN (563). Prophage mTmII -encoded YafO is a hypothetical toxin but a cognate antitoxin was not present. Whether mTmII encodes a previously undescribed YafO silencer, or a chromosomal antitoxin is acting as YafO's counterpart is undetermined. However, toxin-antitoxin systems fundamentally change the ability of bacteria to resist a host of stressors in multi-level interactions (564-569). The two extra toxin genes may be enhancing the fitness of mTmII encoding strains, possibly contributing to enhanced phage resistance (546).

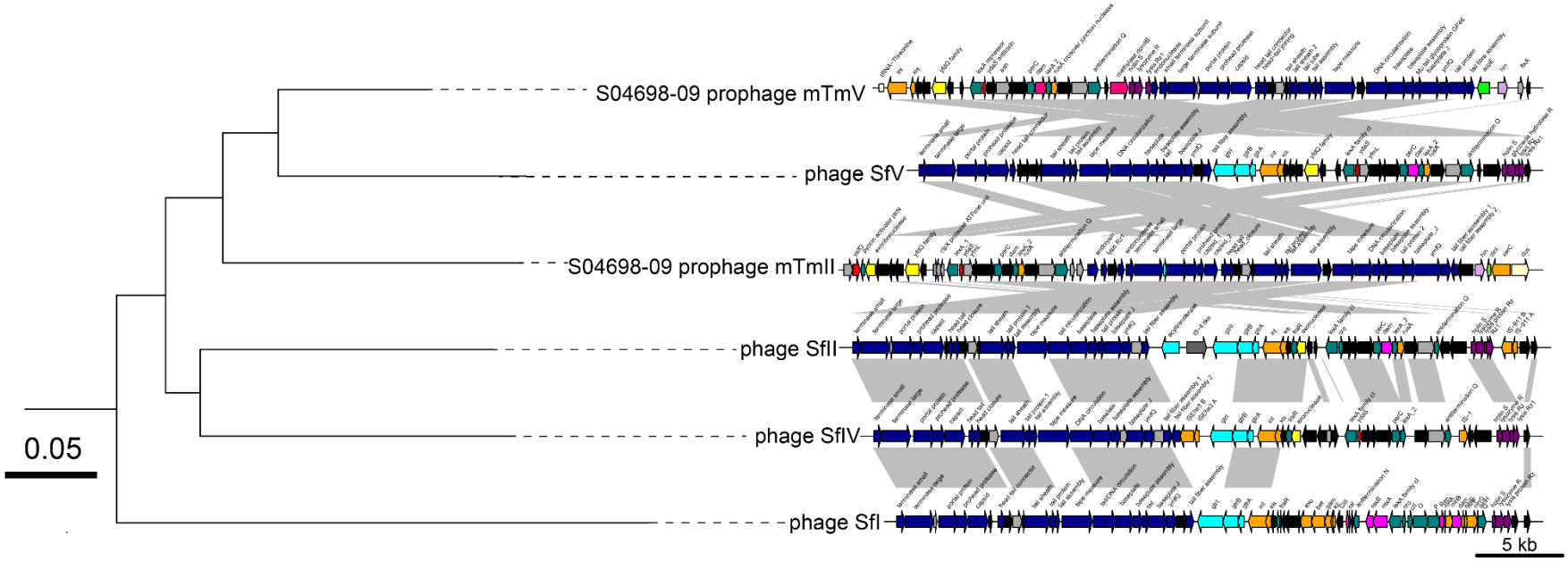


Figure V.5 Phylogenetic relationship and sequence similarity of S04698-09 prophage mTmII and related phage. Phylogenetic relationship and genetic diagrams showing position, size, and orientation of genes for phages similar to prophage mTmII. The neighbour joining phylogenetic tree was constructed using ORF translated amino acid sequences. ORFs are colour coded by function as follows: hypothetical attB sites (white); integration and excision including IS elements (orange); phage lysogeny regulation (dark green); nucleases (yellow); phage defence (bright pink); virulence factor SopE (bright green); phage particle structure and assembly (dark blue); glucosyl and acyl transferases (bright blue); hypothetical toxin-antitoxins (red); host cell lysis (dark purple); phage associated hypothetical proteins (grey), and; other hypotheticals (black).

V.6 Prophage mTmII shows a similar distribution in human infections from 1 year in the UK to strains within a global collection of monophasic *S.* Typhimurium ST34

Prophage mTmII was present largely in proximal lineages of the phylogenetic tree of strains isolated from human infection. The data was consistent with a common ancestor acquiring mTmII and undergoing clonal expansion. To identify if the distribution of mTmII was similar in a broader collection of monophasic ST34 strains, a maximum-likelihood phylogenetic tree of 394 strains isolated globally between 1998-2019 was reconstructed based on 5,274 SNP sites in the core genome. The population displayed a similar distribution of mTmII (Figure V.6 (A)). More basally rooted strains had 11 apparent acquisitions of mTmII, while one apparent acquisition at a more proximal node gave rise to a large sub-clade containing mTmII that exhibited evidence of some limited recombination and loss of mTmII from four strains. The patristic distance of strains containing mTmII was significantly lower than those without mTmII ($p<0.05$) (Figure V.6 (B)). Reduced patristic distance is consistent with more frequent sampling of mTmII containing strains, consistent with it contributing to the continued success of the monophasic ST34 lineage, possibly through enhanced phage resistance.

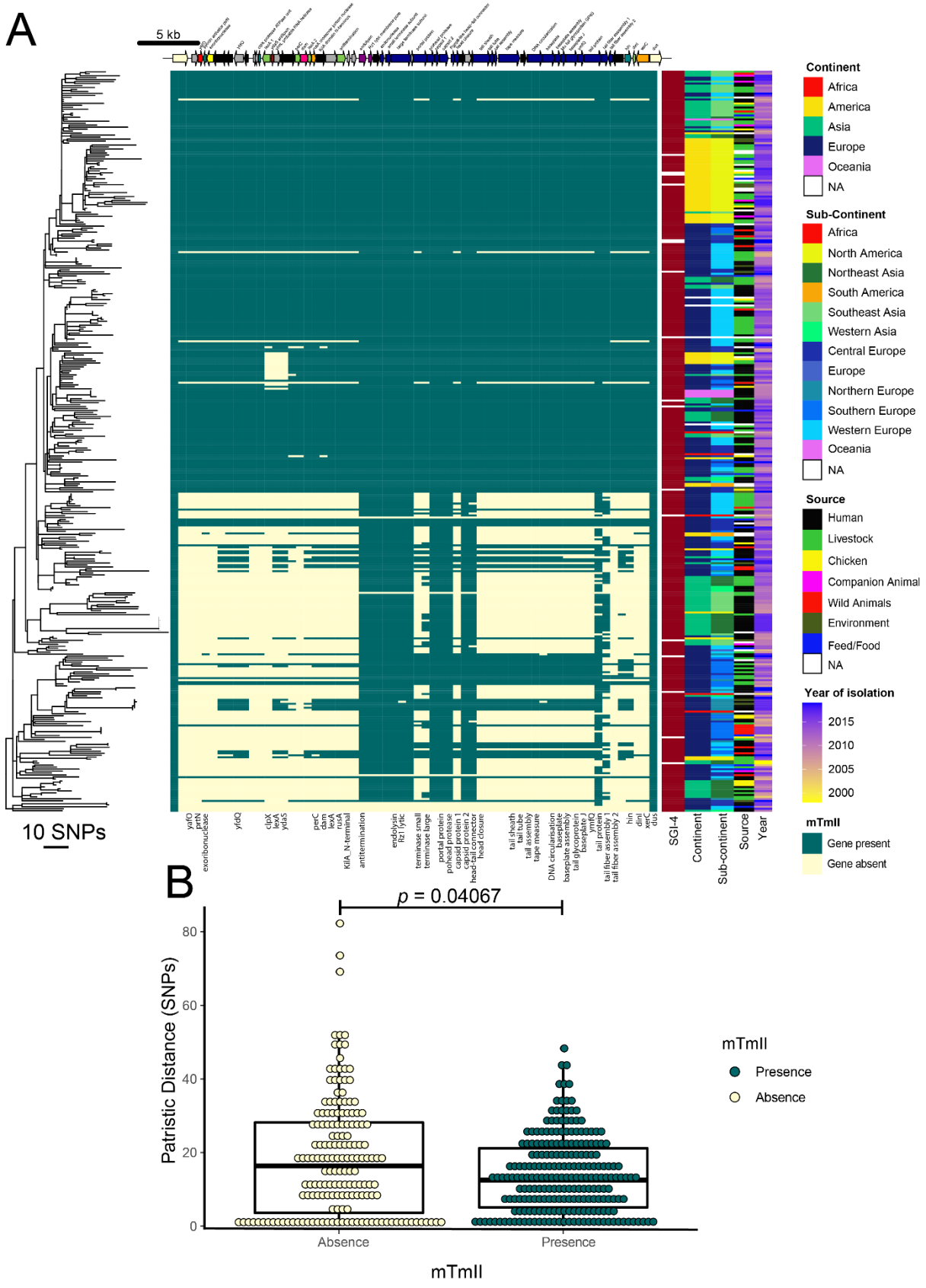

A
5 kb
10 SNPs
Continent
Africa
America
Asia
Europe
Oceania
NA
Sub-Continent
Africa
North America
Northeast Asia
South America
Southeast Asia
Western Asia
Central Europe
Europe
Northern Europe
Southern Europe
Western Europe
Oceania
NA
Source
Human
Livestock
Chicken
Companion Animal
Wild Animals
Environment
Feed/Food
NA
Year of isolation
2015
2010
2005
2000
mTmII
Gene present
Gene absent
SGI-4
Continent
Sub-continent
Source
Year
B
p = 0.04067
Patristic Distance (SNPs)
0
20
40
60
80
Absence
Presence
mTmII
mTmII
Presence
Absence

Figure V.6 Distribution of mTmII within a global collection of monophasic *S.* Typhimurium ST34 strains. (A) Phylogenetic relationship of 394 monophasic *S.* Typhimurium ST34 strains constructed using 5,274 SNP sites and associated with mTmII gene presence or absence based on read files having a gene which has >90 % sequence identity. Presence of SGI-4 has been plotted next to the mTmII gene presence and absence (column 63; SGI-4), indicating SGI-4 presence (red) and absence (white). (B) Patristic distance of strains with mTmII (green) compared to strains without mTmII (white). Probability value for differences in patristic distance between mTmII positive and negative strains determined by the Mann-Whitney-Wilcoxon test.

V.7 Strains containing both mTmII and mTmV display a significant reduction in pairwise core genome SNP distances

During a recent study (Tassinari et al., (2020) (589)), we identified that strains isolated from one year of human gastrointestinal infections in the UK had a reduced patristic distance if prophage encoded *sopE* was present (536). This is consistent with frequent sampling of strains that encode mTmV, consistent with positive selection. To investigate whether mTmII had a similar effect and determine if mTmV or mTmII were more associated with reduced patristic distance, a phylogenetic relationship of monophasic ST34 strains from one year of human infections in the UK from 2014-2015 was reconstructed and presence or absence of prophages mTmII and mTmV associated with each tip (Figure V.5, (A)). The patristic distance for all strains containing mTmII was compared against all strains without mTmII. The same analysis was repeated with mTmV. Presence of either prophage significantly reduced the patristic distance of the strains, consistent with them being under positive selection (Figure V.7 (B)). However, when comparing the patristic distances of strains with only either mTmII or mTmV, neither had significantly reduced patristic distances compared with isolates which had neither prophage. Significant reduction in reduced patristic distance was only observed for strains that have both mTmII and mTmV present in their genome, consistent with both phages being required for enhanced fitness.

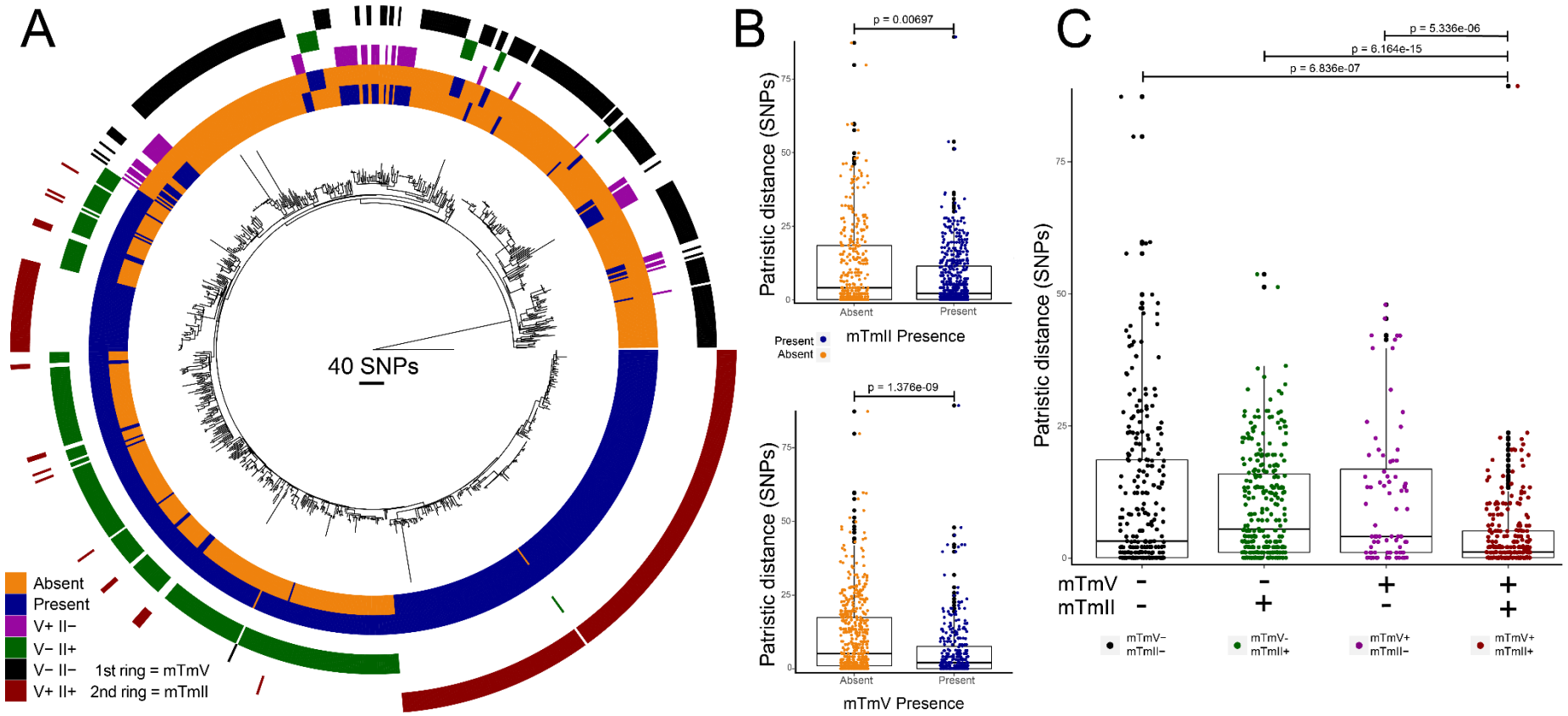


Figure V.7 Patristic distances of strains containing prophages mTmV and mTmII isolated from one year of human gastrointestinal infections, United Kingdom, 2014-2015. (A) Phylogenetic relationship of monophasic *S.* Typhimurium ST34 strains isolated from 1 year of gastrointestinal infections, UK. The inner most ring (first ring) indicates presence or absence of mTmV, and the second ring from the inside (second ring) indicates

presence or absence of mTmII. (B, top) Patristic distance of strains with presence of prophage mTmII (blue) compared to those without mTmII (orange). (B, bottom) Patristic distance of strains with presence of mTmV (blue) compared to those without mTmV (orange). (C) Comparison of: strains lacking both mTmII and mTmV; strains with mTmII but not mTmV; strains with mTmV but not mTmII, and; strains that contain both prophages. Statistical significance was assessed using the Mann-Whitney-Wilcoxon test with resulting *p*-values indicated on graphs.

V.8 Lysogenic mTmII confers phage resistance in DT120 strains

To investigate the hypothesis that acquisition of mTmII conveyed increased resistance to phage in DT193 strains, a donor strain with a selectable marker in mTmII was used to isolate mTmII before introduction into a DT120 strain. To construct the donor strain, a chloramphenicol resistance gene (*cat*) was introduced into an intergenic region upstream of ORF1 of lysogenic mTmII in monophasic ST34 reference strain S04698-09. To provide a selectable marker for a recipient strain, a kanamycin resistance gene (*aphII*) was introduced upstream of *iciA* in DT120 strains L00745-07 and L00979-07. The mTmII prophage was first transferred into a derivative of strain 4/74 in which all prophage had been removed, giving rise to 4/74:mTmII. Prophage mTmII was then transferred into DT120 strains giving rise to mTmII lysogen strains L00745-07:mTmII and L00979-07:mTmII. Phage sensitivity of parent DT120s (L00754-07 and L00979-07), DT193s (S04698-09 and S04332-09), and mTmII lysogenic DT120s (L00745-07:mTmII and L00979-07:mTmII) to three typing phages (8, 18, and 29) was determined in liquid culture assays. In the Anderson phage typing scheme DT120 is defined by varying degrees of lysis or confluent lysis to typing phages 8 and 18, and DT193 strains defined by resistance to these phages. Strains of both types are resistant to lysis by phage 29 and it was used as a control. Each strain was cultured with phage with an MOI of one. Lysis was measured by measuring the OD_{600} of the culture during growth at 37^0C. DT120 strains showed a distinct reduction in their OD_{600} after 8-10 h (Figure V.7 (A)). DT120 strains with lysogenic mTmII displayed resistance to phages 8 and 18. The area under the curve was then calculated using definite integrals from 5 to 15 h (Figure V.7 (C)).

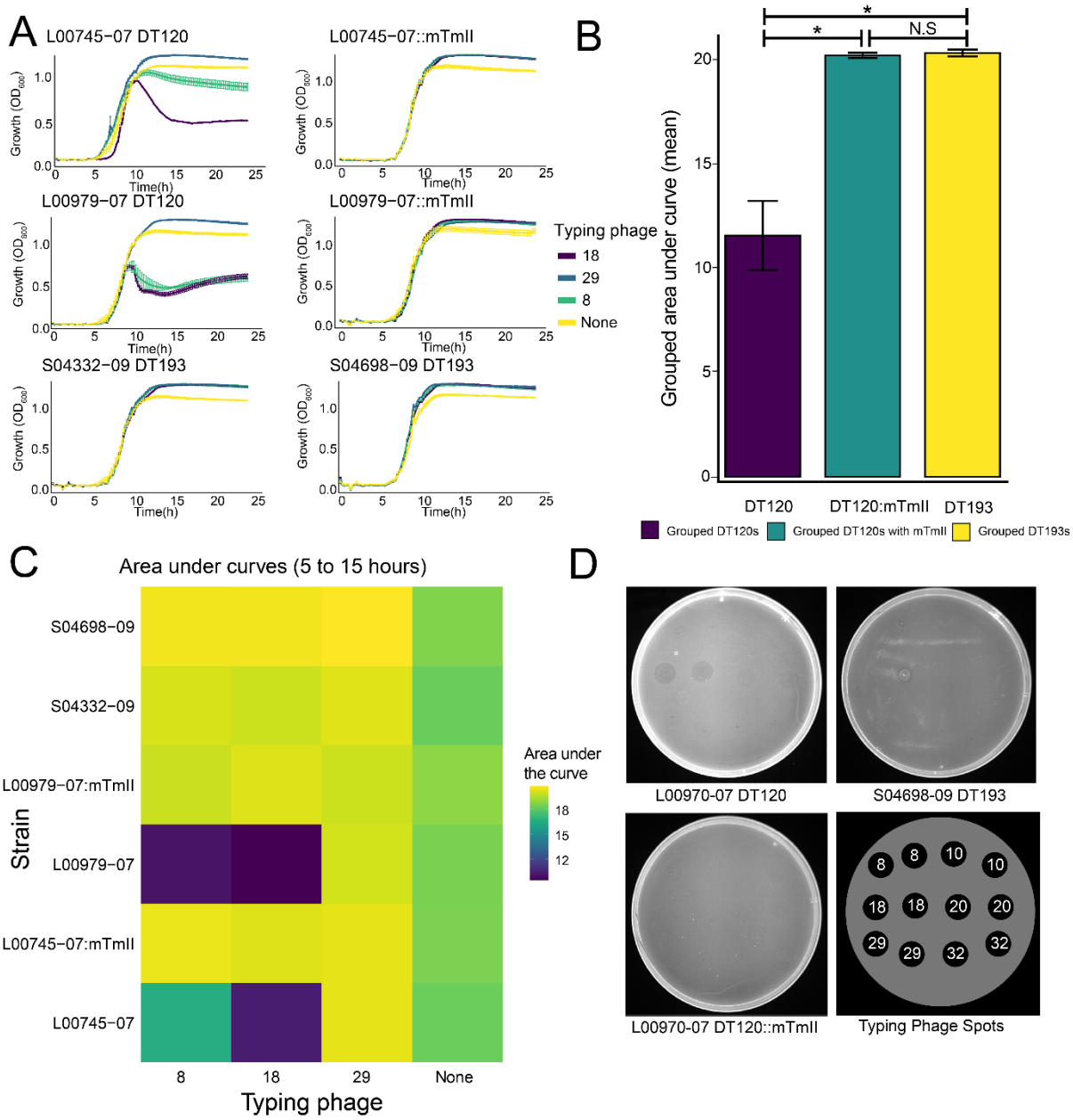


Figure V.8 Effect of phage pressure on the growth of DT120, DT193, and DT120 with lysogenic mTmII strains. (A) Effect of pressure from typing phages 8, 18, and 29 compared with regular growth over 24 h as measured by the optical density of bacterial cultures at 600 nM. (B) Area under the curves between 5 and 15 h from (A) as grouped by DT120, DT193 or DT120s with lysogenic mTmII. Data from phages 8 and 18 were used in this analysis, as these are phages which show a difference between DT120 and DT193 in the Anderson phage typing scheme. * = Mann-Whitney-Wilcoxon test *p*-value of <0.05. N.S = not significant. (C) Area under curve heatmap showing the data from each curve in (A). (D) Phage typing plates using phages 8, 10, 18, 20, and 29 with DT193 strain S04698-09, DT120 strain L00979-

07, and L00979-07 with lysogenic mTmII. Mutants constructed and experiments undertaken by Luke Acton, experimental design and data analysis carried out by this author.

The area under the curves were then grouped by phage type or mTmII presence and a significant difference observed between parent DT120 strains and DT120 strains with lysogenic mTmII, but no difference between DT193 and mTmII lysogenic DT120 strains. The observed resistance in mTmII lysogenic DT120s suggests that mTmII is conferring phage resistance and providing the DT193 phage type.

V.9 Prophage mTmII was acquired before mTmV

Recently, we reported that monophasic *S.* Typhimurium ST34 prophage mTmV was acquired on multiple occasions from 2002 to 2010 using time-scaled phylogenetic analysis (536). From maximum-likelihood phylogenetic analysis, it was apparent the main DT193 clade began with one clone that acquired mTmII and subsequently underwent clonal expansion (Figure V.4). To determine the acquisition time of mTmII for the main DT193 clade, Bactdating was used (465), initially to ascertain whether the global collection of monophasic *S.* Typhimurium ST34 had a suitable temporal signal, and subsequently to infer a time-structured phylogenetic tree (Figure V.8 (B and C)). There was significant probability that a regression line had a good fit to the data (p=1x10^{-4}), suggesting an association between SNP distance and isolation time. The clade in which mTmII was acquired before undergoing clonal expansion was determined to have a common ancestor around 1994 (CI 95 % = 1991 to 1996).

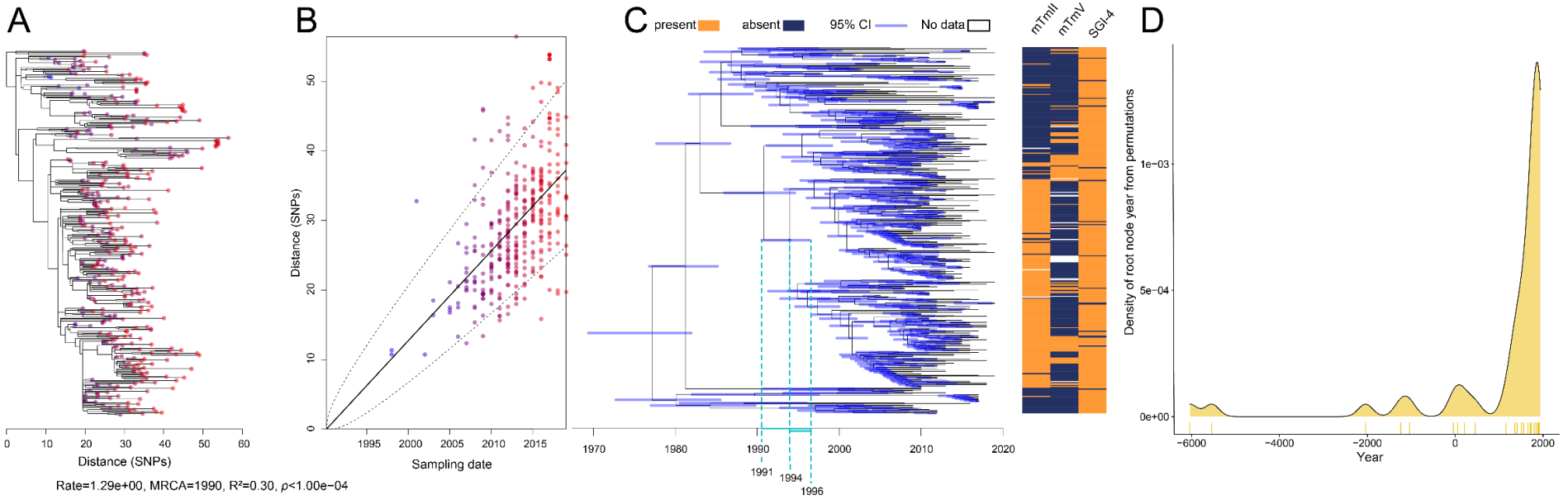


Figure V.9 Temporally structured phylogenetic tree of a global strain collection of monophasic *S.* Typhimurium ST34 showing acquisition time of prophage mTmII within a clade that subsequently underwent clonal expansion. (A) Maximum likelihood phylogenetic tree of global monophasic *S.* Typhimurium strains with isolation date, indicating if a strain was isolated further back in time (blue) or more recently (red). (B) Linear regression of time of strain isolation against SNP distance from root to tip, indicating strains isolated before 2005 (blue), strains isolated between 2005 and 2010 (purple), and strains isolated after 2010 (red). (C) Temporally structured phylogenetic tree showing 95 % confidence intervals and

presence or absence of notable monophasic *S.* Typhimurium ST34 genomic elements. (D) Frequency of root node dates from 50 permutation tests (6000 bc – 1950 ad).

SGI-4 is present in most monophasic *S.* Typhimurium isolates (Figure V.9 (C)), and from this data the lineage was estimated to have a common ancestor around 1976 (95 % CI = 1970-1983). These dates indicate when an initial clone may have gained SGI-4 before undergoing expansion, probably within porcine reservoirs and the clonal expansion event possibly coinciding with the transition of use of antbiotics to use of copper as a growth supplement and antimicrobial in pigs (273). To determine if the inference from Bayesian MCMC analysis was based solely on the structure of the tree and sample dates, 50 permutations were undertaken by randomly assigning the dates to each tip (Figure V.9 (D)). Only 13/50 permutations were significantly different from the real data when using the Mann-Whitney-Wilcoxon test (570, 571), providing a mere 26 % confidence that the model generated by the real data did not occur by chance. However, all permutation root dates were further back in time than the estimation from the real date (permutations from 6000 bc – 1950 ad, real data estimate ~1976).

V.10 Chapter V discussion

Strains within a clade having a reduced core genome SNP distance (patristic distance) is indicative of repeated sampling of closely related strains, consistent with positive selection. The patristic distance of monophasic ST34 strains was shown to have a weak negative correlation with *Ri*. This correlation is consistent with increased phage resistance providing increased sampling, consistent with positive selection of phage resistant strains. However, one of the four most frequent phage types observed in the monophasic ST34 lineage was U323 (*Ri*=0.5667). The presence of this phage type may suggest that different clades have different phage selection pressures, but the host origin of these strains beyond a case of human infection was undetermined. Despite the frequent observations of U323, strains with this phage type had a lower patristic distance than those of phage types representing higher *Ri* values, especially DT193 and U311 (*Ri*=1). DT193 was the most frequent phage type, and is characterised by resistance to all standard phage preparations in the Anderson typing scheme (165). When estimating the ancestral history of phage types a common theme among results is that DT193 is not the monophasic ST34 lineage's ancestral phage type. The frequent observation of DT193 and identification that it is not

the root node ancestor is consistent with phage resistance being selected for in broad-host-range strains. Kmers from 12 loci were associated with DT193. Kmers mapping to regions other than prophage mTmII had an apparent link in their allele changes, with the second most frequent alleles commonly appearing in the same strains. An explanation for this link may be altered modification at common sequences by type I RM specificity subunits, such as HsdS, which has been shown to exhibit phasevarions producing altered modification sites (572, 573), although this hypothesis was not investigated. A genetic element highly associated with DT193 was prophage mTmII. The abundance of mTmII in a time-minimised, human-isolate phylogenetic tree suggests that its acquisition was beneficial. Observing a similar distribution of mTmII within global monophasic ST34 strains corroborates this and is consistent with mTmII providing a fitness advantage to strains circulating in broad-hosts, surviving in food production, and possibly enhanced environmental fitness. These data also suggest mTmII increases host phage resistance and lowers the core genome SNP distance of strains being isolated. This supports the central hypothesis being tested in this chapter - that broad-host-range lineages require increased phage resistance to exclude phages within their extended biome. Due to the abundance of mTmII, particularly with increasing numbers in chromosomes of monophasic *S.* Typhimurium ST34 strains, it can be hypothesised that genes which have second copies within the chromosome may be exerting dominance over their counterparts (194). This may be due to subtle mutations which allow such mechanisms as increased protein binding during protein interactions due to more efficient tertiary structure, or co-transferred genes allowing significantly refined regulation of gene expression. Bacterial gene dominance was recently shown to be important in governing the spread of MGEs, playing a role in determining undiscovered incompatibility through negative selection from an MGE harbouring genetically recessive genes, and discontinuation of the lineage (194). Two notable features of mTmII present in most strains harbouring the phage include a putative toxin gene *yafO* (559), and a less well characterised gene that has literature evidence of being a toxin, yet conversely labelled in pFam-v33 as an antitoxin (*ydaS*)(561). Introduction of new toxin genes can have striking effects on the biology of bacteria, as they play a multifaceted role in minimising metabolic activity in conditions of stress (574). However, their main function appears to be phage defence (546). It is possible that these two genes are increasing the ability of monophasic *S.* Typhimurium to survive in many different environments, from phage stress, to the

stressors encountered within epithelial cells after invasion (574, 575). The extent to which rapid introduction of new toxin genes changes the physiology of pathogens in this manner is not well understood and requires further investigation.

The observation that both mTmII and mTmV are required for a significant decrease in patristic distance is consistent with these two phages both contributing to enhanced fitness of monophasic *S.* Typhimurium ST34 strains. It is possible that harbouring both prophages, one which introduces a virulence gene and the other enhancing phage predation resistance, explains this observation (536). However, the dataset from which these observations were made is limited to human infection isolates, and the possible role of the phages by themselves cannot be linked to strain epidemiology. These data are also consistent with mTmII being introduced to the lineage after SGI-4. It is possible that we are witnessing the stepping-stones required for the continued success of a broad-host-range pathogen. A discrepancy was observed between possible common ancestor dates for the monophasic ST34 lineage. This study observed a common ancestor for the monophasic lineage with 95 % confidence intervals from 1970-1983, where as in Tassinari et al., 2020 it was determined to be ~1985-2000. Both data structures had equal probability best fit linear regression ($p<1x10^{-4}$) and similar R^2 values (0.30 and 0.34). The main difference lies in the range and quantity of strains used for the analysis, this study using more strains, including two strains isolated from an earlier time (1998). It is possible that including strains isolated further back in time is generating the observed discrepancy. Neither strain isolated in 1998 was observed to have the MDR transposon disrupting the *fljB* locus as defined by searching read files with ARIBA (Tassinari et al., Unpublished data). However, SGI-4 was present within one isolate, and they are both phylogenetically related to the monophasic ST34 clade. It is possible that by using an extended strain isolation date range information is being included that was missed in the previous study, but each result has been generated from different datasets.

Chapter VI: Microevolution of O-antigen Polymerase Encoding *wzy* Provides Resistance to Phage Predation

Ancestral reconstruction of phage sensitivity patterns (Chapter IV) revealed the dynamics of phage resistance and its effect on the population structure of *S.* Typhimurium. Here, the effect of phage predation on the population structure of a second epidemic group of *S.* Typhimurium was investigated. This epidemic group was selected since it contains strains that are highly host adapted to one avian family (common ducks: Anatidae), as well as several human outbreaks in the UK and Ireland from contaminated duck egg-containing products (185, 242). Strains from the DT8 lineage (β 2, herein referred to as the DT8 lineage) had remarkably conserved phage sensitivity profiles with only sporadic switches to a profile of increased resistance to phage characterised by the DT30 profile. In contrast to the broad-host-range epidemic represented by the ST34 clade, increased resistance to phage was not followed by increased clonal expansion consistent with an increase in fitness.

Increased phage resistance was predominantly associated with deletion of a single gene involved in O-chain polymerisation of long chain LPS (*wzy*), a function important for *Salmonella* survival and pathogenesis. Deletion of *wzy* was due to precise excision at repeat sequences flanking the gene, consistent with a novel cis-integrating mobilizable element (CIME). These analyses provide insight into how *S.* Typhimurium utilises recombination to generate rapid antigenic variation within closely related bacterial populations that may also represent a rapid reversable mechanism for phage resistance.

VI.1 DT30 strains are sporadically distributed in the duck-associated DT8 lineage and DT30 is not associated with loss of pSLT or gain of an ICE

To investigate the population structure of *S.* Typhimurium DT8, a maximum-likelihood phylogenetic tree was reconstructed of 292 strains isolated between 1993 and 2013 including 182 DT8 strains isolated from ducks and a small number of other species, and 111 strains from 18 other phage types (Figure VI.1 (A)). The majority of DT8 strains (172 of 182) were present in an epidemic clade along with 37 strains of DT30. A small minority were of 16 phage types, including 5 DT9 typed strains. DT30 strains exclude 10 phage preparations compared to DT8 in the Anderson phage typing scheme (241). This is an increase of 0.33 in *Ri* from 0.634 in DT8 to 0.967 for DT30 within strains <110 SNPs apart (Table VI.1).

Table VI.1. Phage sensitivity profile of DT8 and DT30. Sensitivity of *Salmonella* Typhimurium DT8 and DT30 strains to 30 phage preparations in the Anderson typing scheme. CL corresponds to confluent lysis, a strong lytic interaction between the phage and bacteria, SCL means semi-confluent lysis, a strong signal but noticeably different to confluent lysis, and +++, ++, +/- indicate varying degrees of plaques generated by the phage preparations.

	Typing Phage Number/Result																													
Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35
DT8	-	-	-	-	-	-	-	CL	SCL	SCL	-	-	-	-	+++	-	-	-	SCL	-	SCL	SCL		+/-	+/-	-	-	CL	CL	-
DT30	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

The distribution of DT30 strains amongst DT8 strains in the phylogenetic tree is sporadic with no more than 2-5 strains per cluster of DT30 (Figure VI.1 (A)). This is consistent with DT30 emerging multiple times throughout the population with limited clonal expansion.

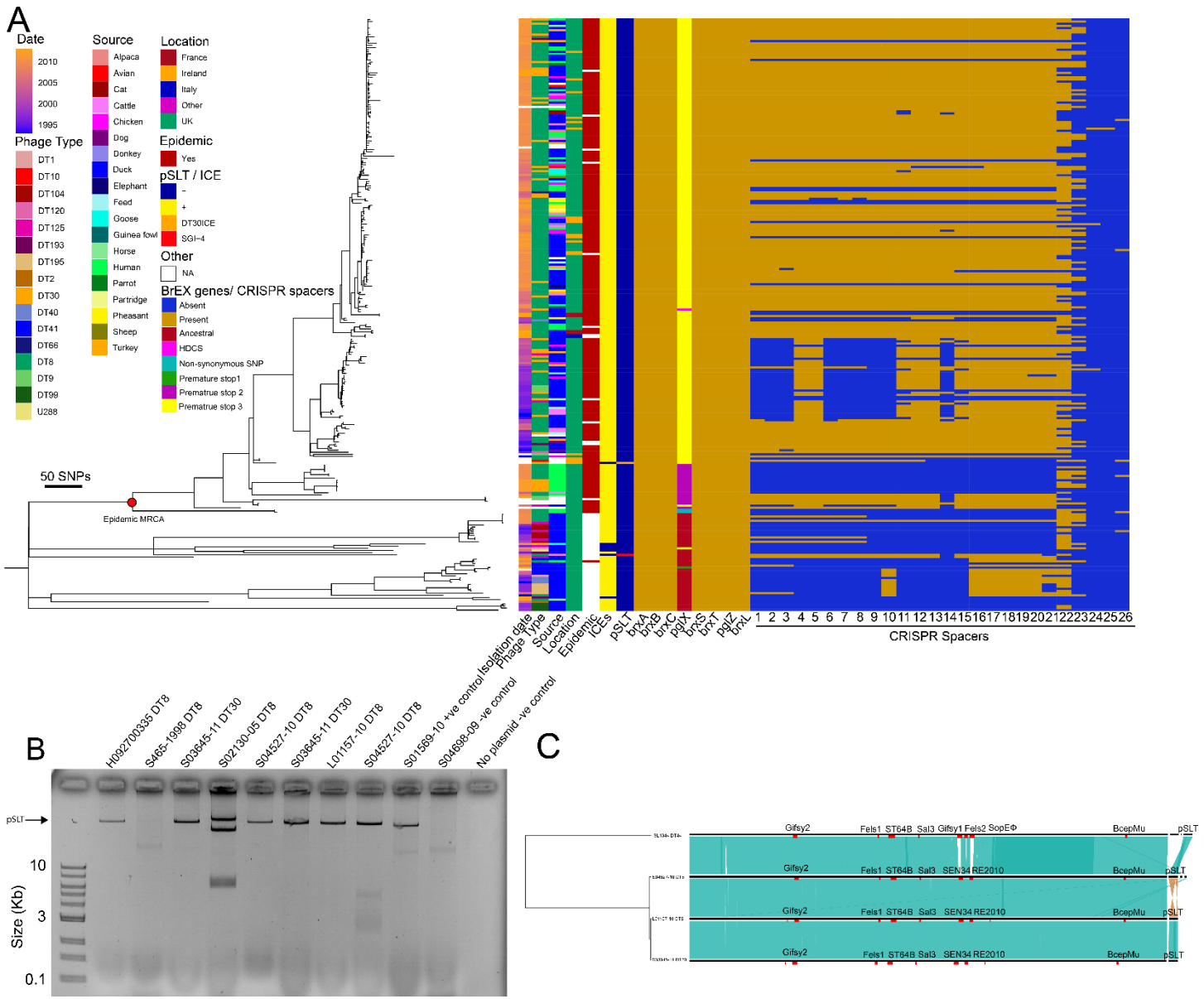

A
Date
2010
2005
2000
1995
Phage Type
DT1
DT10
DT104
DT120
DT125
DT193
DT195
DT2
DT30
DT40
DT41
DT66
DT8
DT9
DT99
U288
Source
Alpaca
Avian
Cat
Cattle
Chicken
Dog
Donkey
Duck
Elephant
Feed
Goose
Guinea fowl
Horse
Human
Parrot
Partridge
Pheasant
Sheep
Turkey
Location
France
Ireland
Italy
Other
UK
Epidemic
Yes
pSLT / ICE
-
+
DT30ICE
SGI-4
Other
NA
BrEX genes/ CRISPR spacers
Absent
Present
Ancestral
HDCS
Non-synonymous SNP
Premature stop1
Prematrue stop 2
Prematrue stop 3
50 SNPs
Epidemic MRCA
Isolation date
Phage type
Source
Location
Epidemic
ICEs
pSLT
brxA
brxB
brxC
pglX
brxS
brxT
pglZ
brxL
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26
CRISPR Spacers
B
H092700335 DT8
S465-1998 DT8
S03645-11 DT30
S02130-05 DT8
S04527-10 DT8
S03645-11 DT30
L01157-10 DT8
S04527-10 DT8
S01569-10 +ve control
S04698-09 -ve control
No plasmid -ve control
pSLT
10
3
0.1
Size (Kb)
C
Gifsy2
Fels1 ST64B Sal3 Gifsy1 Fels2 SopEΦ
BcepMu
pSLT
Gifsy2
Fels1 ST64B Sal3 SEN34 RE2010
BcepMu
pSLT
Gifsy2
Fels1 ST64B Sal3 SEN34 RE2010
BcepMu
pSLT
Gifsy2
Fels1 ST64B Sal3 SEN34 RE2010
BcepMu
pSLT

Figure VI.1 Phylogenetic relationship of duck associated *S.* Typhimurium DT8 lineage strains and identification of plasmids in DT8 and DT30 phage typed strains. (A) Maximum likelihood phylogenetic tree and associated data of 292 *S.* Typhimurium strains constructed using 6,690 recombination-purged core genome SNP sites and associated strain data, including the presence of *S.* Typhimurium virulence plasmid pSLT, ICEs, BrEX genes, and CRISPR spacers with >90 % sequence identity to the spacer array of L01157-10. (B) Plasmids extracted from 8 *S.* Typhimurium strains and separated by size on an agarose gel. (C) Whole genome alignments of *S.* Typhimurium reference SL1344, DT8 reference L01157-10, DT30 reference S03645-11, and a further DT8 strain, S04527-10.

To address the hypothesis from Mohammed et al., (543), that loss of pSLT and gain of ICE_{DT30} determined a switch in phage type from DT8 to DT30, the presence of pSLT and ICE_{DT30} in the whole genome sequence of *S.* Typhimurium strains isolated from ducks were identified using ARIBA (464). Detection of the pSLT sequence was supported by molecular evidence through extraction of plasmid DNA, separation, and visualisation by agarose gel electrophoresis, as described by Kado and Liu, (1981) (576) (Figure VI.1 (B)). The nucleotide sequence of pSLT was missing (<90 % sequence identity) from the WGS of 12/292 *S.* Typhimurium isolates, of which two were duck associated DT193 typed strains, one DT1 strain, and one DT30 strain (1/39; 2.57 %, Figure VI.1 (A)), and the plasmid was observed in reference DT30 strain S03645-11 consistent with most DT30 strains being typed as such due to other genetic determinants (Figure VI.1 (B)). ICE_{DT30} was not present in any of the strains analysed in this study despite its presence in a single DT30 strain reported previously (596). The presence of BrEX genes, CRISPR spacers, and prophage elements previously reported to contribute to phage resistance were then determined to investigate if they were involved in the observed phage resistance increased of DT30 strains. The distribution of BrEX genes and CRISPR spacers was determined through mapping and local assemblies with ARIBA (464) (Figure VI.1 (A)). The BrEX system was present within all our isolates at 4.7Mbp (with respect to a linearized chromosome with *thrA* as ORF1) adjacent to a prophage integrase, DNA damage response DNA helicase protein encoding *recD2*, and a restriction endonuclease. This was a type one BrEX system, similar to that described by Goldfarb et al., (2015)(577), but with differences from the typical genetic organisation due to insertion of two hypothetical protein encoding

genes between *pglZ* and *pglX* (577). *PglX* had two stop codons in strains of the DT8 lineage, which was clade specific (Figure VI.2). There was no association between phage type and CRISPR spacers or BrEX gene differences (Figure VI.1 (A)).

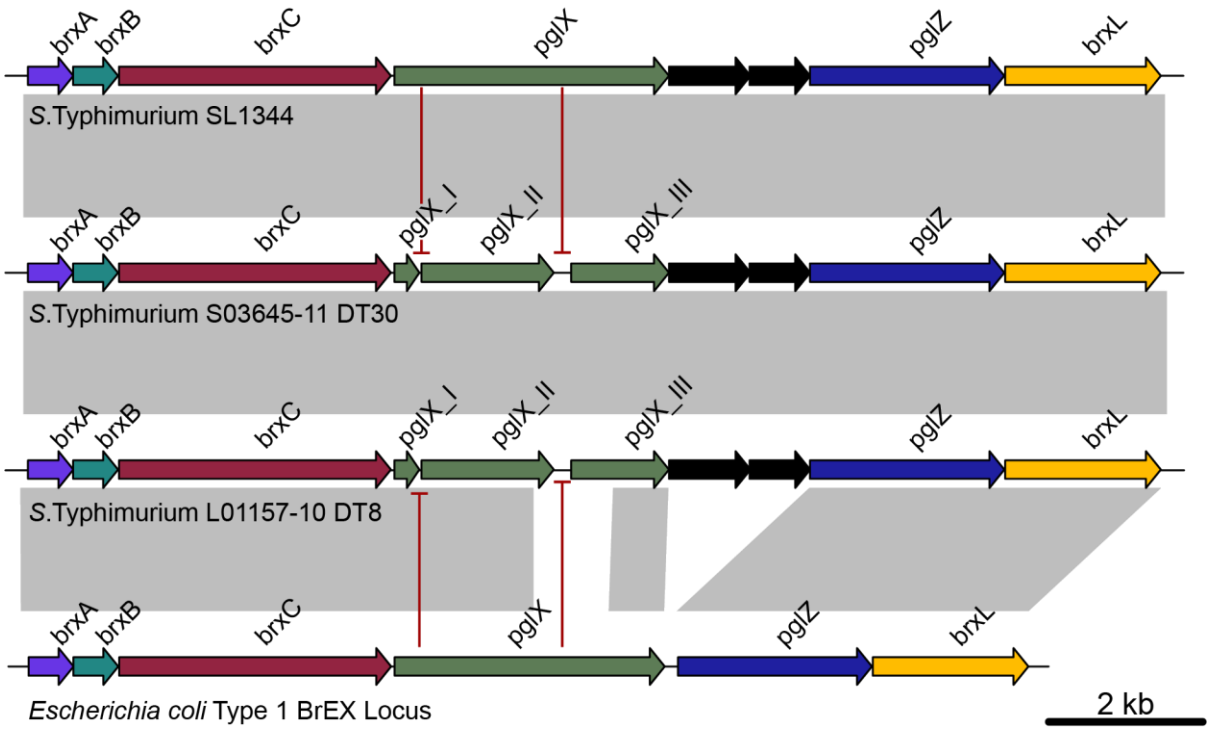


Figure VI.2. Genetic organisation and direction of Bacteriophage exclusion (BrEX) system loci BrEX genetic organisation differences between a typical type 1 BrEX system from *Escherichia coli*, compared with the loci discovered within reference *S.* Typhimurium SL1344, DT8 L01157-10, and DT30 S03645-11. Vertical red lines indicate location of stop codons.

Prophage differences were investigated in 3 long read WGSs of reference strains, two DT8 and one DT30, and compared to *S.* Typhimurium reference strain SL1344 (Figure VI.1 (C)). Prophages were identified by PHASTER (549), and regions of genomic similarity identified as previously described (39). There were no differences in prophages between our DT8 and DT30 reference strains, consistent with prophage differences not determining phage type.

VI.2 The DT8 phage sensitivity profile is ancestral with potentially reversable transitions to DT30

The ancestral state with respect to the DT8 and DT30 phage sensitivity profiles were assessed within a phylogenetic tree of 162 DT8 strains and 34 DT30 strains reconstructed using 2,297 core genome SNP

sites. A likely history of changes from the phage sensitive to phage resistant phenotype (DT8 to DT30) was determined to investigate the direction of change of phage sensitivity profiles. Due to ambiguities in the literature over the performance of ancestral reconstruction methodologies (550), two complimentary methods were employed, a maximum-likelihood method and a Bayesian MCMC method (Figure VI.4 (A, B, and C)). Bacterial propensity for recombination, horizontal gene transfer, and phase variation can change phenotypes at varying rates. Therefore, it was reasoned that appropriate transition rate matrices (Q) would require testing for significance. The sporadic distribution of DT30 amongst DT8 isolates provided a further challenge when determining which model is closest to the true history of changes, as different values for Q generated disparate models. To determine significance, likelihood ratio tests (LRT) were conducted between nested models (578). LRT tests between models using an equal rate Q and between a tip data estimated MCMC distribution for Q suggested the extra parameter of a different transition rate for the model was significant (p=1.88x10^{-15}) (Figure VI.3). A significant LRT test between the models suggests that the transition rate between DT8 and DT30 is not equal, and that transition of DT8 to DT30 occurs at a faster rate than DT30 to DT8.

To investigate whether the highest likelihood model where phage type transition rates were unequal was significant due to increased sampling of DT8 strains, rather than based on phylogenetic structure, permutation tests were undertaken. Twenty permutations were undertaken and the Mann-Whitney-Wilcoxon test used to assess if the probability distributions of each node being DT30 from each permutation were significantly different to the real data (Figure VI.4 (E)). A total of 18 permutations out of 20 were not significantly different from the data, indicating 90 % confidence that the observed data did not occur by chance.

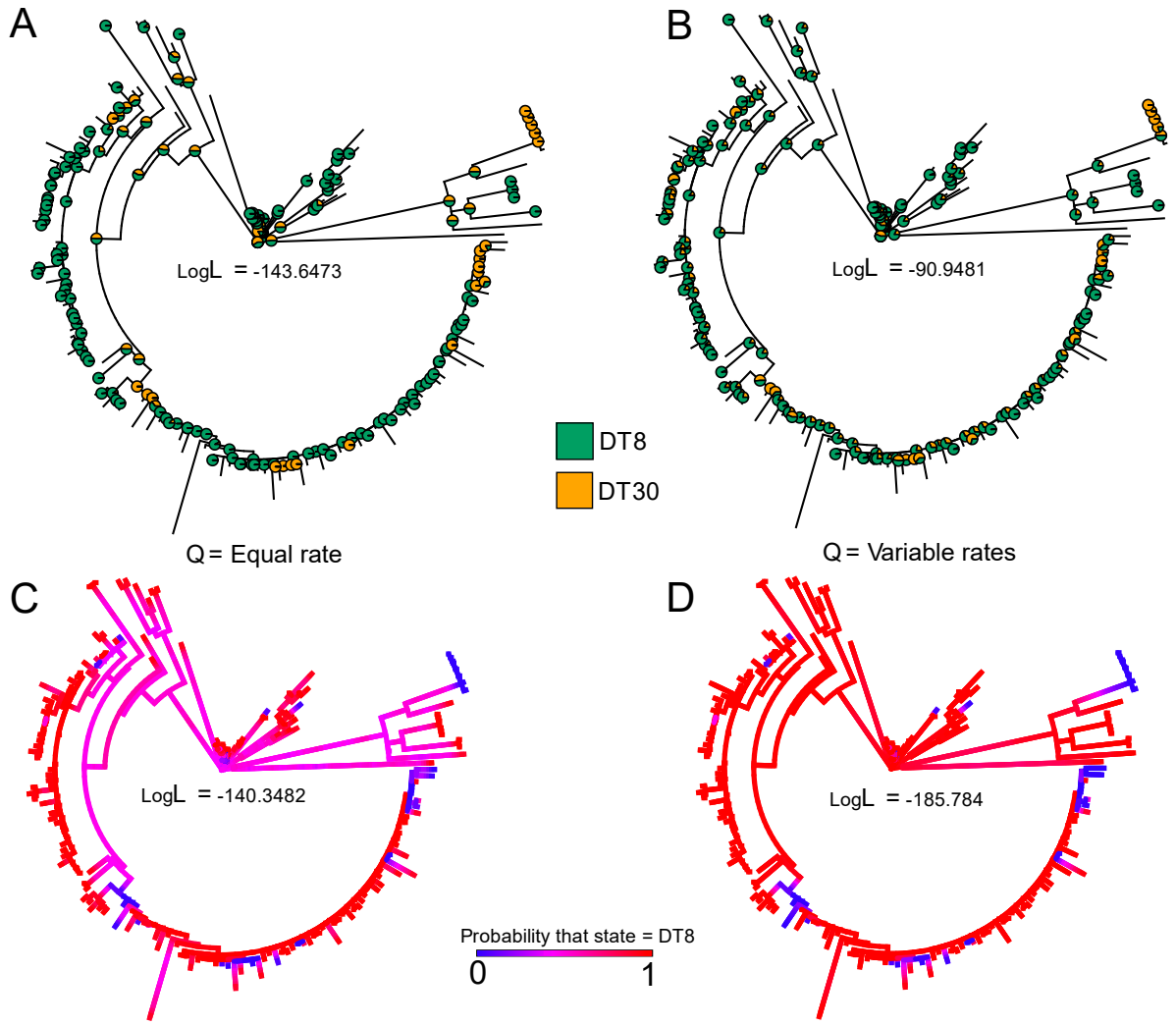


Figure VI.3 Model and method selection for ancestral history estimation of DT8 and DT30 typed strains.

Four representative examples of outcomes from various model tests. (A) and (B) show maximum-likelihood generated probability distributions as pie charts at each node. Each node represents a common ancestor within the history of the lineage, with the root node in the centre of each phylogeny. The probability of each ancestor having a resistant DT30 phenotype (orange circle), and phage susceptible DT8 phenotype (green circle). (C and D) The probability of ancestral states being DT8 (red) or DT30 (blue) across branches and at each node, as determined from a distribution of 1000 trees sampled from 100,000 iterations of MCMC. (C, pink and D, pink) Probability tends toward 50 %, displaying uncertainty, (C, blue and D, blue) probability of ancestors being DT30. (C, red and D, red) Probability of ancestors being DT8. (A) and (C) were determined by restricting Q to equal transition rates between

phenotypes. (B) and (D) allowed different rates of transition. Log-likelihoods are shown in the centre of each phylogenetic tree.

An estimate of the probability distribution at each node was undertaken using maximum posterior probabilities (MPP) (Figure VI.4 (A)). The quantity of state changes tracking back to each node from tips was elucidated (Figure VI.4 (B)), giving an indication of the quantity of state changes from root to tip. The root node was ancestral to 84 tips which had no changes in state to DT30, with 343 nodes (including tips) estimated DT8 and 76 estimated DT30. There were 45 predicted phage resistance changes across the phylogeny including 26 predicted from DT8 to DT30 and 19 predicted from DT30 to DT8. Similarly, MCMC analysis identified a mean average of 39.659 changes in phenotype state between DT8 and DT30. The mean posterior values for Q were:

$$Q = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix} = \begin{pmatrix} -10.69519 & 10.69519 \\ 14.19633 & -14.19633 \end{pmatrix}$$

where:

$$\alpha = r(DT30 \rightarrow DT8)$$

$$\beta = r(DT8 \rightarrow DT30)$$

These changes of state included an average of 25.754 switches from DT8 to DT30, and 13.905 for DT30 to DT8, which corresponds to the posterior values for rates of change between the two states. A common theme between analyses was at least one resistance phenotype state change from DT8 to DT30, back to DT8, and vice versa (Figure VI.4). Changes of phage type from DT8 to DT30, and back to DT8 is consistent with a naturally complementable or reversable mechanism producing the resistant DT30 phenotype potentially occurring throughout the population. This may include excision and integration of the region within single strains or frequent loss and transfer. Despite different approaches the outcomes of resulting models were correlated, (Figure VI.3 (D)) with an R^2 of 0.675 between probability outcomes of each common ancestor being DT30 from MPP approximation and MCMC stochastic mapping. Best fit models of the evolutionary history provided the least phenotype changes across branches, suggesting that if ancestral history estimates are incorrect it is more likely the changes of state were underestimated.

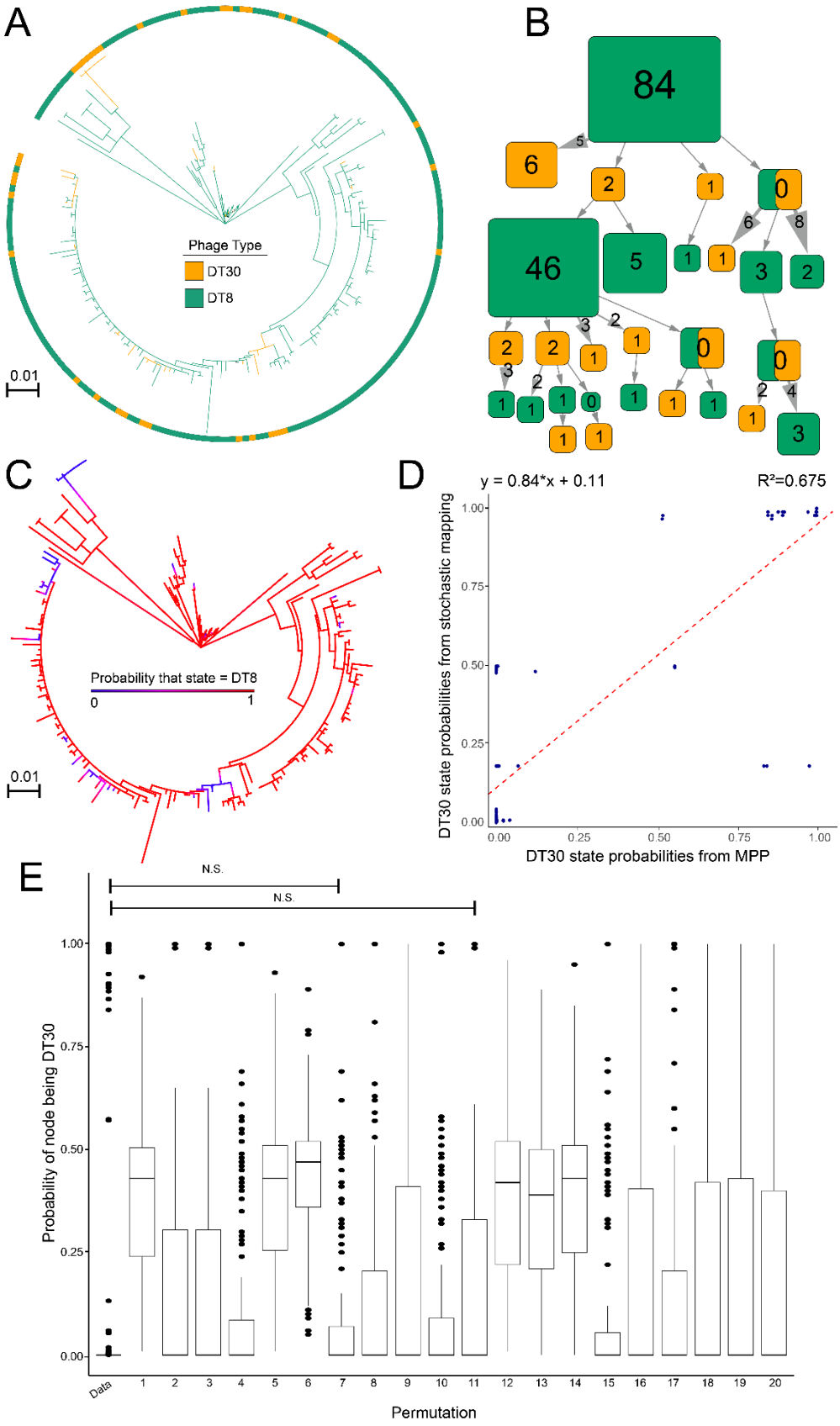

A
Phage Type
DT30
DT8
0.01
B
84
46
C
Probability that state = DT8
0
1
0.01
D
y = 0.84*x + 0.11
R²=0.675
DT30 state probabilities from stochastic mapping
DT30 state probabilities from MPP
E
N.S.
N.S.
Probability of node being DT30
Data
Permutation

Figure VI.4 Most likely models for ancestral states across a maximum-likelihood phylogenetic tree of DT8 and DT30 phage typed *S.* Typhimurium. (A) Maximum posterior probability (MPP) ancestral states across branches and ancestral bifurcating nodes. (A, orange) Character state being phage resistance DT30. (A, green) Character state being phage susceptible DT8. (A, black) Character state undetermined due to equal probabilities. (A, outer ring) Phenotype of the strain at corresponding tips. Scale bar displays SNPs per site in corresponding branch lengths. (B) Collapsed branch version of (A), leaving nodes and tips, and showing changes of state across the phylogeny of DT8 (green) and DT30 (orange). Square size indicates the number of tips tracking to that node without a change of state, a larger square representing a node with more tips tracked to it unchanged. The number within each square demonstrates how many tips track back to a particular node, 0 indicating a node having split probability between DT8 or DT30. (C) Probability based ancestral estimation and most significant model of the data generated from a distribution of 1000 samples from 100,000 generated ancestral histories estimated using stochastic mapping via Markov Chains with Monte-Carlo (MCMC) sampling based SIMMAP (554). (C, red) Probability toward DT8. (C, blue) probability toward DT30. (C, pink) Equal probabilities. (D) Linear regression of the probability that each ancestor in the tree is DT30 from maximum posterior probability estimates or MCMC estimates. (E) Probability of all nodes being DT30 in 20 stochastic mapping permutation tests. Each permutation is significantly different from the real data unless stated (N.S = $p > 0.05$). Significance determined by the Mann-Whitney-Wilcoxon test.

VI.3 Polymorphisms in the O-antigen polymerase encoding *wzy* are significantly associated with phage resistant DT30 phenotype

A GWAS approach was undertaken to identify genetic traits associated with the DT8 or DT30 phage sensitivity profiles. An anticipated limitation was the small dataset size, which was likely to generate low Kmer association probabilities due to reduced statistical power. All Kmers were kept that had any association with each phenotype, except those associated with the phylogenetic signal. A total of 2,617 Kmers were identified that met these criteria. Kmers mapped to four locations of the chromosome sequence of DT8 strain L01157-10. Two hundred Kmers mapped to *yjbR* that encodes an uncharacterised DNA-binding domain containing protein. Two-thousand-and-fifteen Kmers mapped to

wzy which encodes Wzy-Wzx dependant O-antigen polymerase of *S.* Typhimurium and α 1-2 linked O-unit containing *S. enterica* serovars of serogroup B (579). Two-hundred-and-one Kmers mapped to *fucO* which encodes lactaldehyde reductase, but without statistical significance ($p > 0.001$). A further 201 Kmers mapped to *gss* that encodes a bifunctional glutathionylspermidine synthetase/amidase, similarly without statistical significance ($p > 0.001$). All Kmers in the most significant 1 % of the range of p-values ($p < 0.001$) mapped to *wzy* and a hypothetical, uncharacterised upstream gene that was designated *wjx* for '*wzy* juxtaposed' (Figure VI.5 (B)). The potential significance of *wzy-wjx* in affecting sensitivity to phage was recognised through examining the dataset with Kmer-based GWAS approach that assembles Kmers for understanding polymorphism structure, DBGWAS (Table VI.2) (471). The largest *q*-value for a genomic region was determined to be *wzy-wjx* through aligning significant Kmers (Wald test $p<0.05$) to reference proteomes using BLASTx (488). These data are consistent with differences in the *wzy-wjx* locus resulting in DT8 or DT30 phage types.

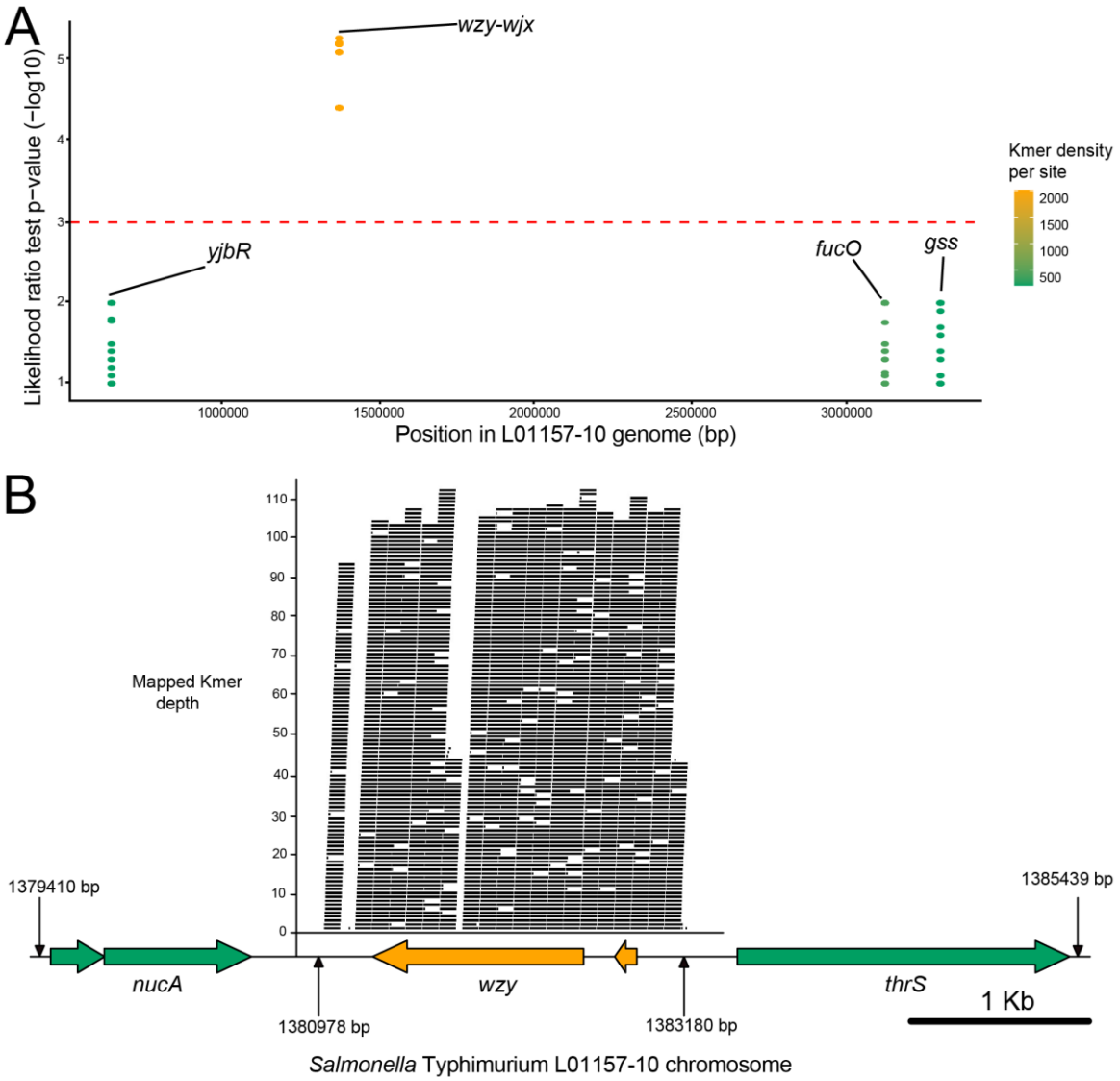


Figure VI.5 Significant Kmers from genome-wide association mapped to the chromosome of *S.* Typhimurium DT8 reference strain L01157-10. (A) Kmer positions and likelihood ratio test (LRT) probability values of significant Kmers mapped to chromosomal locations of *S.* Typhimurium DT8 reference strain L01157-10. (B) Location and depth of Kmers mapping to the *wzy-wjx* locus.

Region	Node numbe	Allele frequency	Pheno0	Pheno1	Significant	q-value	Effect estimation	Wald stat	Length
wzy locus	n4947	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	338
wzy locus	n1180	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	1253
wzy locus	n2899	156	128/139	28/29	Yes	1.00E+00	1.18E-01	-1.0382	96
wzy locus	n4640	156	138/139	18/29	Yes	4.27E-11	-7.95E-01	8.28351	58
wzy locus	n852	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	128
wzy locus	n2002	156	138/139	18/29	Yes	4.27E-11	-7.95E-01	8.26644	61
wzy locus	n7458	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	87
wzy locus	n5331	156	138/139	18/29	Yes	4.27E-11	-7.93E-01	8.25461	61
wzy locus	n5431	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	116
wzy locus	n1205	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	265
wzy locus	n3322	157	138/139	19/29	Yes	1.26E-09	-7.80E-01	7.62715	265
wzy locus	n7710	156	137/139	19/29	Yes	4.44E-08	-7.02E-01	7.0156	61
wzy locus	n8591	156	138/139	18/29	Yes	4.27E-11	-7.93E-01	8.24864	61
wzy locus	n5150	133	120/139	13/29	Yes	5.25E-05	-4.90E-01	5.24164	61
wzy locus	n5118	119	107/139	12/29	Yes	2.09E-04	-5.98E-01	4.8514	61

Table VI.2 Assembled significant Kmers and associated data from De Bruijn graph-based bacterial genome-wide association. Tabulated results from DBGWAS (471) when associating Kmers from WGS of DT8 and DT30 strains with either phage type.

VI.4 *Wzy* is variable in the DT8 lineage

To investigate the polymorphic structure of *wzy* within the DT8 lineage the *wzy* locus was extracted from 162 DT8 and 34 DT30 strains and examined in relation to their phylogenetic relationship (Figure VI.6). A reduced read depth indicating a deletion was visible in a 2,202bp locus of 14 strains, 13 of which exhibited DT30. WGS assemblies of strains with this genotype had 259bp between *wzy*-*wjx* flanking genes *nucA* and *thrS*, consistent with a base specific mechanism deleting the region such as recombination. Two DT30 isolates had a stop codon introduced within the *wzy* gene, and a further strain encoded an amino acid change from lysine to asparagine, hypothetically disrupting *wzy* function. Twelve isolates had a reduced read depth in the innermost 500bp section of *wzy*, although these isolates were of the susceptible DT8 phenotype, consistent with common heterogeneity in the region. Reference DT30 strain S03645-11 exhibited a frameshift in *wzy* hypothetically resulting in a premature stop codon.

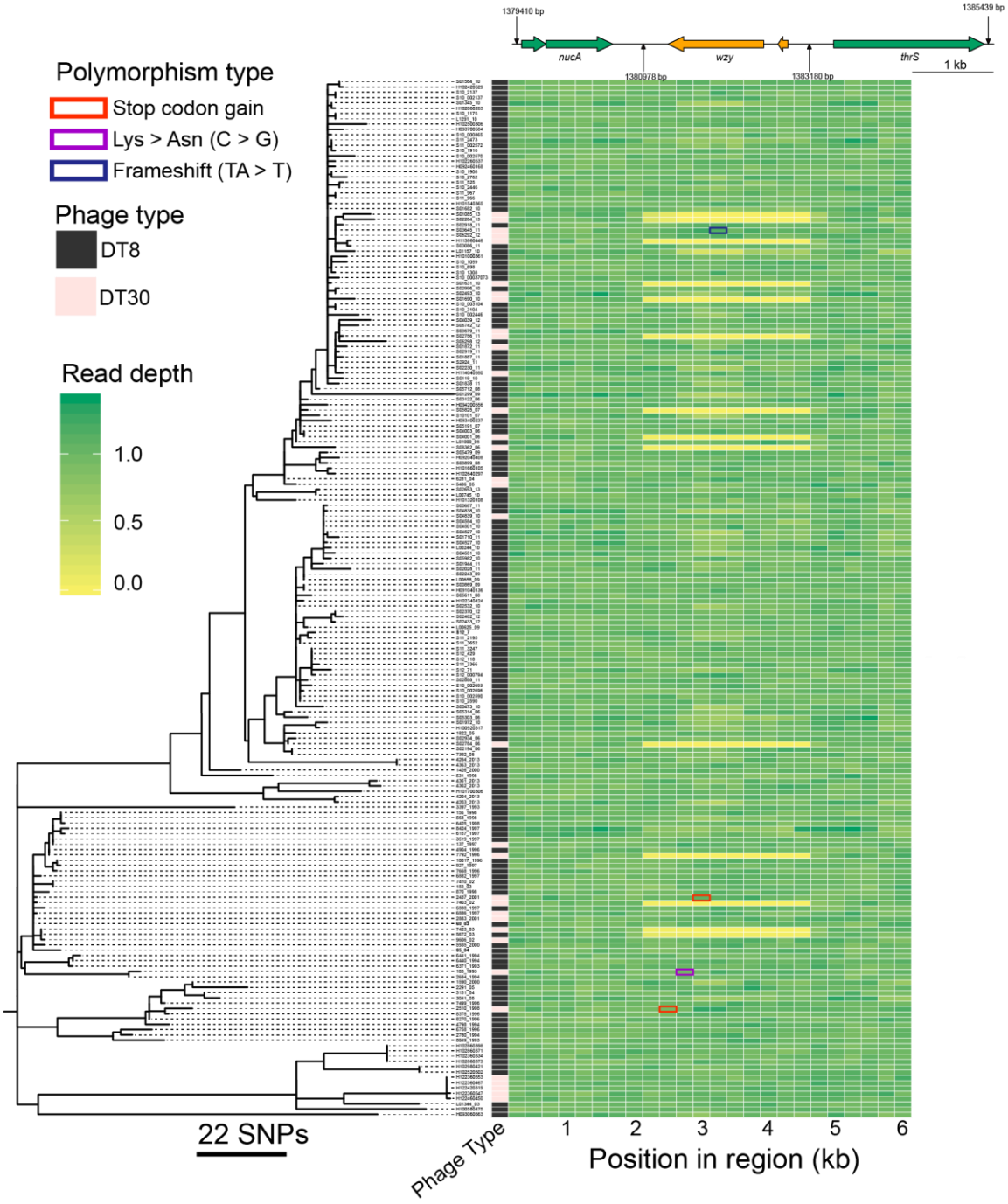


Figure VI.6 Sequence polymorphisms in the *wzy* locus of *S.* Typhimurium DT8 clade β2. Maximum-likelihood phylogenetic tree constructed using 2,297 core genome SNP sites from 162 DT8 strains and 34 DT30 strains. The heatmap displays reads mapped to the *wzy-wjx* locus of DT8 reference strain L01157-10 in 250bp bins. Boxes with coloured outlines denote the location of polymorphisms hypothetically disrupting Wzy function.

VI.5 The *wzy*- genotype is selected for by phage predation

To investigate if deletion of *wzy* can be selected due to phage predation, cultures of DT8 and DT30 strains were challenged with Anderson typing scheme phage preparations 8, 10, 18, 20, 29 and 32. First, phage typing, as described in the PHE phage typing protocol, was undertaken to ensure our reference strains displayed the expected phage sensitivity (Figure VI.7 (A)). Both wild type DT8 strains were susceptible to typing phages 8, 10, 20, 29, and 32 (Figure V.7 (A)). Wild type DT30 strain S03645-11 was susceptible to only typing phage 8 (Figure VI.7 (A)). Both outcomes are consistent with the Anderson phage typing scheme. Phage resistant colonies of DT8 were observed growing in plaques, particularly within plaques due to typing phage preparation 10 (phage 10). The colonies from the plaques were harvested and presence or absence of the *wzy-wjx* locus assessed through amplification of the region using primers that anneal just outside of a direct repeat sequence identified at distal ends of the locus. The primers were designed to polymerise both *wzy+* and *wzy-*. When polymerising the *wzy-wjx* locus, two distinct bands were visible for strains typed DT8 and DT30 after agarose gel separation (Figure VI.7 (C)), consistent with both *wzy+* and *wzy-* genotypes existing in a single colony. Both DT8 strains isolated from phage 10 plaques exhibited DT30 phage sensitivity profiles upon re-phage typing. From sequence analysis it was evident that reference DT30 strain S03645-11 was possibly typed as such due to a *wzy* frameshift mutation (Figure VI.6). However, the presence of both *wzy+* and *wzy-* genotypes within a single colony of this strain suggested that DT8 and DT30 strains undergo deletion of the region in clonal populations, even if the gene is already non-functional (Figure VI.7 (C)). To investigate the dynamics of *wzy+* and *wzy-* strains when under selection pressure from phage predation in liquid culture, the optical density of a mixed culture of L01157-10, L01157-10:*wzy-,* or S03645-11 with each typing phage was measured every 15 minutes for 24 h. This was then compared with optical density measurements when each strain was grown without phage predation selection pressure. Wild type DT8 strain L01157-10 displayed delayed lag phase when challenged with Anderson typing scheme phage preparations 8, 10, 18, 20, 29, and 32, consistent with the Anderson typing scheme, except for phage 18 which produces no lysis on agar plates (Figure VI.7 (A)). Both wild type DT30 strain S03645-11 and phage-pressure-selected L01157-10:*wzy-* displayed no lag phase extension when cultured with phages 10, 18, 20, 29, and 32, consistent with resistance to these phage preparations. S03645-11 and L01157-

10:*wzy*- both had an extended lag phase when under pressure from phage 8, also consistent with the Anderson typing scheme, and suggesting both are DT30.

Quantitative RT-PCR was used to investigate the dynamics of *wzy* genotypes under selection pressure by phage predation. The *wzy*- genotype was observed dominant at ~4 h post phage inoculum in a culture inoculated with DT8 *wzy*+ strain L01157-10, and the mutant genotype was dominant in the reaction mixture at 24 h. There were no significant differences between relative increases of *wzy+* or *wzy*- levels in the absence of phage selection, and no significant difference in relative change of *wzy*- compared to *rpoD* when a natural *wzy*- mutant was grown with typing phage 10, consistent with this genotype providing phage resistance (Figure VI.7 (B)). Absolute quantification of *wzy*- and *wzy+* of wild type L01157-10 grown in LB broth overnight suggested a copy number ratio of ~1:500 respectively (Figure VI.7 (D)). This is consistent with the observed deletion of *wzy-wjx* occurring without strong selection in a small proportion of the population, giving rise to cells that are naturally *wzy-.*

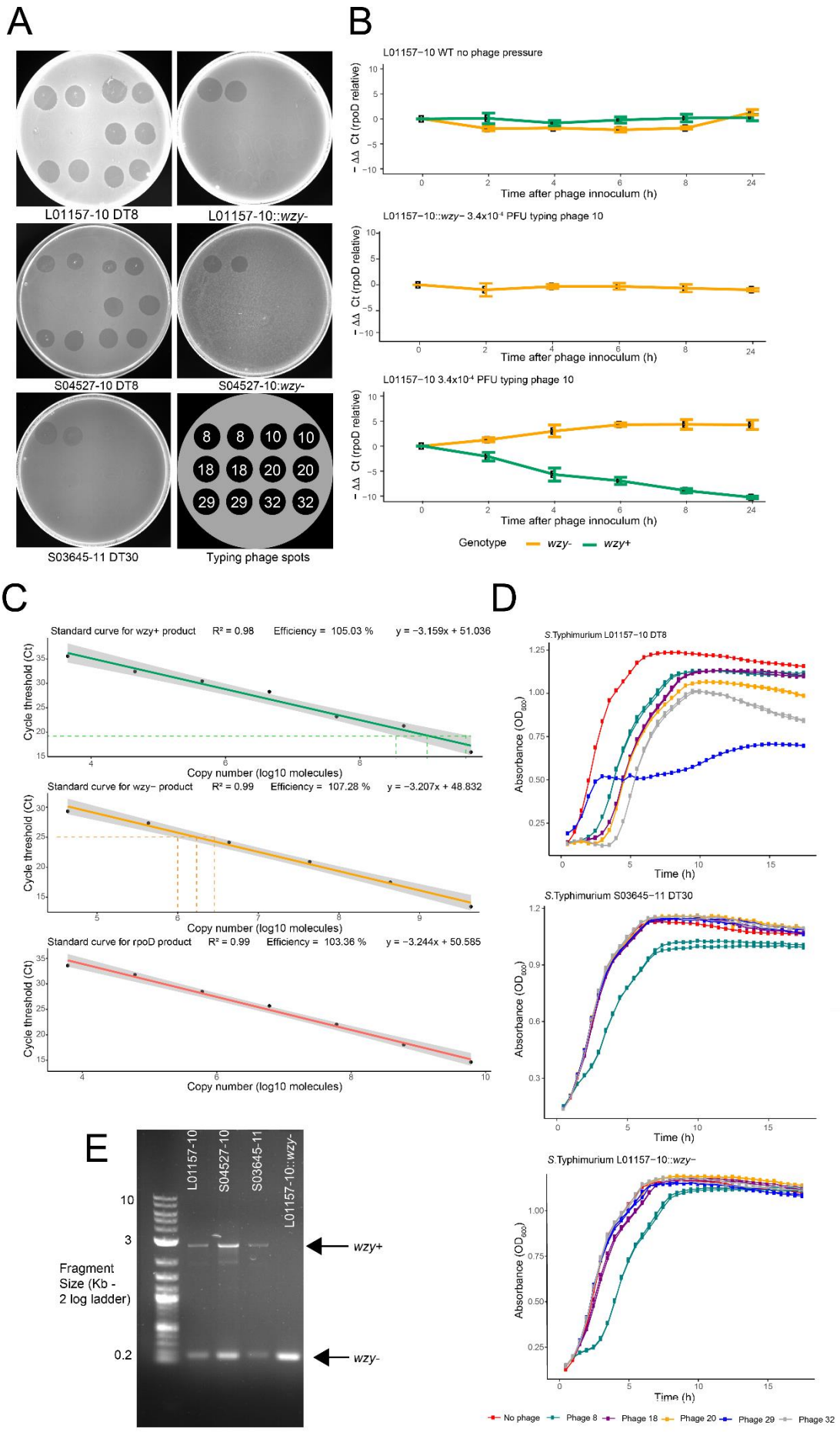

A
L01157-10 DT8
L01157-10::wzy-
S04527-10 DT8
S04527-10:wzy-
S03645-11 DT30
Typing phage spots
8 8 10 10
18 18 20 20
29 29 32 32
B
L01157−10 WT no phage pressure
L01157−10::wzy− 3.4x10⁻⁴ PFU typing phage 10
L01157−10 3.4x10⁻⁴ PFU typing phage 10
- ΔΔ Ct (rpoD relative)
Time after phage innoculum (h)
Genotype wzy- wzy+
C
Standard curve for wzy+ product R² = 0.98 Efficiency = 105.03 % y = −3.159x + 51.036
Standard curve for wzy− product R² = 0.99 Efficiency = 107.28 % y = −3.207x + 48.832
Standard curve for rpoD product R² = 0.99 Efficiency = 103.36 % y = −3.244x + 50.585
Cycle threshold (Ct)
Copy number (log10 molecules)
D
S.Typhimurium L01157−10 DT8
S.Typhimurium S03645−11 DT30
S.Typhimurium L01157−10::wzy−
Absorbance (OD600)
Time (h)
No phage Phage 8 Phage 18 Phage 20 Phage 29 Phage 32
E
L01157-10
S04527-10
S03645-11
L01157-10::wzy-
10
3
0.2
Fragment Size (Kb - 2 log ladder)
wzy+
wzy-

Figure VI.7 Phenotyping *wzy+* and *wzy-* strains with pressure from Anderson typing scheme phage preparations. (A) Phage typing of DT8, DT30, and isolated *wzy-* selection mutants of *S.* Typhimurium. (B) Quantitative PCR of *wzy+* and *wzy-* strains of *S.* Typhimurium with and without pressure from Anderson phage typing scheme phage preparation 10. (C) Quantitative PCR standard curves of dilutions of amplified DNA of 214bp between *nucA* and *thrS* for *wzy-*, 207bp inside of *wzy* for *wzy+*, and 154bp of housekeeping gene *rpoD*. Horizontal lines indicate inferred quantity of genotypes in an overnight culture of L01157-10. (D) Liquid culture growth assessment of *wzy-,* DT8, and DT30 *S.* Typhimurium strains when under pressure from Anderson typing scheme phage preparations 8, 10, 18, 20, 29, and 32. (E) PCR products from polymerising the *wzy* locus of reference DT8 strain L01157-10, reference DT30 strain S03645-11, and L01157-10:*wzy-* after selection with phage pressure from plaques in (A).

VI.6 Deletion of *wzy* is reversable within a mixed population containing the wild type gene

The *wzy-wjx* locus was predicted to be a cis-integrating mobilizable element (CIME) due to *in silico* observation of a base specific deletion consistent with recombination, identification of a pair of 112bp identical nucleotide repeats with one at each end of the locus, and no predicted genes for transfer. CIMEs utilise other genetic element's transfer machinery such as prophage integrases for excision and integration or T4SSs for cell to cell transfer (580, 581). To test this hypothesis a donor strain was constructed containing a selectable chloramphenicol resistance gene (*cat*) introduced via allelic exchange into *wjx* of L01157-10, and the insertion location confirmed by PCR (primers in Appendix II). To construct a recipient strain, a kanamycin resistance cassette (*aphII*) was inserted as a knock-in into an intergenic region upstream of *iciA* in DT8 reference strain L01157-10 and the location of insertion confirmed by PCR (primers in Appendix II). A derivative of L01157-10:*aphII* in which the *wzy* gene had been deleted was selected by culture with phage 10. Strains with the *wzy-* genotype were isolated and confirmed through amplification using primers flanking the deletion as previously described (Figure VI.7 (E)). To exclude the possibility that the L01157-10:*aphII* gene was transferred to the donor strain, a second mutation conferring resistance to nalidixic acid was selected by culturing L01157-10:*wzy-,aphII* on selective media supplemented with 30 μg/ml nalidixic acid. Since determinants of resistance to kanamycin and nalidixic acid were not genetically linked, it was reasoned that they would need to occur

together in a single strain, and that this was highly improbable. The frequency of nalidixic acid resistance development for L01157-10:*aphII* was ~1 in $5x10^7$ CFU/ml. Next, L01157-10:*aphII, nal*^{R} was grown in mixed culture for 24 h with L01157-10 Δ *wjx*::*cat* before plating on agar plates with selection for the donor (25 μg/ml chloramphenicol), the recipient (30 μg/ml nalidixic acid and 50 μg/ml kanamycin), and possible *wzy* locus transductants (25 μg/ml chloramphenicol, 30 μg/ml nalidixic acid and 50 μg/ml kanamycin). The frequency of cells showing resistance to nalidixic acid, chloramphenicol, and kanamycin after mixed culture was 10^{-9} per recipient cell (Figure VI.8 (C)). This was significantly increased (p<0.01) when 0.5 μg/ml of prophage inducing, DNA intercalating molecule mitomycin C was added to the culture (Figure VI.8 (C)), consistent with transduction being the method of *wzy* transfer. Therefore, the probability that generalised transduction could transduce the kanamycin cassette was ~$1/10^9$ without phage induction, and ~$1/10^7$ when supplemented with mitomycin C. The probability of spontaneous nalidixic acid resistance was estimated at $5/10^7$. Together, these suggest the probability of both kanamycin transfer through generalised transduction and spontaneous nalidixic acid resistance occurring within a single cell was $5/10^{16}$ without mitomycin C, and $5/10^{14}$ with mitomycin C, much greater than the estimate for probability of chloramphenicol cassette transfer ($5/10^9$ without mitomycin C and $5/10^7$ with mitomycin C). This provided confidence that emergence of resistance to chloramphenicol, nalidixic acid, and kanamycin was conferred through the *cat* gene being transferred to the recipient. To investigate the genotypes of mutants resulting from mixed culture and triple antibiotic resistance the *wzy-wjx* locus was amplified using primers annealing in adjacent genes; 207 bp of *wzy* was amplified, and 433 bp of *cat* was amplified in 17 isolated mutant colonies (Figure VI.8 (B)). Eleven of the 17 had DNA amplified consistent with presence of the *wzy-wjx* locus, 15 of 17 potential *wzy* revertants displayed amplified DNA consistent with presence of *wzy*, and 16 of 17 displayed amplified DNA consistent with presence of *cat* (Figure VI.8 (B)). Two revertants of *wzy* (revertant 14 and revertant 15) were subject to re-phage typing, and both displayed altered phenotypes compared to WT L01157-10, but also different plaque morphologies compared with L01157-10:*wzy*- (Figure VI.8 (D)). To investigate the genotypes in more detail, L01157-10 Δ *wjx*::*cat* and *wzy* revertant numbers 1, 2, 8, 9, 12, and 15 were subject to WGS, core genome SNP-based phylogenetic relationship reconstructed, and *wzy-wjx* locus extracted and viewed against their phylogenetic relationship (Figure VI.8 (E)). The WGS

of strains proposed to have *wzy* and *cat* transferred into them exhibited *wzy* and *cat* inserted into the same locus as the original donor strain, discordant with PCR results for revertants 2, 8, and 12 (Figure VI.8 (B)), possibly indicating PCR errors, but consistent with *wzy* being transferred between the strains. Each revertant also had *aphII* in the expected intergenic region upstream of *iciA*, and a mutation in *gyrA* resulting in aspartic acid substitution with tyrosine at amino acid position 87. These data suggest that the *wzy* locus is transferable within a mixed genotype population of a single strain, but possibly does not restore the original phenotype, although a SNP within *rfaL* was also observed, suggesting the lack of phenotype restoration (data not shown).

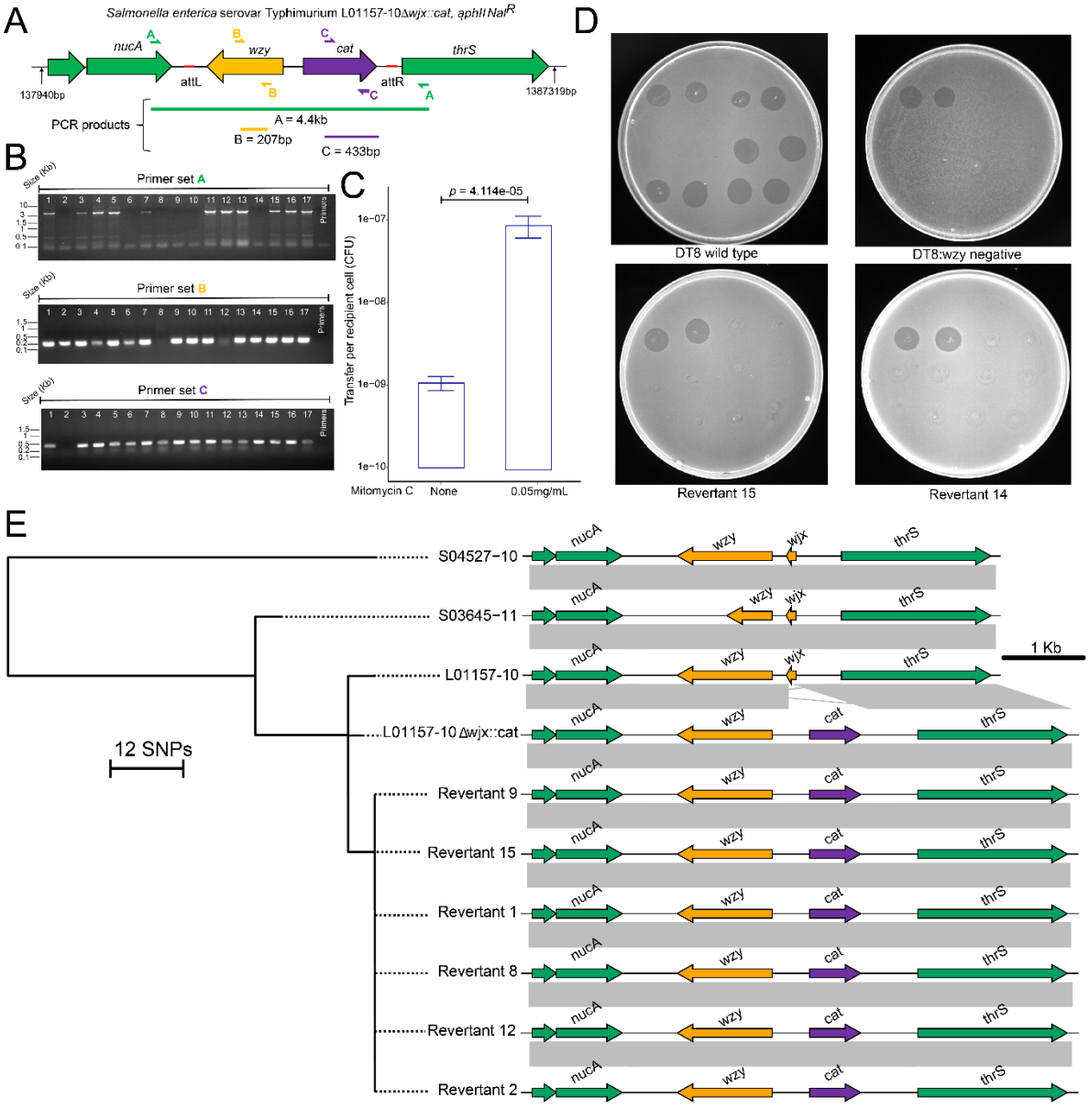
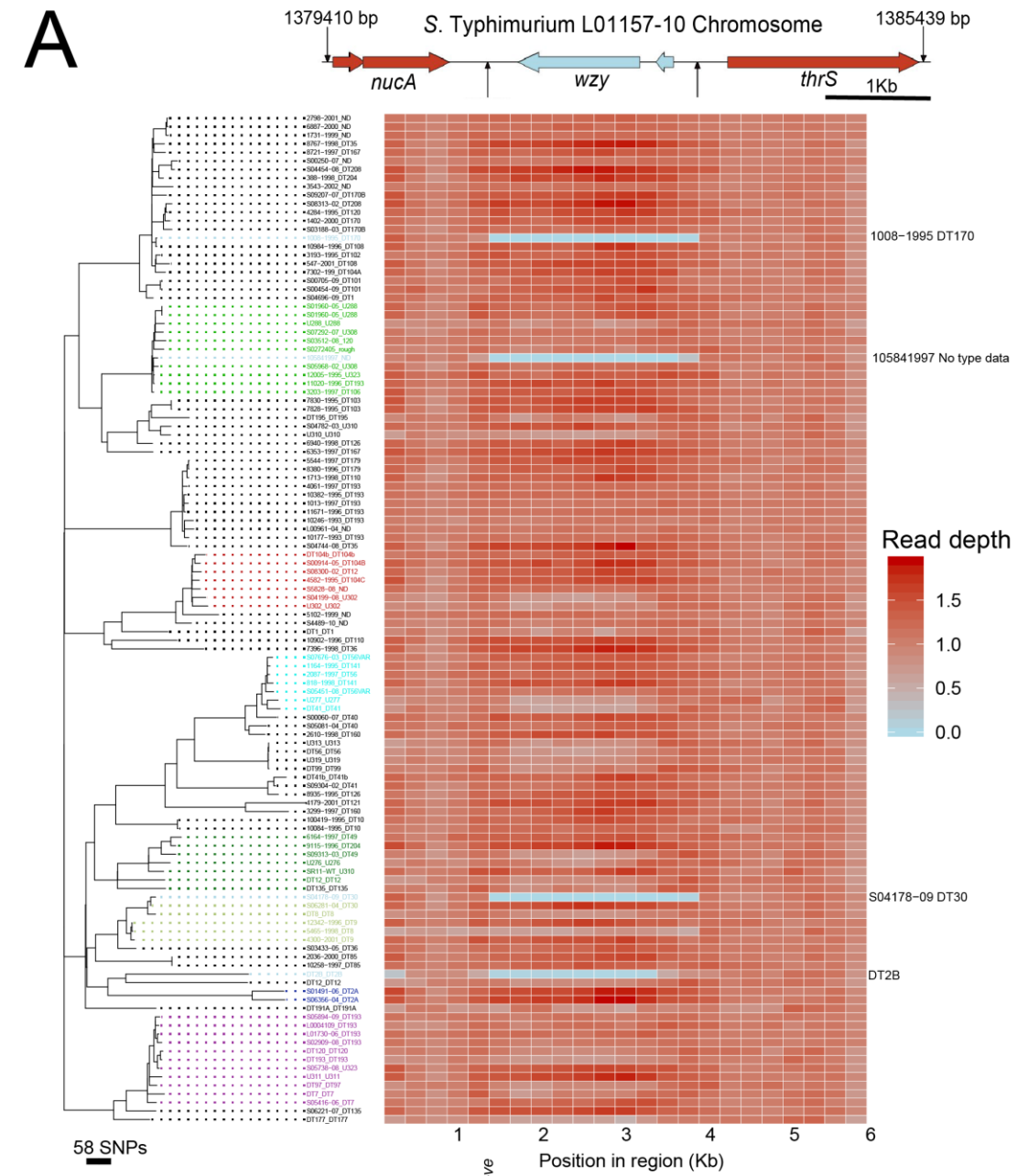


Figure VI.8 Transfer of the *wzy* locus within a mixed genotype population of a single strain. (A) Genetic organisation and direction of genes in the *wzy* locus, including position of chloramphenicol resistance gene (*cat*), and length of PCR products when amplifying targets using primers pairs for: the whole *wzy* locus with primers in flanking genes (A, A); 207bp of *wzy* (A, B), and; 433bp of *cat*. (B) Agarose gel separated DNA products from amplifying regions of17 *wzy+* revertant colonies as shown in (A), using primer sets A, B, and C. (C) *Wzy* locus transfer frequency calculated as transfer per recipient cell with and without DNA damage-inducing mitomycin C. (D) Representative phage typing of DT8 wild type, DT8 *wzy-*, and L01157-10 *wzy* revertants 14 and 15 (as in B). (E, left) Maximum likelihood phylogenetic tree of two wild type DT8 strains (L01157-10 and S04527-10), one wild type DT30 strain (S03645-11), L01157-10 Δ *wjx*::*cat*, and 6 *wzy* revertants of L01157-10::*wzy-, aphII*. (E, right). Genetic diagram of *wzy-wjx* loci from each strain or mutant as denoted by the phylogenetic tree labels.

VI.7 Deletion of the *wzy* locus has occurred sporadically throughout *S.* Typhimurium

To investigate evidence of the deletion of *wzy* in other lineages of *S.* Typhimurium, the reads from 120 population representative WGSs were mapped to the *wzy* locus of L01157-10 using Bowtie-2 and read depth determined through Bedtools (472, 473) (Figure VI.9 (A)). Four strains had *wzy* deleted from the WGSs: a porcine associated DT170 lineage strain (1008-1995), an untyped strain from the U288 linage (105841997), an avian associated reference strain phage typed DT2B (DT2B), and a duck associated DT8 lineage strain phage typed DT30 (S04178-09). The deletion was base specific as observed in DT8 and DT30 strains, beginning precisely at the predicted attL an attR site and resulting in 159bp between adjacent genes *nucA* and *thrS* in WGS assemblies of strains with the deletion. This is consistent with deletion of *wzy* occurring in strains from other *S.* Typhimurium lineages at low frequencies, and *wzy-wjx* behaving as a CIME. The observed frequency of deletion in these strains may be an artefact of the strain collection or sampling bias and more investigation is required to understand the dynamics of this locus within the *S.* Typhimurium population. Next, to investigate whether *wzy* could be deleted from clonal populations as observed for DT8 lineage strains, the *wzy* locus was amplified from 11 reference strains from across the *S.* Typhimurium population using primers annealing just outside of the direct nucleotide repeats predicted to be attachment sites for *wzy-wjx* (Figure VI.9 (B)). L01157-

10:*wzy*- was used as a *wzy*- control, L01157-10 Δ *wjx*::*cat* also polymerised to show that the locus is present, and the *wzy* locus from L01157-10 Δ *wjx*::*cat* after being challenged with phage pressure subject to amplification to identify that the locus can be deleted from this mutant. The *wzy* locus was present in all strains except L01157-10 *wzy*-, L01157-10 Δ *wjx*::*cat wzy*-, a DT2 reference strain, and SL1344. The loss of *wzy* from the two reference strains is consistent with possible phage contamination. Never-theless, the data are consistent with *wzy* deletion occurring within strains from other lineages of *S.* Typhimurium in a recombination-mediated, CIME-dependant manner.


A
1379410 bp
S. Typhimurium L01157-10 Chromosome
1385439 bp
nucA
wzy
thrS
1Kb
1008−1995 DT170
105841997 No type data
Read depth
1.5
1.0
0.5
0.0
S04178−09 DT30
DT2B
58 SNPs
1
2
3
4
5
6
Position in region (Kb)

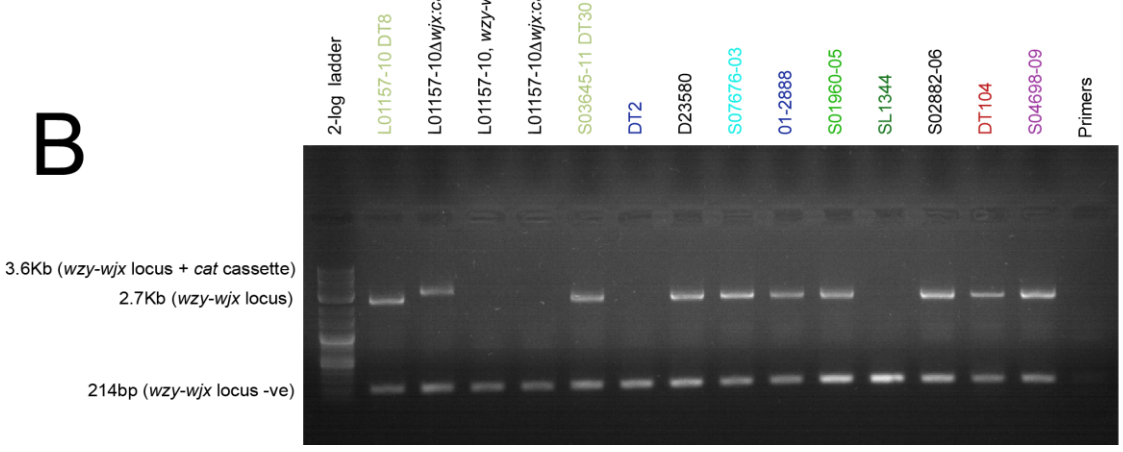

B
2-log ladder
L01157-10 DT8
L01157-10Δwjx:cat
L01157-10, wzy-wjx -ve
L01157-10Δwjx:cat, wzy-wjx-ve
S03645-11 DT30
DT2
D23580
S07676-03
01-2888
S01960-05
SL1344
S02882-06
DT104
S04698-09
Primers
3.6Kb (wzy-wjx locus + cat cassette)
2.7Kb (wzy-wjx locus)
214bp (wzy-wjx locus -ve)

Figure VI.9 Assessment of *wzy* locus in *S.* Typhimurium population representatives. (A) Maximum-likelihood phylogenetic tree of 120 strains representing the genomic diversity of *S.* Typhimurium, constructed using 17,263 SNP sites, showing the read depth of respective *wzy* loci, normalised to read depth of single copy gene *thrS*. (B) Agarose gel separated DNA products from polymerising the *wzy* locus from 11 reference *S.* Typhimurium strains, 2 phage selected L01157-10*wzy*-, and L01157-10Δ*wjx*::*cat*. Reference strain labels are coded by third order clades (B), and have respective related strains coloured (A): DT8 lineage β2 (pale green), DT2 lineage β3 (blue), DT56 lineage β5 (bright blue), U288 lineage α12 (bright green), DT204 lineage β1 (dark green), DT104 lineage α15 (red), and monophasic ST34 lineage α17 (purple).

VI.8 Chapter VI Discussion

This Chapter displayed evidence that the DT30 phage type is, in part, determined by a base specific deletion of the *wzy* locus. Initially, it was identified that DT30 strains are sporadically distributed within the DT8 lineage. Seventeen of 163 duck associated isolates (10.42 %) had a possible phage resistant genotype (Figure VI.5), as well as two of the 33 human isolates (6.06 %). The *wzy*- genotype was present in 14 of 196 epidemic isolates tested for *wzy* presence or absence (7.1 %), despite 34 of these isolates exhibiting phage type DT30 (17.35 %). Therefore, 16 isolates typed DT30 have unexplained changes producing the phenotype, despite discovery of five different *wzy* polymorphisms within 17 DT30 strains. Strains exhibiting DT30 that did not have *wzy* polymorphisms may be due to the difference in techniques of phage typing and WGS producing disparate phenotypes, or other changes that are difficult to extrapolate. The BReX system in DT8 lineage strains encoded a fragmented DNA adenine methyltansferase, *pglX* (577). Redundancy of this system may partly be determining reduced phage resistance observed in DT8 strains, possibly explaining why *wzy*- strains are more frequent in the lineage. Redundancy in one phage exclusion system will logically require other systems to complement the function and exclude phages. Ancestral state reconstruction determined that multiple switches between DT8 and DT30 occurred in the history of the DT8 lineage, and the best fit model to the data was one with the least switches between DT8 and DT30, consistent with a complementable or reversable mechanism producing DT30. Bacterial Kmer based GWAS identified polymorphisms in *wzy* as

significantly associated with DT30 strains, and evidence was generated consistent with *wzy* determining DT30 and the difference in phage sensitivities of DT8 and DT30 strains. *Wzy* deficient mutants of *S.* Typhimurium exhibit attenuated virulence raising the LD_{50} by 10^6 in mice when coupled with an *asd* deficient mutant (582). *Wzy* deficient mutants have also been shown to have reduced mouse colonisation (583). However, reduced colonisation is not to the extent of other LPS biosynthesis gene mutants, probably due to retaining one immune stimulating O-unit tetramer (583). Due to deletion of *wzy* producing attenuation for infection, the loss of *wzy* could be an evolutionary dead-end. The hypothesis formulated from ancestral reconstruction and *wzy* phenotype investigations was that the *wzy* locus was transferable between mixed *wzy* genotype strains. Therefore, for evidence that *wzy* deletion was not an evolutionary dead-end and to test the hypothesis of reversibility from ancestral state reconstruction, mixed genotype mutants were co-cultured and plated on selective media. The *wzy* gene was determined to be transferable within a mixed genotype strain population, with an increase in transfer when cultured with DNA SOS response-inducing mitomycin C. Mitomycin C has been shown to induce the lytic cycle in otherwise lysogenic prophages (584-586). Increased *wzy* transfer in the presence of mitomycin C is consistent with *wzy* transfer through generalised transduction. Despite observation of *wzy* being genotypically reversable, phage typing plates of selected strains with *wzy* reversion did not show reversion of the original phenotype. This may be due to changes introduced into the strains through the phage or nalidixic acid selection procedures, selection of LPS biosynthesis gene mutants as a by-product of recombination-based allelic exchange of the *cat* gene into the *wzy* locus, or simply evidence that reverting the gene presence does not revert the phenotype. Understanding this requires more investigation.

The distribution of the *wzy*- genotype amongst 120 representative *S.* Typhimurium isolates included a DT30 strain, a strain from a UK lineage of ST313 which may be avian adapted, a bovine and porcine associated DT170 strain, and a strain residing in the porcine adapted U288 clade (39, 91, 587). Absence of *wzy* in these WGS is consistent with the genotype being selected for in ducks and other wild avian or agricultural species. Whether there is evolutionary pressure in the duck biome from phage selecting the *wzy*- genotype more frequently than in other lineages, or perhaps lack of evolutionary pressure to

perpetuate a long, energy expensive O-chain, requires investigation. Reduced read depth of the innermost 500 bp of *wzy* is more prevalent throughout the *S.* Typhimurium population (Figure VI.5 and Figure VI.9). The reason for this reduction of read depth is undetermined.

The ability of bacteria to modify LPS is well known to change the phage host range {Labrie, 2010 #5281}(331). However, *wzy* gene diversity as a method of phage evasion has not been well described, with a distinct lack of experimental evidence of hypothesised transposable behaviour of *wzy* genes (380, 579). There are studies hypothesising the evolution of *wzy* genes in serogroup B, where the O27 antigen is established through tetramer O-unit α1-6 linkage polymerised by $wzy_{\alpha 1\text{-}6}$ that lies adjacent to an IS1617-like element in some O-antigen gene clusters of subspecies II, and 57 serovars of subspecies I (380, 579). However, *S.* Typhimurium is O27- and retains a remnant of $wzy_{\alpha 1\text{-}6}$ within the O-antigen biosynthesis gene cluster between *wbaV* and *wbaU* (380, 579). An IS element is hypothesised to have deleted the ancestral $wzy_{\alpha 1\text{-}6}$, and subsequently an α1-2 linkage encoding *wzy* has been acquired, first described in *S.* Typhimurium strain LT2 (13). WGS data analysis in this study was consistent with other literature reports - the *S.* Typhimurium $wzy_{\alpha 1\text{-}2}$ was 2.6Mb away from the O-antigen gene cluster, between *nucA* and *thrS*, with no significant sequence identity to serogroup D3 $wzy_{\alpha 1\text{-}6}$ from serovar Schleissheim. $Wzy_{\alpha 1\text{-}6}$ is hypothesised to have been deleted from serogroup B and D O-antigen biosynthesis gene clusters and re-entered serogroup D3 through recombination via an adjacent IS element (380). The data presented here is, in part, consistent with these observations, as *wzy* was identified to be genotypically reversable. *Wzy* genes exhibit functional similarity but sequence divergence, indicating convergent evolution. Gaining different *wzy* genes will produce different carbon bonds between O-units of LPS, and deleting *wzy* changes the immunogenicity of strains (583), consistent with recombination mediating antigenic variation through gain and loss of *wzy* genes.

Intra-serovar variation of the O-antigen has been described in *S.* Typhimurium, where the genes encoding chain length determination protein opvAB is subject to phase variation (383). Phasevarions of opvAB generate sub-populations of *S.* Typhimurium with $opvAB^{on}$ and $opvAB^{off}$ allowing defence against phages that utilise LPS as a receptor, but at a cost of reduced virulence. All colonies selected for by phage

predation in this study exhibited *wzy*-, consistent with the idea that, for typing phage preparation interactions, more subtle alterations of LPS were ineffective at phage defence.

If deletion of *wzy* is beneficial for fitness in circumstances other than phage defence requires further investigation. However, the energy cost of a multiple O-unit LPS is likely a significant burden for the cell. Inhibiting polymerisation could be beneficial for persistence in certain niches where a single O-unit tetramer does not reduce fitness. Evolutionarily, the common ancestor of Enterobacteriacae would have endured ~100 millions of years of co-evolution with bacteriophage, where the ability to lose and gain new O-antigen polymerase genes may be a useful and quick method of defence.

A mechanism for the deletion of *wzy* remains elusive from the data in this study, but due to the attachment sites at each distal ends of the *wzy* locus, observation of a base specific deletion, and *wzy*- being selected for in L01157-10 strains with phage pressure, a hypothesis, that the *wzy* locus is deleted and transferred in a CIME-dependant through utilising other present phage integrases, was proposed (Figure VI.10). Whether this is the case requires more investigation.

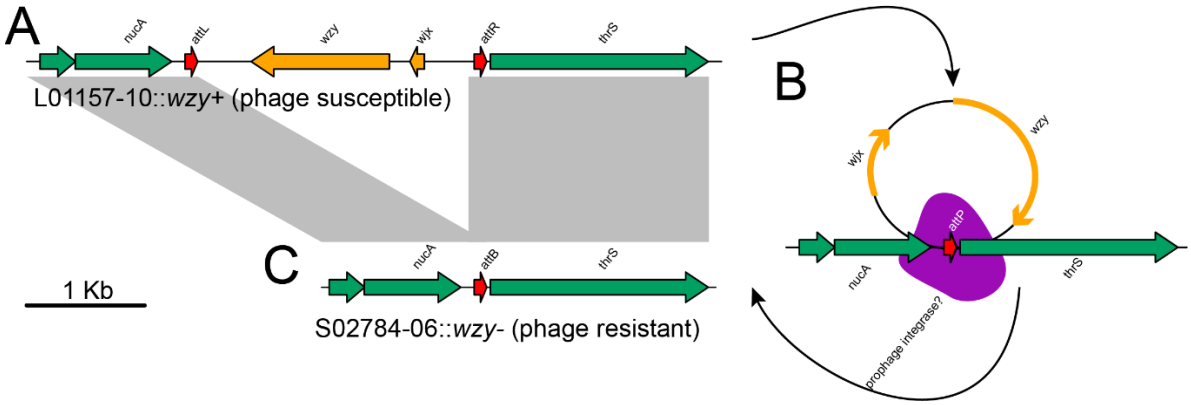


Figure VI.10 Hypothesis for phage integrase-mediated excision and deletion of the *wzy* locus from *S.* Typhimurium chromosomes. (A) Genome map of L01157-10 (a wild type *wzy+*) and (C) S02784-06 (a wild type *wzy*-) with direction and size of genetic content. (Green arrows) Chromosomal genes not within the hypothetical *wzy* locus attachment sites. (Red arrows) Attachment sites. (Orange arrows) Genes inside the proposed attachment sites. (B) Hypothetical mechanism for excision that leads to *wzy* transfer and deletion (B), showing crossover junction hypothetically mediated by a prophage integrase (purple).

Concluding remarks

S. Typhimurium has evolved alongside animals, displaying an inherent ability to evolve lineages that infect multifarious animal hosts and lineages that become pandemic clones (39). Similarly, research shows that strains of *S.* Typhimurium can occupy certain reservoirs and become host adapted (78). All *S.* Typhimurium lineages appear human pathogenic (39, 169). This equates to numerous evolutionary selection pressures: host animal microbiomes, host immune systems, environmental stressors, competing pathogens, antimicrobial agents, and phage predation. This study questioned what genomic characteristics are required for such an organism to counter these selection pressures.

From analysis of the population, two lineages of *S.* Typhimurium were identified that had phage resistance differences with respect to phage sensitivity differences observed by the Anderson phage typing scheme and potentially novel acquired genomic material, one broad-host-range, widespread lineage from clade α, and a duck associated lineage from clade β. The study characterised an acquired genomic island that was associated with initial clonal expansion of monophasic *S.* Typhimurium ST34, and due to strains showing increased resistance to copper levels found regularly in pig feed, this can be proposed as an initial stressor counteracted by a single clone that has been most frequently isolated from pigs (184, 541). It was determined to be self-transferable and enhance resistance to copper, silver, and arsenic. During the start of this study (2016) monophasic *S.* Typhimurium ST34 was recognised as an important pathogen due to several outbreaks and subsequent epidemics within Europe, hence the interest in SGI-4 and its role in initiating clonal expansion. Since then the clone has spread globally, gained resistance to last-resort antibiotic *coli*stin, and been identified as causing invasive disease in immunocompromised human hosts. Unpublished data suggests that the clone emerged initially from Europe, possibly Italy, probably from a pig farm (Tassinari et al., Unpublished data). This suggests that agricultural use of heavy metals require re-evaluation, and alternative practises need to be proposed to reduce the possibility of emergence of pathogens in this manner, such as phage-based therapeutics and improved agricultural practises. This study also identified bacteriophage resistance determinants within the most common phage type of monophasic *S.* Typhimurium, DT193, and duck associated phage type

DT8 and DT30. Through a combination of phenotypic data from the *S.* Typhimurium phage typing scheme and WGS we were able to link phenotype to genotype. The data presented a common theme for both lineages – phage predation shapes *S.* Typhimurium genome evolution. Resisting this predation appears important, not just for survival when encountering new phages in different niches, but possibly to outcompete different bacteria and other lineages of *S.* Typhimurium with different prophage repertoires. The genomic plasticity provided through introduction of new genetic material via phage transduction may also enhance fitness such as AMR genes, phage defence genes, virulence genes, or colonization factors. In the case of monophasic *S.* Typhimurium ST34, exclusion of phage is associated with acquisition of a prophage (mTmII) revealing that phage themselves require defence from other phages to stabilise their lysogenic host. Acquiring mTmII also correlated with further clonal expansion, suggesting synergistic bacteria-prophage relationships can be fundamental to evolutionary success of both.

Within the duck associated DT8/DT30 lineage of clade β, a frequent polyphyletic deletion of a genetic region encoding the O-antigen polymerase was identified. This excluded 10 phages of the Anderson typing scheme and was associated with phage type DT30. A long LPS O-antigen chain is a fundamental molecule for cell stability, protection, adhesion, and infection. Inhibiting O-chain polymerisation in one evolutionary event through chromosomal deletion is probably costly for fitness, producing attenuated infection (583). This is consistent with a trade-off between virulence and resisting phage predation. It also may be an evolutionary dead-end unless another O-antigen polymerising gene is introduced. This is evident from the literature (583). Selection pressure from phage, and possibly mammalian cellular immunity, appear to have evolved *wzy* genes that are associated with transposable elements (380). A common ancestor of *S.* Typhimurium is hypothesised to have had its ancestral *wzy* deleted and a different *wzy* gene inserted elsewhere in the genome, likely through HGT. *Wzy* genes encode protein orthologues with sequence divergence but similar function. This diversity produces different carbon position bonds between repeat units (379). If regaining the same *wzy*, the cell would exhibit the same O-antigens, but a sequence divergent *wzy* could produce a different antigen, meaning the swapping of these genes promotes antigenic variation. Varying antigens can change the dynamics of phage-bacteria interactions,

host-pathogen interactions, and produce serotype conversion, ultimately changing the biology of the lineage, potentially generating variation for continued survival in different niches.

The data presented in this study suggest phage resistance occurs rapidly and is a requirement for pathogen survival. The surge in use of phage and phage-based therapeutics as antimicrobials will require careful monitoring and detailed understanding of resistance mechanisms to help retain their efficacy, including further study of interactions between phage and bacteria. A lesson learned from use of antibiotics is that we should extend phage research to understand how introducing new therapeutic phage can influence existing biomes and ecosystems (588). The knowledge gained through studying phage-bacteria interactions has the potential to produce safe, effective, and resistance free future therapies, as well as halt the development of bacterial pathogens by understanding how phage predation shapes their evolution.

References

1. Lamas A, Miranda JM, Regal P, Vázquez B, Franco CM, Cepeda A. A comprehensive review of non-*enterica* subspecies of *Salmonella enterica*. Microbiological research. 2018;206:60-73.
2. Adeolu M, Alnajar S, Naushad S, Gupta RS. Genome-based phylogeny and taxonomy of the 'Enterobacteriales': proposal for Enterobacterales ord. nov. divided into the families Enterobacteriaceae, Erwiniaceae fam. nov., Pectobacteriaceae fam. nov., Yersiniaceae fam. nov., Hafniaceae fam. nov., Morganellaceae fam. nov., and Budviciaceae fam. nov. International journal of systematic and evolutionary microbiology. 2016;66(12):5575-99.
3. Sanderson KE, Nair S, Barrow P, Methner U. Taxonomy and species concepts in the genus *Salmonella*. *Salmonella* in domestic animals. 2013;2:1-19.
4. Alikhan N-F, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of *Salmonella*. PLoS genetics. 2018;14(4):e1007261.
5. Hoszowski A, Wasyl D. Taxonomy and nomenclature of genus *Salmonella*. Med Weter. 2000;56(2):75-8.
6. Criscuolo A, Issenhuth-Jeanjean S, Didelot X, Thorell K, Hale J, Parkhill J, et al. The speciation and hybridization history of the genus *Salmonella*. Microbial genomics. 2019;5(8).
7. Popoff M, Le Minor L. Antigenic formulas of the *Salmonella* serovars 7th revision: World Health Organization Collaborating Centre for Reference and Research on *Salmonella*. Pasteur Institute, Paris, France. 1997.
8. Dobzhansky T. Nothing in biology makes sense except in the light of evolution. The american biology teacher. 1973;35(3):125-9.
9. Ochman H. Evolutionary history of enteric bacteria. *Escherichia coli* and *Salmonella* typhimurium: cellular and molecular biology. 1987:1649-54.
10. Ochman H, Groisman E. The origin and evolution of species differences in *Escherichia coli* and *Salmonella* typhimurium. Molecular Ecology and Evolution: Approaches and Applications: Springer; 1994. p. 479-93.
11. Roth JR, Benson N, Galitski T, Haack K, Lawrence JG, Miesel L. Rearrangements of the bacterial chromosome: formation and applications.
12. Doolittle RF, Feng D-F, Tsang S, Cho G, Little E. Determining divergence times of the major kingdoms of living organisms with a protein clock. Science. 1996;271(5248):470-7.
13. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, et al. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature. 2001;413(6858):852-6.
14. Mitchell MA, Shane SM, editors. *Salmonella* in reptiles. Seminars in avian and exotic pet medicine; 2001: Elsevier.
15. Bäumler AJ. The record of horizontal gene transfer in *Salmonella*. Trends in microbiology. 1997;5(8):318-22.
16. Bäumler AJ, Tsolis RM, Heffron F. The lpf fimbrial operon mediates adhesion of *Salmonella* typhimurium to murine Peyer's patches. Proceedings of the National Academy of Sciences. 1996;93(1):279-83.

17. Lan R, Reeves PR. Gene transfer is a major factor in bacterial evolution. Molecular biology and evolution. 1996;13(1):47-55.
18. Hensel M. Evolution of pathogenicity islands of *Salmonella enterica*. International Journal of Medical Microbiology. 2004;294(2-3):95-102.
19. Müller K, Collinson SK, Kay W. Type 1 fimbriae of *Salmonella* enteritidis. Journal of bacteriology. 1991;173(15):4765-72.
20. Hinsley AP, Berks BC. Specificity of respiratory pathways involved in the reduction of sulfur compounds by *Salmonella enterica*. Microbiology. 2002;148(11):3631-8.
21. Hensel M, Hinsley AP, Nikolaus T, Sawers G, Berks BC. The genetic basis of tetrathionate respiration in *Salmonella* typhimurium. Molecular microbiology. 1999;32(2):275-87.
22. Lerminiaux NA, MacKenzie KD, Cameron AD. *Salmonella* Pathogenicity Island 1 (SPI-1): The Evolution and Stabilization of a Core Genomic Type Three Secretion System. Microorganisms. 2020;8(4):576.
23. Rehman T, Yin L, Latif MB, Chen J, Wang K, Geng Y, et al. Adhesive mechanism of different *Salmonella* fimbrial adhesins. Microbial pathogenesis. 2019;137:103748.
24. Barlag B, Hensel M. The giant adhesin SiiE of *Salmonella enterica*. Molecules. 2015;20(1):1134-50.
25. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, et al. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. Nature. 2010;467(7314):426-9.
26. Zhou J, Xiong X, Wang K-X, Zou L-J, Ji P, Yin Y-L. Ethanolamine enhances intestinal functions by altering gut microbiome and mucosal anti-stress capacity in weaned rats. British Journal of Nutrition. 2018;120(3):241-9.
27. Kofoid E, Rappleye C, Stojiljkovic I, Roth J. The 17-Gene Ethanolamine (eut) Operon of*Salmonella* typhimurium Encodes Five Homologues of Carboxysome Shell Proteins. Journal of bacteriology. 1999;181(17):5317-29.
28. Porwollik S, Wong RM-Y, McClelland M. Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis. Proceedings of the National Academy of Sciences. 2002;99(13):8956-61.
29. Boyd EF, Wang F-S, Whittam TS, Selander RK. Molecular genetic relationships of the *Salmonella*e. Applied and environmental microbiology. 1996;62(3):804-8.
30. Kingsley RA, Van Amsterdam K, Kramer N, Bäumler AJ. The shdA Gene Is Restricted to Serotypes of*Salmonella enterica* Subspecies I and Contributes to Efficient and Prolonged Fecal Shedding. Infection and immunity. 2000;68(5):2720-7.
31. Kisiela DI, Chattopadhyay S, Libby SJ, Karlinsey JE, Fang FC, Tchesnokova V, et al. Evolution of *Salmonella enterica* virulence via point mutations in the fimbrial adhesin. PLoS Pathog. 2012;8(6):e1002733.
32. Guiney DG, Fierer J. The role of the spv genes in *Salmonella* pathogenesis. Frontiers in microbiology. 2011;2:129.
33. Yue M, Rankin SC, Blanchet RT, Nulton JD, Edwards RA, Schifferli DM. Diversification of the *Salmonella* fimbriae: a model of macro-and microevolution. PloS one. 2012;7(6):e38596.

34. McQuiston JR, Fields PI, Tauxe RV, Logsdon Jr JM. Do *Salmonella* carry spare tyres? Trends in microbiology. 2008;16(4):142-8.
35. Mortimer CK, Gharbia SE, Logan JM, Peters TM, Arnold C. Flagellin gene sequence evolution in *Salmonella*. Infection, Genetics and Evolution. 2007;7(4):411-5.
36. Yamamoto S, Kutsukake K. FljA-mediated posttranscriptional control of phase 1 flagellin expression in flagellar phase variation of *Salmonella enterica* serovar Typhimurium. Journal of bacteriology. 2006;188(3):958-67.
37. Kutsukake K, Nakashima H, Tominaga A, Abo T. Two DNA invertases contribute to flagellar phase variation in *Salmonella enterica* serovar Typhimurium strain LT2. Journal of bacteriology. 2006;188(3):950-7.
38. Naberhaus SA, Krull AC, Arruda BL, Arruda P, Sahin O, Schwartz KJ, et al. Pathogenicity and competitive fitness of *Salmonella enterica* serovar 4,[5], 12: i:-compared to *Salmonella* typhimurium and *Salmonella* derby in swine. Frontiers in veterinary science. 2020;6:502.
39. Bawn M, Alikhan N-F, Thilliez G, Kirkwood M, Wheeler NE, Petrovska L, et al. Evolution of *Salmonella enterica* serotype Typhimurium driven by anthropogenic selection and niche adaptation. PLoS genetics. 2020;16(6):e1008850.
40. Shin H, Lee J-H, Kim H, Choi Y, Heu S, Ryu S. Receptor diversity and host interaction of bacteriophages infecting *Salmonella enterica* serovar Typhimurium. PloS one. 2012;7(8):e43392.
41. Senevirathne A, Hewawaduge C, Lee JH. *Salmonella enterica* serovar Enteritidis ghosts displaying a surface FliC adjuvant elicit a robust immune response and effective protection against virulent challenge. Veterinary Microbiology. 2020:108633.
42. Grimont PA, Weill F-X. Antigenic formulae of the *Salmonella* serovars. WHO collaborating centre for reference and research on *Salmonella*. 2007;9:1-166.
43. Tanner JR, Kingsley RA. Evolution of *Salmonella* within hosts. Trends in microbiology. 2018;26(12):986-98.
44. Langridge GC, Fookes M, Connor TR, Feltwell T, Feasey N, Parsons BN, et al. Patterns of genome evolution that have accompanied host adaptation in *Salmonella*. Proceedings of the National Academy of Sciences. 2015;112(3):863-8.
45. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, et al. Host adapted serotypes of *Salmonella enterica*. Epidemiology & Infection. 2000;125(2):229-55.
46. Rabsch W, Andrews HL, Kingsley RA, Prager R, Tschäpe H, Adams LG, et al. *Salmonella enterica* serotype Typhimurium and its host-adapted variants. Infection and immunity. 2002;70(5):2249-55.
47. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, et al. Distinct *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nature genetics. 2016;48(10):1211-7.
48. Kingsley RA, Bäumler AJ. Host adaptation and the emergence of infectious disease: the *Salmonella* paradigm. Molecular microbiology. 2000;36(5):1006-14.
49. Song J, Willinger T, Rongvaux A, Eynon EE, Stevens S, Manz MG, et al. A mouse model for the human pathogen *Salmonella* typhi. Cell host & microbe. 2010;8(4):369-76.

50. Key FM, Posth C, Esquivel-Gomez LR, Hübler R, Spyrou MA, Neumann GU, et al. Emergence of human-adapted *Salmonella enterica* is linked to the Neolithization process. Nature ecology & evolution. 2020;4(3):324-33.
51. Parkhill J, Dougan G, James K, Thomson N, Pickard D, Wain J, et al. Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. Nature. 2001;413(6858):848-52.
52. Holt KE, Thomson NR, Wain J, Langridge GC, Hasan R, Bhutta ZA, et al. Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi. BMC genomics. 2009;10(1):36.
53. Xie S, Li Y, Zhao S, Lv Y, Yu Q. *Salmonella* infection induced intestinal crypt hyperplasia through Wnt/β-catenin pathway in chicken. Research in Veterinary Science. 2020.
54. Feng Y, Johnston RN, Liu G-R, Liu S-L. Genomic comparison between *Salmonella* Gallinarum and Pullorum: differential pseudogene formation under common host restriction. PLoS One. 2013;8(3):e59427.
55. Fei X, Li Q, Olsen JE, Jiao X. A bioinformatic approach to identify core genome difference between *Salmonella* Pullorum and *Salmonella* Enteritidis. Infection, Genetics and Evolution. 2020:104446.
56. APHA. *Salmonella* serovars isolated from livestock 2014 - 2018. 2019.
57. PHE. 2018.
58. Chiu C-H, Su L-H, Chu C. *Salmonella enterica* serotype Choleraesuis: epidemiology, pathogenesis, clinical disease, and treatment. Clinical microbiology reviews. 2004;17(2):311-22.
59. Liu W-Q, Feng Y, Wang Y, Zou Q-H, Chen F, Guo J-T, et al. *Salmonella* paratyphi C: genetic divergence from *Salmonella* choleraesuis and pathogenic convergence with *Salmonella* typhi. PloS one. 2009;4(2):e4510.
60. Matthews TD, Schmieder R, Silva GG, Busch J, Cassman N, Dutilh BE, et al. Genomic comparison of the closely-related *Salmonella enterica* serovars Enteritidis, Dublin and Gallinarum. PloS one. 2015;10(6):e0126883.
61. Palomares Velosa JE, Salman MD, Roman-Muniz IN, Reynolds S, Linke L, Magnuson R, et al. Socio-ecological Factors of Zoonotic Diseases Exposure in Colorado Dairy Workers. Journal of Agromedicine. 2020:1-11.
62. Kudirkiene E, Sørensen G, Torpdahl M, de Knegt LV, Nielsen LR, Rattenborg E, et al. Epidemiology of *Salmonella enterica* Serovar Dublin in Cattle and Humans in Denmark, 1996 to 2016: a Retrospective Whole-Genome-Based Study. Applied and Environmental Microbiology. 2020;86(3).
63. Hong S-F, Chiu C-H, Chu C, Feng Y, Ou JT. Complete nucleotide sequence of a virulence plasmid of *Salmonella enterica* serovar Dublin and its phylogenetic relationship to the virulence plasmids of serovars Choleraesuis, Enteritidis and Typhimurium. FEMS microbiology letters. 2008;282(1):39-43.
64. Han J, Lynne AM, David DE, Nayak R, Foley SL. Sequencing of plasmids from a multi-antimicrobial resistant *Salmonella enterica* serovar Dublin strain. Food Research International. 2012;45(2):931-4.

65. Schroll C, Huang K, Ahmed S, Kristensen BM, Pors SE, Jelsbak L, et al. The SPI-19 encoded type-six secretion-systems (T6SS) of *Salmonella enterica* serovars Gallinarum and Dublin play different roles during infection. Veterinary microbiology. 2019;230:23-31.
66. Porwollik S, Santiviago C, Cheng P, Florea L, McClelland M. Differences in gene content between *Salmonella enterica* serovar Enteritidis isolates and comparison to closely related serovars Gallinarum and Dublin. Journal of bacteriology. 2005;187(18):6545-55.
67. Chiu C-H, Tang P, Chu C, Hu S, Bao Q, Yu J, et al. The genome sequence of *Salmonella enterica* serovar Choleraesuis, a highly invasive and resistant zoonotic pathogen. Nucleic acids research. 2005;33(5):1690-8.
68. Nair S, Fookes M, Corton C, Thomson NR, Wain J, Langridge GC. Genetic Markers in S. Paratyphi C Reveal Primary Adaptation to Pigs. Microorganisms. 2020;8(5):657.
69. Liu W-Q, Liu G-R, Li J-Q, Xu G-M, Qi D, He X-Y, et al. Diverse genome structures of *Salmonella* paratyphi C. BMC genomics. 2007;8(1):290.
70. Banerji S, Simon S, Tille A, Fruth A, Flieger A. Genome-based *Salmonella* serotyping as the new gold standard. Scientific reports. 2020;10(1):1-10.
71. Gantois I, Ducatelle R, Pasmans F, Haesebrouck F, Gast R, Humphrey TJ, et al. Mechanisms of egg contamination by *Salmonella* Enteritidis. FEMS microbiology reviews. 2009;33(4):718-38.
72. Debra, Stevens M. Analysis of the role of 13 major fimbrial subunits in colonisation of the chicken intestines by *Salmonella enterica* serovar Enteritidis reveals a role for a novel locus. BMC Microbiology. 2008;8:228.
73. Lei C-W, Zhang Y, Kang Z-Z, Kong L-H, Tang Y-Z, Zhang A-Y, et al. Vertical transmission of *Salmonella* Enteritidis with heterogeneous antimicrobial resistance from breeding chickens to commercial chickens in China. Veterinary Microbiology. 2020;240:108538.
74. Guard J, Cao G, Luo Y, Baugher JD, Davison S, Yao K, et al. Genome sequence analysis of 91 *Salmonella* Enteritidis isolates from mice caught on poultry farms in the mid 1990s. Genomics. 2020;112(1):528-44.
75. Gordon MA, Graham SM, Walsh AL, Wilson L, Phiri A, Molyneux E, et al. Epidemics of invasive *Salmonella enterica* serovar enteritidis and *S. enterica* Serovar typhimurium infection associated with multidrug resistance among adults and children in Malawi. Clinical Infectious Diseases. 2008;46(7):963-9.
76. Okoro CK, Kingsley RA, Connor TR, Harris SR, Parry CM, Al-Mashhadani MN, et al. Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. Nature genetics. 2012;44(11):1215.
77. Kingsley RA, Kay S, Connor T, Barquist L, Sait L, Holt KE, et al. Genome and transcriptome adaptation accompanying emergence of the definitive type 2 host-restricted *Salmonella enterica* serovar Typhimurium pathovar. mBio. 2013;4(5).
78. Kirkwood M, Vohra P, Bawn M, Thilliez G, Pye H, Tanner J, et al. Ecological niche adaptation of a bacterial pathogen associated with reduced zoonotic potential. bioRxiv. 2020.
79. Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, et al. Signatures of adaptation in human invasive *Salmonella* Typhimurium ST313 populations from sub-Saharan Africa. PLoS Negl Trop Dis. 2015;9(3):e0003611.

80. Van Puyvelde S, Pickard D, Vandelannoote K, Heinz E, Barbé B, de Block T, et al. An African *Salmonella* Typhimurium ST313 sublineage with extensive drug-resistance and signatures of host adaptation. Nature communications. 2019;10(1):1-12.
81. Seribelli AA, Gonzales JC, de Almeida F, Benevides L, Medeiros MIC, dos Prazeres Rodrigues D, et al. Phylogenetic analysis revealed that *Salmonella* Typhimurium ST313 isolated from humans and food in Brazil presented a high genomic similarity. Brazilian Journal of Microbiology. 2020;51(1):53-64.
82. Kirk MD, Pires SM, Black RE, Caipo M, Crump JA, Devleesschauwer B, et al. World Health Organization estimates of the global and regional disease burden of 22 foodborne bacterial, protozoal, and viral diseases, 2010: a data synthesis. PLoS medicine. 2015;12(12):e1001921.
83. Organization WH. *Salmonella* (non-typhoidal). Fact sheets Retrived from http://www who int/mediacentre/factsheets/fs139/en/(1202 2018). 2018.
84. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The global burden of nontyphoidal *Salmonella* gastroenteritis. Clinical infectious diseases. 2010;50(6):882-9.
85. Adhikari BB, Angulo F, Meltzer M. Economic burden of *Salmonella* infections in the United States. 2004.
86. Suijkerbuijk AW, Bouwknegt M, Mangen M-JJ, de Wit GA, van Pelt W, Bijkerk P, et al. The economic burden of a *Salmonella* Thompson outbreak caused by smoked salmon in the Netherlands, 2012–2013. The European Journal of Public Health. 2017;27(2):325-30.
87. Esan OB, Perera R, McCarthy N, Violato M, Fanshawe TR. Incidence, risk factors, and health service burden of sequelae of *Campylobacter* and Non-typhoidal *Salmonella* infections in England, 2000-2015: a retrospective cohort study using linked electronic health records: Sequelae of gastrointestinal infections. Journal of Infection. 2020.
88. Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. The Lancet. 2012;379(9835):2489-99.
89. Balasubramanian R, Im J, Lee J-S, Jeon HJ, Mogeni OD, Kim JH, et al. The global burden and epidemiology of invasive non-typhoidal *Salmonella* infections. Human vaccines & immunotherapeutics. 2019;15(6):1421-6.
90. Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, et al. The global burden of non-typhoidal *Salmonella* invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet Infectious Diseases. 2019;19(12):1312-24.
91. Parsons BN, Humphrey S, Salisbury AM, Mikoleit J, Hinton JC, Gordon MA, et al. Invasive non-typhoidal *Salmonella* typhimurium ST313 are not host-restricted and have an invasive phenotype in experimentally infected chickens. PLoS Negl Trop Dis. 2013;7(10):e2487.
92. Mead PS, Slutsker L, Dietz V, McCaig LF, Bresee JS, Shapiro C, et al. Food-related illness and death in the United States. Emerging infectious diseases. 1999;5(5):607.
93. Voetsch AC, Van Gilder TJ, Angulo FJ, Farley MM, Shallow S, Marcus R, et al. FoodNet estimate of the burden of illness caused by nontyphoidal *Salmonella* infections in the United States. Clinical Infectious Diseases. 2004;38(Supplement_3):S127-S34.

94. Krause G, Altmann D, Faensen D, Porten K, Benzler J, Pfoch T, et al. SurvNet electronic surveillance system for infectious disease outbreaks, Germany. 2007.
95. PHE. Public Health England Gastrointestinal Infections Data https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/337647/*Salmonella*_surveillance_tables.pdf: Public Health England; 2013.
96. Kogut MH, Swaggerty CL, Byrd JA, Selvaraj R, Arsenault RJ. Chicken-specific kinome array reveals that *Salmonella enterica* serovar Enteritidis modulates host immune signaling pathways in the cecum to establish a persistence infection. International journal of molecular sciences. 2016;17(8):1207.
97. Cogan T, Humphrey T. The rise and fall of *Salmonella* Enteritidis in the UK. Journal of applied microbiology. 2003;94:114-9.
98. ECDC. Incidence of *Salmonella* in Europe [2014 data]. 2016.
99. ECDC. Salmonellosis - Annual Epidemiological Report for 2017. 2020.
100. Campos J, Cristino L, Peixe L, Antunes P. MCR-1 in multidrug-resistant and copper-tolerant clinically relevant *Salmonella* 1, 4,[5], 12: i:-and S. Rissen clones in Portugal, 2011 to 2015. Eurosurveillance. 2016;21(26).
101. Mandilara G, Lambiri M, Polemis M, Passiotou M, Vatopoulos A. Phenotypic and molecular characterisation of multiresistant monophasic *Salmonella* Typhimurium (1, 4,[5], 12: i:-) in Greece, 2006 to 2011. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2013;18:20496.
102. Nigro G, Bottone G, Maiorani D, Trombatore F, Falasca S, Bruno G. Pediatric Epidemic of *Salmonella enterica* Serovar Typhimurium in the Area of L'Aquila, Italy, Four Years after a Catastrophic Earthquake. Int J Env Res Pub He. 2016;13(5):475.
103. Petrovska L, Mather AE, AbuOun M, Branchu P, Harris SR, Connor T, et al. Microevolution of Monophasic *Salmonella* Typhimurium during Epidemic, United Kingdom, 2005-2010. Emerging infectious diseases. 2016;22(4):617-24.
104. Raguenaud M, Le Hello S, Salah S, Weill F-X, Brisabois A, Delmas G, et al. Epidemiological and microbiological investigation of a large outbreak of monophasic *Salmonella* Typhimurium 4, 5, 12: i:-in schools associated with imported beef in Poitiers, France, October 2010. Eurosurveillance. 2010;17:40.
105. Wasyl D, Hoszowski A. Occurrence and characterization of monophasic *Salmonella enterica* serovar Typhimurium (1, 4,[5], 12: i:-) of non-human origin in Poland. Foodborne pathogens and disease. 2012;9(11):1037-43.
106. Sun H, Wan Y, Du P, Bai L. The epidemiology of monophasic *Salmonella* Typhimurium. Foodborne Pathogens and Disease. 2020;17(2):87-97.
107. Sun R-Y, Ke B-X, Fang L-X, Guo W-Y, Li X-P, Yu Y, et al. Global clonal spread of mcr-3-carrying MDR ST34 *Salmonella enterica* serotype Typhimurium and monophasic 1, 4,[5], 12: i:− variants from clinical isolates. Journal of Antimicrobial Chemotherapy. 2020.
108. dos Santos AM, Ferrari RG, Conte-Junior CA. Virulence factors in *Salmonella* Typhimurium: the sagacity of a bacterium. Current microbiology. 2019;76(6):762-73.

109. Buchwald DS, Blaser MJ. A review of human salmonellosis: II. Duration of excretion following infection with nontyphi *Salmonella*. Reviews of infectious diseases. 1984;6(3):345-56.
110. Blanc-Potard A-B, Solomon F, Kayser J, Groisman EA. The SPI-3 pathogenicity island of*Salmonella enterica*. J Bacteriol. 1999;181(3):998-1004.
111. Paesold G, Guiney DG, Eckmann L, Kagnoff MF. Genes in the *Salmonella* pathogenicity island 2 and the *Salmonella* virulence plasmid are essential for *Salmonella*-induced apoptosis in intestinal epithelial cells. Cellular microbiology. 2002;4(11):771-81.
112. Pickard D, Wain J, Baker S, Line A, Chohan S, Fookes M, et al. Composition, acquisition, and distribution of the Vi exopolysaccharide-encoding *Salmonella enterica* pathogenicity island SPI-7. Journal of Bacteriology. 2003;185(17):5055-65.
113. Waterman SR, Holden DW. Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. Cellular microbiology. 2003;5(8):501-11.
114. Schmidt H, Hensel M. Pathogenicity islands in bacterial pathogenesis. Clin Microbiol Rev. 2004;17(1):14-+.
115. Seth-Smith HMB. SPI-7: *Salmonella*'s Vi-Encoding Pathogenicity Island. Journal of infection in developing countries. 2008;2(4):267-71.
116. Elder JR, Chiok KL, Paul NC, Haldorson G, Guard J, Shah DH. The *Salmonella* pathogenicity island 13 contributes to pathogenesis in streptomycin pre-treated mice but not in day-old chickens. Gut pathogens. 2016;8(1):1.
117. Hurley D, McCusker MP, Fanning S, Martins M. *Salmonella*–host interactions–modulation of the host innate immune system. Frontiers in immunology. 2014;5:481.
118. Mintz E, Cartter M, Hadler J, Wassell J, Zingeser J, Tauxe R. Dose–response effects in an outbreak of *Salmonella* enteritidis. Epidemiology and infection. 1994;112(01):13-23.
119. Ashkenazi S, Amir Y, Dinari G, Schonfeld T, Nitzan M. Differential Leukocyte Count in Acute Gastroenteritis An Aid to Early Diagnosis. Clinical Pediatrics. 1983;22(5):356-8.
120. Ehrbar K, Friebel A, Miller SI, Hardt W-D. Role of the *Salmonella* pathogenicity island 1 (SPI-1) protein InvB in type III secretion of SopE and SopE2, two *Salmonella* effector proteins encoded outside of SPI-1. J Bacteriol. 2003;185(23):6950-67.
121. Tükel Ç, Raffatellu M, Humphries AD, Wilson RP, Andrews‐Polymenis HL, Gull T, et al. CsgA is a pathogen‐associated molecular pattern of *Salmonella enterica* serotype Typhimurium that is recognized by Toll‐like receptor 2. Molecular microbiology. 2005;58(1):289-304.
122. Zeng H, Wu H, Sloane V, Jones R, Yu Y, Lin P, et al. Flagellin/TLR5 responses in epithelia reveal intertwined activation of inflammatory and apoptotic pathways. American Journal of Physiology-Gastrointestinal and Liver Physiology. 2006;290(1):G96-G108.
123. Hapfelmeier S, Stecher B, Barthel M, Kremer M, Müller AJ, Heikenwalder M, et al. The *Salmonella* pathogenicity island (SPI)-2 and SPI-1 type III secretion systems allow *Salmonella* serovar typhimurium to trigger colitis via MyD88-dependent and MyD88-independent mechanisms. The Journal of Immunology. 2005;174(3):1675-85.
124. Ehrbar K, Hapfelmeier S, Stecher B, Hardt WD. InvB is required for type III-dependent secretion of SopA in *Salmonella enterica* serovar Typhimurium. J Bacteriol. 2004;186(4):1215-9.

125. Hapfelmeier S, Ehrbar K, Stecher B, Barthel M, Kremer M, Hardt WD. Role of the *Salmonella* pathogenicity island 1 effector proteins SipA, SopB, SopE, and SopE2 in *Salmonella enterica* subspecies 1 serovar Typhimurium colitis in streptomycin-pretreated mice. Infect Immun. 2004;72(2):795-809.
126. Figueira R, Watson KG, Holden DW, Helaine S. Identification of *Salmonella* pathogenicity island-2 type III secretion system effectors involved in intramacrophage replication of *S. enterica* serovar typhimurium: implications for rational vaccine design. mBio. 2013;4(2).
127. Jennings E, Thurston TL, Holden DW. *Salmonella* SPI-2 type III secretion system effectors: molecular mechanisms and physiological consequences. Cell host & microbe. 2017;22(2):217-31.
128. Blanc-Potard AB, Solomon F, Kayser J, Groisman EA. The SPI-3 pathogenicity island of *Salmonella enterica* [In Process Citation]. J Bacteriol. 1999;181(3):998-1004.
129. Rychlik I, Karasova D, Sebkova A, Volf J, Sisak F, Havlickova H, et al. Virulence potential of five major pathogenicity islands (SPI-1 to SPI-5) of *Salmonella enterica* serovar Enteritidis for chickens. BMC microbiology. 2009;9(1):268.
130. Wille T, Wagner C, Mittelstädt W, Blank K, Sommer E, Malengo G, et al. SiiA and SiiB are novel type I secretion system subunits controlling SPI 4 - mediated adhesion of S almonella *enterica*. Cellular microbiology. 2014;16(2):161-78.
131. Kiss T, Morgan E, Nagy G. Contribution of SPI-4 genes to the virulence of *Salmonella enterica*. FEMS microbiology letters. 2007;275(1):153-9.
132. Reis BP, Zhang S, Tsolis RM, Bäumler AJ, Adams LG, Santos RL. The attenuated sopB mutant of *Salmonella enterica* serovar Typhimurium has the same tissue distribution and host chemokine response as the wild type in bovine Peyer's patches. Veterinary microbiology. 2003;97(3-4):269-77.
133. Pezoa D, Yang H-J, Blondel CJ, Santiviago CA, Andrews-Polymenis HL, Contreras I. The type VI secretion system encoded in SPI-6 plays a role in gastrointestinal colonization and systemic spread of *Salmonella enterica* serovar Typhimurium in the chicken. PloS one. 2013;8(5):e63917.
134. Mulder DT, Cooper CA, Coombes BK. Type VI secretion system-associated gene clusters contribute to pathogenesis of *Salmonella enterica* serovar Typhimurium. Infection and immunity. 2012;80(6):1996-2007.
135. Sana TG, Flaugnatti N, Lugo KA, Lam LH, Jacobson A, Baylot V, et al. *Salmonella* Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. Proceedings of the National Academy of Sciences. 2016;113(34):E5044-E51.
136. Brugiroux S, Beutler M, Pfann C, Garzetti D, Ruscheweyh H-J, Ring D, et al. Genome-guided design of a defined mouse microbiota that confers colonization resistance against *Salmonella enterica* serovar Typhimurium. Nature microbiology. 2016;2(2):1-12.
137. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, et al. *Salmonella enterica* serovar typhimurium exploits inflammation to compete with the intestinal microbiota. PLoS Biol. 2007;5(10):2177-89.
138. Sana TG, Flaugnatti N, Lugo KA, Lam LH, Jacobson A, Baylot V, et al. *Salmonella* Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. Proceedings of the National Academy of Sciences. 2016:201608858.

139. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, et al. *Salmonella enterica* serovar typhimurium exploits inflammation to compete with the intestinal microbiota. PLoS Biol. 2007;5(10):e244.
140. Thiennimitr P, Winter SE, Winter MG, Xavier MN, Tolstikov V, Huseby DL, et al. Intestinal inflammation allows *Salmonella* to use ethanolamine to compete with the microbiota. Proceedings of the National Academy of Sciences. 2011;108(42):17480-5.
141. Ibarra JA, Steele‐Mortimer O. *Salmonella*–the ultimate insider. *Salmonella* virulence factors that modulate intracellular survival. Cellular microbiology. 2009;11(11):1579-86.
142. Haraga A, Ohlson MB, Miller SI. *Salmonella*e interplay with host cells. Nature Reviews Microbiology. 2008;6(1):53-66.
143. Horstmann JA, Lunelli M, Cazzola H, Heidemann J, Kühne C, Steffen P, et al. Methylation of *Salmonella* Typhimurium flagella promotes bacterial adhesion and host cell invasion. Nature Communications. 2020;11(1):1-11.
144. Ibarra JA, Knodler LA, Sturdevant DE, Virtaneva K, Carmody AB, Fischer ER, et al. Induction of *Salmonella* pathogenicity island 1 under different growth conditions can affect *Salmonella*–host cell interactions in vitro. Microbiology. 2010;156(4):1120-33.
145. Golubeva YA, Ellermeier JR, Chubiz JEC, Slauch JM. Intestinal Long-Chain Fatty Acids Act as a Direct Signal To Modulate Expression of the *Salmonella* Pathogenicity Island 1 Type III Secretion System. Mbio. 2016;7(1):e02170-15.
146. Curkić I, Müller B, Studer N, Sturm A, Hardt W-D. 4 Chapter 4-The quest for cause and consequence on the expression of the *Salmonella* Typhimurium type three secretion system 1 and the RpoE-mediated stress response. Single cell analysis tools for gene expression by *Salmonella* Typhimurium. 2016:92.
147. Bakowski MA, Braun V, Lam GY, Yeung T, Do Heo W, Meyer T, et al. The phosphoinositide phosphatase SopB manipulates membrane surface charge and trafficking of the *Salmonella*-containing vacuole. Cell host & microbe. 2010;7(6):453-62.
148. Knodler LA, Steele‐Mortimer O. Taking Possession: Biogenesis of the *Salmonella*‐Containing Vacuole. Traffic. 2003;4(9):587-99.
149. Fu Y, Galán JE. The *Salmonella* typhimurium tyrosine phosphatase SptP is translocated into host cells and disrupts the actin cytoskeleton. Molecular microbiology. 1998;27(2):359-68.
150. Mallo GV, Espina M, Smith AC, Terebiznik MR, Alemán A, Finlay BB, et al. SopB promotes phosphatidylinositol 3-phosphate formation on *Salmonella* vacuoles by recruiting Rab5 and Vps34. The Journal of cell biology. 2008;182(4):741-52.
151. Hensel M. *Salmonella* pathogenicity island 2. Molecular microbiology. 2000;36(5):1015-23.
152. Brumell JH, Goosney DL, Finlay BB. SifA, a Type III Secreted Effector of *Salmonella* typhimurium, Directs *Salmonella*‐Induced Filament (Sif) Formation Along Microtubules. Traffic. 2002;3(6):407-15.
153. Fitzsimmons LF, Liu L, Kant S, Kim J-S, Till JK, Jones-Carson J, et al. SpoT Induces Intracellular *Salmonella* Virulence Programs in the Phagosome. mBio. 2020;11(1).

154. Helaine S, Cheverton AM, Watson KG, Faure LM, Matthews SA, Holden DW. Internalization of *Salmonella* by macrophages induces formation of nonreplicating persisters. Science. 2014;343(6167):204-8.
155. Barrett D, Tilling O, Button E, Hart K, MacGillivray F, Jansen J, et al. Youngstock health: Effective disease prevention today ensuring tomorrow's profitable herd. Livestock. 2020;25(Sup2):1-24.
156. Forsythe SJ. The microbiology of safe food: John Wiley & Sons; 2020.
157. Cohen N, Messou M-AA, Calabrese S, Narayanan E. *Salmonella* vaccines. Google Patents; 2019.
158. Cully M. The politics of antibiotics. Nature. 2014;509(7498):S16.
159. Anadón A. WS14 The EU ban of antibiotics as feed additives (2006): alternatives and consumer safety. Journal of Veterinary Pharmacology and Therapeutics. 2006;29:41-4.
160. Debski B. Supplementation of pigs diet with zinc and copper as alternative to conventional antimicrobials. Pol J Vet Sci. 2016;19(4):917-24.
161. Luvsansharav UO, Vieira A, Bennett S, Huang J, Healy JM, Hoekstra RM, et al. *Salmonella* Serotypes: A Novel Measure of Association with Foodborne Transmission. Foodborne Pathogens and Disease. 2020;17(2):151-5.
162. Rabsch W. *Salmonella* Typhimurium Phage Typing for Pathogens. In: Schatten H, Eisenstark A, editors. *Salmonella*, Methods and Protocols. Methods in Molecular Biology. 394 ed. Totowa, New Jersey: Humana Press; 2007. p. 177-212.
163. Achtman M, Wain J, Weill FX, Nair S, Zhou Z, Sangal V. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. PLoS Pathog. 2012;8.
164. Reuter S, Ellington MJ, Cartwright EJ, Köser CU, Török ME, Gouliouris T, et al. Rapid bacterial whole-genome sequencing to enhance diagnostic and public health microbiology. JAMA internal medicine. 2013;173(15):1397-404.
165. Anderson E, Ward LR, de Saxe MJ, De Sa J. Bacteriophage-typing designations of *Salmonella* typhimurium. Epidemiology & Infection. 1977;78(2):297-300.
166. Felix A. Phage typing of *Salmonella* typhimurium: its place in epidemiological and epizootiological investigations. Microbiology. 1956;14(1):208-22.
167. Callow BR. A new phage-typing scheme for *Salmonella typhimurium*. J Hyg. 1959;57:346-59.
168. Hauser E, Tietze E, Helmuth R, Junker E, Blank K, Prager R, et al. Pork contaminated with *Salmonella enterica* serovar 4,[5],12:i:-, an emerging health risk for humans. Appl Environ Microbiol. 2010;76(14):4601-10.
169. Branchu P, Bawn M, Kingsley RA. Genome variation and molecular epidemiology of *Salmonella enterica* serovar Typhimurium pathovariants. Infection and immunity. 2018;86(8):e00079-18.
170. Thornton L, Gray S, Bingham P, Salmon R, Hutchinson D, Rowe B, et al. The problems of tracing a geographically widespread outbreak of salmonellosis from a commonly eaten food: *Salmonella* typhimurium DT193 in north west England and north Wales in 1991. Epidemiology & Infection. 1993;111(3):465-72.

171. Pang S, Octavia S, Reeves PR, Wang Q, Gilbert GL, Sintchenko V, et al. Genetic relationships of phage types and single nucleotide polymorphism typing of *Salmonella enterica* serovar Typhimurium. Journal of clinical microbiology. 2012;50(3):727-34.
172. Schmieger H. Molecular survey of the *Salmonella* phage typing system of Anderson. Journal of bacteriology. 1999;181(5):1630-5.
173. Switt AIM, Sulakvelidze A, Wiedmann M, Kropinski AM, Wishart DS, Poppe C, et al. *Salmonella* phages and prophages: genomics, taxonomy, and applied aspects. *Salmonella*: Springer; 2015. p. 237-87.
174. Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, Force CPT. PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. Emerging infectious diseases. 2001;7(3):382.
175. Sandvang D, Jensen LB, Baggesen D, Baloda SB. Persistence of a *Salmonella enterica* serotype Typhimurium clone in

Danish pig production units and farmhouse environment studied by

pulsed field gel electrophoresis (PFGE). FEMS Microbiol Lett. 2000;187:21-5.
176. Fabre L, Zhang J, Guigon G, Le Hello S, Guibert V, Accou-Demartin M, et al. CRISPR typing and subtyping for improved laboratory surveillance of *Salmonella* infections. PloS one. 2012;7(5):e36995.
177. Torpdahl M, S√∏rensen G, Lindstedt, Nielsen EM. Tandem repeat analysis for surveillance of human *Salmonella* Typhimurium infections. Emerging Infectious Diseases. 2007;13(3):388-95.
178. Xie X, Wang Z, Zhang K, Li Y, Hu Y, Pan Z, et al. Pig as a reservoir of CRISPR type TST4 *Salmonella enterica* serovar Typhimurium monophasic variant during 2009–2017 in China. Emerging Microbes & Infections. 2020;9(1):1-4.
179. Jolley K, Chan M, Maiden M. mlstdbNet - distributed multi-locus sequence typing (MLST) databases. BMC Bioinformatics. 2004;5(1):86.
180. Branchu P, Bawn M, Kingsley RA. Genome variation and molecular epidemiology of *Salmonella enterica* serovar Typhimurium pathovariants. Infection and Immunity. 2018;86(8).
181. Work TM, Dagenais J, Stacy BA, Ladner JT, Lorch JM, Balazs GH, et al. A novel host-adapted strain of *Salmonella* Typhimurium causes renal disease in olive ridley turtles (Lepidochelys olivacea) in the Pacific. Scientific reports. 2019;9(1):1-13.
182. Handeland K, Refsum T, Johansen B, Holstad G, Knutsen G, Solberg I, et al. Prevalence of *Salmonella* Typhimurium infection in Norwegian hedgehog populations associated with two human disease outbreaks. Epidemiology & Infection. 2002;128(3):523-7.
183. Rabsch W, Tschäpe H, Bäumler AJ. Non-typhoidal salmonellosis: emerging problems. Microbes and infection. 2001;3(3):237-47.
184. Petrovska L, Mather AE, AbuOun M, Branchu P, Harris SR, Connor T, et al. Microevolution of monophasic *Salmonella* typhimurium during epidemic, United Kingdom, 2005–2010. Emerging infectious diseases. 2016;22(4):617.
185. Garvey P, McKeown P, Kelly P, Cormican M, Anderson W, Flack A, et al. Investigation and management of an outbreak of *Salmonella* Typhimurium DT8 associated with duck eggs, Ireland 2009 to 2011. Eurosurveillance. 2013;18(16):20454.

186. Hughes LA, Shopland S, Wigley P, Bradon H, Leatherbarrow AH, Williams NJ, et al. Characterisation of *Salmonella enterica*serotype Typhimurium isolates from wild birds in northern England from 2005–2006. BMC Veterinary Research. 2008;4(1):4.
187. Mather AE, Lawson B, De Pinna E, Wigley P, Parkhill J, Thomson NR, et al. Genomic analysis of *Salmonella enterica* serovar Typhimurium from wild passerines in England and Wales. Applied and environmental microbiology. 2016;82(22):6728-35.
188. Andrews-Polymenis HL, Rabsch W, Porwollik S, McClelland M, Rosetti C, Adams LG, et al. Host restriction of *Salmonella enterica* serotype Typhimurium pigeon isolates does not correlate with loss of discrete genes. Journal of bacteriology. 2004;186(9):2619-28.
189. Bloomfield SJ, Benschop J, Biggs PJ, Marshall JC, Hayman DT, Carter PE, et al. Genomic analysis of *Salmonella enterica* serovar Typhimurium DT160 associated with a 14-year outbreak, New Zealand, 1998–2012. Emerging infectious diseases. 2017;23(6):906.
190. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. Genome research. 2009;19(12):2279-87.
191. Zhang S, Li S, Gu W, den Bakker H, Boxrud D, Taylor A, et al. Zoonotic source attribution of *Salmonella enterica* serotype Typhimurium using genomic surveillance data, United States. Emerging infectious diseases. 2019;25(1):82.
192. McClelland M, Sanderson KE, Clifton SW, Latreille P, Porwollik S, Sabo A, et al. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. Nature genetics. 2004;36(12):1268-74.
193. Mather AE, Phuong TLT, Gao Y, Clare S, Mukhopadhyay S, Goulding DA, et al. New variant of multidrug-resistant *Salmonella enterica* serovar Typhimurium associated with invasive disease in immunocompromised patients in Vietnam. mBio. 2018;9(5).
194. Rodríguez-Beltrán J, Sørum V, Toll-Riera M, de la Vega C, Peña-Miller R, San Millán Á. Genetic dominance governs the evolution and spread of mobile genetic elements in bacteria. Proceedings of the National Academy of Sciences. 2020;117(27):15755-62.
195. Mueller-Doblies D, Speed K, Davies RH. A retrospective analysis of *Salmonella* serovars isolated from pigs in Great Britain between 1994 and 2010. Preventive veterinary medicine. 2013;110(3-4):447-55.
196. Ethelberg S, Wingstrand A, Jensen T, Sørensen G, Muller L, Lisby M, et al. Large outbreaks of *Salmonella* Typhimurium infection in Denmark in 2008. Eurosurveillance. 2008;13(44):19023.
197. Aarestrup F. Get pigs off antibiotics. Nature. 2012;486(7404):465-6.
198. Barton MD. Impact of antibiotic use in the swine industry. Current opinion in microbiology. 2014;19:9-15.
199. Kim K-R, Owens G, Kwon S-I, So K-H, Lee D-B, Ok YS. Occurrence and environmental fate of veterinary antibiotics in the terrestrial environment. Water, Air, & Soil Pollution. 2011;214(1-4):163-74.
200. Manzetti S, Ghisi R. The environmental release and fate of antibiotics. Marine pollution bulletin. 2014;79(1-2):7-15.

201. Threlfall E, Frost J, Ward L, Rowe B. Epidemic in cattle and humans of *Salmonella* typhimurium DT 104 with chromosomally integrated multiple drug resistance. Veterinary Record. 1994;134(22):577-.
202. Helms M, Ethelberg S, Mølbak K, Group DS. International *Salmonella* typhimurium DT104 infections, 1992–2001. Emerging infectious diseases. 2005;11(6):859.
203. Pritha ST, Rahman S, Punom SA, Rahman MM, Nazir KNH, Islam MS. Isolation, Molecular Detection and Antibiogram of Multi-drug Resistant *Salmonella* Typhimurium DT104 from Selected Dairy Farms in Mymensingh, Bangladesh. American Journal of Microbiological Research. 2020;8(4):136-40.
204. Leekitcharoenphon P, Hendriksen RS, Le Hello S, Weill F-X, Baggesen DL, Jun S-R, et al. Global genomic epidemiology of *Salmonella enterica* serovar Typhimurium DT104. Applied and environmental microbiology. 2016;82(8):2516-26.
205. Soliman AM, Ramadan H, Ghazy E, Yu L, Hisatsune J, Kayama S, et al. Emergence of *Salmonella* genomic island 1 variant SGI1-C in a multidrug-resistant clinical isolate of Klebsiella pneumoniae ST485 from Egypt. Antimicrobial Agents and Chemotherapy. 2020.
206. Cummins ML, Hamidian M. *Salmonella* Genomic Island 1 is Broadly Disseminated within Gammaproteobacteriaceae. Microorganisms. 2020;8(2):161.
207. de Curraize C, Siebor E, Varin V, Neuwirth C, Hall RM. Two New SGI1-LK Variants Found in Proteus mirabilis and Evolution of the SGI1-HKL Group of *Salmonella* Genomic Islands. Msphere. 2020;5(2).
208. Amavisit P, Boonyawiwat W, Bangtrakulnont A. Characterization of *Salmonella enterica* serovar Typhimurium and monophasic *Salmonella* serovar 1, 4,[5], 12: i:-isolates in Thailand. Journal of clinical microbiology. 2005;43(6):2736-40.
209. Mossong J, Marques P, Ragimbeau C, Huberty-Krau P, Losch S, Meyer G, et al. Outbreaks of monophasic *Salmonella enterica* serovar 4,[5],12:i:- in Luxembourg, 2006. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2007;12(6):E11-2.
210. Laorden L, Herrera-Leon S, Martinez I, Sanchez A, Kromidas L, Bikandi J, et al. Genetic evolution of the Spanish multidrug-resistant *Salmonella enterica* 4,5,12:i:- monophasic variant. J Clin Microbiol. 2010;48(12):4563-6.
211. Nicholson F, Chambers B, Williams J, Unwin R. Heavy metal contents of livestock feeds and animal manures in England and Wales. Bioresource Technol. 1999;70(1):23-31.
212. Khot LR, Sankaran S, Maja JM, Ehsani R, Schuster EW. Applications of nanomaterials in agricultural production and crop protection: a review. Crop protection. 2012;35:64-70.
213. Jondreville C, Revy P, Dourmad J. Dietary means to better control the environmental impact of copper and zinc by pigs from weaning to slaughter. Livestock Production Science. 2003;84(2):147-56.
214. Li Y, Zeng Z, Chen Z, QIU J-p. Contamination of roxarsone to the environment around pig farm. Chin J Vet Sci. 2006;26(6):665-7.
215. Frost DV. Considerations on the safety of arsanilic acid for use in poultry feeds. Poultry Science. 1953;32(2):217-27.

216. Wang F-m, Chen Z-l, Sun Y-x, Gao Y, Yu J. Investigation on the pollution of organoarsenical additives to animal feed in the surroundings and farmland near hog farms. Acta Ecologica Sinica. 2006;1.
217. King T, Sheridan R. Determination of 27 Elements in Animal Feed by Inductively Coupled Plasma-Mass Spectrometry. Journal of AOAC International. 2019;102(2):434-44.
218. Arai N, Sekizuka T, Tamamura Y, Tanaka K, Barco L, Izumiya H, et al. 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 4 decades. Journal of clinical microbiology. 2018;56(5).
219. Proroga YTR, Mancusi A, Peruzy MF, Carullo MR, Montone AMI, Fulgione A, et al. Characterization of *Salmonella* Typhimurium and its monophasic variant 1, 4,[5], 12: i:-isolated from different sources. Folia Microbiologica. 2019;64(6):711-8.
220. Maurischat S, Baumann B, Martin A, Malorny B. Rapid detection and specific differentiation of *Salmonella enterica* subsp. *enterica* Enteritidis, Typhimurium and its monophasic variant 4,[5], 12: i:− by real-time multiplex PCR. International journal of food microbiology. 2015;193:8-14.
221. Raguenaud M, Le Hello S, Salah S, Weill F-X, Brisabois A, Delmas G, et al. Epidemiological and microbiological investigation of a large outbreak of monophasic *Salmonella* Typhimurium 4, 5, 12: i:-in schools associated with imported beef in Poitiers, France, October 2010. Eurosurveillance. 2012;17(40):20289.
222. Grattarola C, Gallina S, Giorda F, Pautasso A, Ballardini M, Iulini B, et al. First report of *Salmonella* 1, 4,[5], 12: i:-in free-ranging striped dolphins (Stenella coeruleoalba), Italy. Scientific reports. 2019;9(1):1-14.
223. Li X, Jiang Y, Wu K, Zhou Y, Liu R, Cao Y, et al. Whole-genome sequencing identification of a multidrug-resistant *Salmonella enterica* serovar Typhimurium strain carrying blaNDM-5 from Guangdong, China. Infection, Genetics and Evolution. 2017;55:195-8.
224. Fernández J, Guerra B, Rodicio MR. Resistance to carbapenems in non-typhoidal *Salmonella enterica* Serovars from humans, animals and food. Veterinary sciences. 2018;5(2):40.
225. Biswas S, Li Y, Elbediwi M, Yue M. Emergence and Dissemination of mcr-Carrying Clinically Relevant *Salmonella* Typhimurium Monophasic Clone ST34. Microorganisms. 2019;7(9):298.
226. Elbediwi M, Pan H, Jiang Z, Biswas S, Li Y, Yue M. Genomic Characterization of mcr-1-carrying *Salmonella enterica* Serovar 4,[5], 12: i:-ST 34 Clone Isolated From Pigs in China. Frontiers in bioengineering and biotechnology. 2020;8.
227. Luo Q, Wan F, Yu X, Zheng B, Chen Y, Gong C, et al. MDR *Salmonella enterica* serovar Typhimurium ST34 carrying mcr-1 isolated from cases of bloodstream and intestinal infection in children in China. Journal of Antimicrobial Chemotherapy. 2020;75(1):92-5.
228. Threlfall E, Frost J, Ward L, Rowe B. Plasmid profile typing can be used to subdivide phage-type 49 of *Salmonella* typhimurium in outbreak investigations. Epidemiology & Infection. 1990;104(2):243-51.
229. Bayram LC, AYDIN F. An abdominal cavity abscess associated with *Salmonella enterica* serovar Typhimurium phage type DT2 in a dog: a case report. Veterinarni Medicina. 2016;61(5).

230. Alley M, Connolly J, Fenwick S, Mackereth G, Leyland M, Rogers L, et al. An epidemic of salmonellosis caused by *Salmonella* Typhimurium DT160 in wild birds and humans in New Zealand. New Zealand Veterinary Journal. 2002;50(5):170-6.
231. Fukui D, Takahashi K, Kubo M, Une Y, Kato Y, Izumiya H, et al. Mass mortality of Eurasian tree sparrows (Passer montanus) from *Salmonella* Typhimurium DT40 in Japan, winter 2008–09. Journal of Wildlife Diseases. 2014;50(3):484-95.
232. Hall AJ, Saito EK. Avian wildlife mortality events due to salmonellosis in the United States, 1985–2004. Journal of wildlife diseases. 2008;44(3):585-93.
233. Giovannini S, Pewsner M, Hüssy D, Hächler H, Degiorgis M-PR, Hirschheydt Jv, et al. Epidemic of salmonellosis in passerine birds in Switzerland with spillover to domestic cats. Veterinary pathology. 2013;50(4):597-606.
234. Hernandez SM, Keel K, Sanchez S, Trees E, Gerner-Smidt P, Adams JK, et al. Epidemiology of a *Salmonella enterica* subsp. *enterica* serovar Typhimurium strain associated with a songbird outbreak. Applied and environmental microbiology. 2012;78(20):7290-8.
235. Söderlund R, Jernberg C, Trönnberg L, Pääjärvi A, Ågren E, Lahti E. Linked seasonal outbreaks of *Salmonella* Typhimurium among passerine birds, domestic cats and humans, Sweden, 2009 to 2016. Eurosurveillance. 2019;24(34):1900074.
236. Rabsch W. *Salmonella* typhimurium phage typing for pathogens. *Salmonella*: Springer; 2007. p. 177-211.
237. Threlfall E, Ward L, Rowe B. Epidermic spread of a chloramphenicol-resistant strain of *Salmonella* typhimurium phage type 204 in bovine animals in Britain. Veterinary Record. 1978;103(20):438-40.
238. Threlfall E, Rowe B, Ferguson J, Ward L. Increasing incidence of resistance to gentamicin and related aminoglycosides in *Salmonella* typhimurium phage type 204c in England, Wales and Scotland. The Veterinary record. 1985;117(14):355-7.
239. Ashton PM, Owen SV, Kaindama L, Rowe WP, Lane CR, Larkin L, et al. Public health surveillance in the UK revolutionises our understanding of the invasive *Salmonella* Typhimurium epidemic in Africa. Genome medicine. 2017;9(1):1-13.
240. Ashton PM, Owen SV, Kaindama L, Rowe WP, Lane C, Larkin L, et al. *Salmonella enterica* serovar Typhimurium ST313 responsible for gastroenteritis in the UK are genetically distinct from isolates causing bloodstream infections in Africa. BioRxiv. 2017:139576.
241. Noble D, Lane C, Little C, Davies R, De Pinna E, Larkin L, et al. 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. Epidemiology & Infection. 2012;140(1):146-9.
242. Ashton PM, Peters T, Ameh L, McAleer R, Petrie S, Nair S, et al. Whole genome sequencing for the retrospective investigation of an outbreak of *Salmonella* Typhimurium DT 8. PLoS currents. 2015;7.
243. Garvey P, McKeown P, Kelly P, Cormican M, Anderson W, Flack A, et al. Investigation and management of an outbreak of *Salmonella* Typhimurium DT8 associated with duck eggs, Ireland 2009 to 2011. 2013.

244. Garrod LP, Mc IM. Hospital outbreak of enteritis due to duck eggs. Br Med J. 1949;2(4639):1259-61.
245. Owen M, Jorgensen F, Willis C, McLauchlin J, Elviss N, Aird H, et al. The occurrence of *Salmonella* spp. in duck eggs on sale at retail or from catering in England. Letters in applied microbiology. 2016;63(5):335-9.
246. Mohammed M, Cormican M. Whole genome sequencing provides possible explanations for the difference in phage susceptibility among two *Salmonella* Typhimurium phage types (DT8 and DT30) associated with a single foodborne outbreak. BMC research notes. 2015;8(1):728.
247. Rabsch W, Andrews HL, Kingsley RA, Prager R, Tschape H, Adams LG, et al. *Salmonella enterica* serotype Typhimurium and its host-adapted variants. Infect Immun. 2002;70(5):2249-55.
248. Wray C, McLaren I, Jones YE. The epidemiology of *Salmonella* typhimurium in cattle: plasmid profile analysis of definitive phage type (DT) 204c. Journal of medical microbiology. 1998;47(6):483-7.
249. Jacob W, Kühn H, Kürschner H, Rabsch W. The epidemiological analysis of *Salmonella* typhimurium infections in cattle--results of lysotyping and biochemotyping in the region of East Thuringia from 1974 to 1991. Berliner und Munchener tierarztliche Wochenschrift. 1993;106(8):265-9.
250. Mulvey MR, Boyd DA, Olson AB, Doublet B, Cloeckaert A. The genetics of *Salmonella* genomic island 1. Microbes and Infection. 2006;8(7):1915-22.
251. Threlfall EJ. Epidemic *Salmonella* typhimurium DT 104—a truly international multiresistant clone. Journal of Antimicrobial Chemotherapy. 2000;46(1):7-10.
252. Petrovska L, Mather AE, AbuOun M, Branchu P, Harris SR, Connor T, et al. Microevolution of monophasic *Salmonella* Typhimurium during epidemic, United Kingdom, 2005–2010. Emerging infectious diseases. 2016;22(4):617.
253. Branchu P, Charity OJ, Bawn M, Thilliez G, Dallman TJ, Petrovska L, et al. SGI-4 in monophasic *Salmonella* Typhimurium ST34 is a novel ICE that enhances resistance to copper. bioRxiv. 2019:518175.
254. Gencay YE, Gambino M, Prüssing TF, Brøndsted L. The genera of bacteriophages and their receptors are the major determinants of host range. Environmental microbiology. 2019;21(6):2095-111.
255. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: The agents of open source evolution. Nature Reviews Microbiology. 2005;3(9):722-32.
256. Fitzpatrick DA, Logue ME, Butler G. Evidence of recent interkingdom horizontal gene transfer between bacteria and Candida parapsilosis. BMC evolutionary biology. 2008;8(1):1.
257. Figge RM, Schubert M, Brinkmann H, Cerff R. Glyceraldehyde-3-phosphate dehydrogenase gene diversity in eubacteria and eukaryotes: evidence for intra-and inter-kingdom gene transfer. Mol Biol Evol. 1999;16(4):429-40.
258. Bellanger X, Payot S, Leblond-Bourget N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiology Reviews. 2014;38(4):720-60.
259. Vernikos GS, Parkhill J. Interpolated variable order motifs for identification of horizontally acquired DNA: revisiting the *Salmonella* pathogenicity islands. Bioinformatics. 2006;22(18):2196-203.

260. Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. Nature. 2000;405(6784):299-304.
261. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS microbiology reviews. 2009;33(2):376-93.
262. Navarre WW, Porwollik S, Wang Y, McClelland M, Rosen H, Libby SJ, et al. Selective silencing of foreign DNA with low GC content by the H-NS protein in *Salmonella*. Science. 2006;313(5784):236-8.
263. Doublet B, Boyd D, Mulvey MR, Cloeckaert A. The *Salmonella* genomic island 1 is an integrative mobilizable element. Molecular microbiology. 2005;55(6):1911-24.
264. Rotger R, Casadesús J. The virulence plasmids of *Salmonella*. International Microbiology. 1999;2:177-84.
265. Branchu P, Charity OJ, Bawn M, Thilliez G, Dallman TJ, Petrovska L, et al. SGI-4 in monophasic *Salmonella* Typhimurium ST34 is a novel ICE that enhances resistance to copper. Frontiers in microbiology. 2019;10:1118.
266. Acman M, van Dorp L, Santini JM, Balloux F. Large-scale network analysis captures biological features of bacterial plasmids. Nature communications. 2020;11(1):1-11.
267. Sheppard RJ, Beddis AE, Barraclough TG. The role of hosts, plasmids and environment in determining plasmid transfer rates: A meta-analysis. Plasmid. 2020;108:102489.
268. Ramirez MS, Traglia GM, Lin DL, Tran T, Tolmasky ME. Plasmid‐mediated antibiotic resistance and virulence in gram‐negatives: the Klebsiella pneumoniae paradigm. Plasmids: Biology and Impact in Biotechnology and Discovery. 2015:459-74.
269. Carattoli A. Plasmid-mediated antimicrobial resistance in *Salmonella enterica*. Current issues in molecular biology. 2003;5(4):113-22.
270. Cazares A, Hall JP, Wright LL, Grimes M, Emond-Rhéault J-G, Pongchaikul P, et al. Characterisation of a new megaplasmid family associated with the spread of multidrug resistance in Pseudomonas aeruginosa. Access Microbiology. 2020;2(7A):646.
271. Bethke JH, Davidovich A, Cheng L, Lopatkin AJ, Song W, Thaden JT, et al. Environmental and genetic determinants of plasmid mobility in pathogenic *Escherichia coli*. Science advances. 2020;6(4):eaax3173.
272. Velappan N, Sblattero D, Chasteen L, Pavlik P, Bradbury AR. Plasmid incompatibility: more compatible than previously thought? Protein Engineering, Design and Selection. 2007;20(7):309-13.
273. Boyd EF, Hartl DL. *Salmonella* virulence plasmid: modular acquisition of the spv virulence region by an F-plasmid in *Salmonella enterica* subspecies I and insertion into the chromosome of subspecies II, IIIa, IV and VII isolates. Genetics. 1998;149(3):1183-90.
274. Gao Y, Wen J, Wang S, Xu X, Zhan Z, Chen Z, et al. Plasmid-Encoded bla NDM-5 Gene That Confers High-Level Carbapenem Resistance in *Salmonella* Typhimurium of Pork Origin. Infection and Drug Resistance. 2020;13:1485-90.
275. Wyrsch ER, Hawkey J, Judd LM, Haites R, Holt KE, Djordjevic SP, et al. Z/I1 Hybrid Virulence Plasmids Carrying Antimicrobial Resistance genes in S. Typhimurium from Australian Food Animal Production. Microorganisms. 2019;7(9):299.

276. Tassinari E, Duffy G, Bawn M, Burgess CM, McCabe EM, Lawlor PG, et al. Microevolution of antimicrobial resistance and biofilm formation of *Salmonella* Typhimurium during persistence on pig farms. Scientific Reports. 2019;9(1):8832.
277. Xiang Y, Li F, Dong N, Tian S, Zhang H, Du X, et al. Investigation of a salmonellosis outbreak caused by multidrug resistant *Salmonella* Typhimurium in China. Frontiers in microbiology. 2020;11:801.
278. Carroll L, Gaballa A, Guldimann C, Sullivan G, Henderson L, Wiedmann M. Identification of novel mobilized colistin resistance gene mcr-9 in a multidrug-resistant, colistin-susceptible *Salmonella enterica* serotype Typhimurium isolate. mBio 10: e00853-19. 2019.
279. Dong N, Li Y, Zhao J, Ma H, Wang J, Liang B, et al. The phenotypic and molecular characteristics of antimicrobial resistance of *Salmonella enterica* subsp. *enterica* serovar Typhimurium in Henan Province, China. BMC Infectious Diseases. 2020;20(1):1-11.
280. Kidwell MG. Transposable elements. The evolution of the genome: Elsevier; 2005. p. 165-221.
281. Perween S, Kumar D, Kumar A. A Review on Transposons and its Utilization as Genetic Tool. Int J Curr Microbiol App Sci. 2020;9(2):1874-84.
282. Ladd M, Bordoni B. Genetics, Transposons. StatPearls [Internet]: StatPearls Publishing; 2020.
283. Kaufman J, Terrizzano I, Nayar G, Seabolt E, Agarwal A, Slizovskiy IB, et al. Integrative and Conjugative Elements (ICE) and Associated Cargo Genes within and across Hundreds of Bacterial Genera. bioRxiv. 2020.
284. Te Poele EM, Bolhuis H, Dijkhuizen L. Actinomycete integrative and conjugative elements. Antonie Van Leeuwenhoek. 2008;94(1):127-43.
285. Botelho J, Schulenburg H. The Role of Integrative and Conjugative Elements in Antibiotic Resistance Evolution. Trends in Microbiology. 2020.
286. Guglielmini J, Quintais L, Garcillán-Barcia MP, de La Cruz F, Rocha EP. The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. PLoS genet. 2011;7(8):e1002222.
287. Boyd EF, Almagro-Moreno S, Parent MA. Genomic islands are dynamic, ancient integrative elements in bacterial evolution. Trends in microbiology. 2009;17(2):47-53.
288. Brochet M, Couvé E, Glaser P, Guédon G, Payot S. Integrative conjugative elements and related elements are major contributors to the genome diversity of Streptococcus agalactiae. Journal of bacteriology. 2008;190(20):6913-7.
289. Esposito D, Scocca JJ. The integrase family of tyrosine recombinases: evolution of a conserved active site domain. Nucleic Acids Research. 1997;25(18):3605-14.
290. Ringwald K, Yoneji S, Gardner J. Resolution of mismatched overlap Holliday junction intermediates by the tyrosine recombinase IntDOT. Journal of bacteriology. 2017;199(10).
291. Johnson CM, Grossman AD. Integrative and conjugative elements (ICEs): what they do and how they work. Annual review of genetics. 2015;49:577-601.
292. Koraimann G, Wagner MA. Social behavior and decision making in bacterial conjugation. Frontiers in cellular and infection microbiology. 2014;4:54.

293. Das B, Martínez E, Midonet C, Barre F-X. Integrative mobile elements exploiting Xer recombination. Trends in microbiology. 2013;21(1):23-30.
294. Cury J, Touchon M, Rocha EP. Integrative and conjugative elements and their hosts: composition, distribution and organization. Nucleic acids research. 2017;45(15):8943-56.
295. Santoro F, Romeo A, Pozzi G, Iannelli F. Excision and Circularization of Integrative Conjugative Element Tn5253 of Streptococcus pneumoniae. Frontiers in microbiology. 2018;9:1779.
296. Weiss E, Spicher C, Haas R, Fischer W. Excision and transfer of an integrating and conjugative element in a bacterial species with high recombination efficiency. Scientific reports. 2019;9(1):1-14.
297. Menard KL, Grossman AD. Selective pressures to maintain attachment site specificity of integrative and conjugative elements. PLoS Genet. 2013;9(7):e1003623.
298. Wozniak RA, Waldor MK. A toxin–antitoxin system promotes the maintenance of an integrative conjugative element. PLoS Genet. 2009;5(3):e1000439.
299. Fraikin N, Goormaghtigh F, Van Melderen L. Type II toxin-antitoxin systems: evolution and revolutions. Journal of Bacteriology. 2020;202(7).
300. Switt AIM, den Bakker HC, Cummings CA, Rodriguez-Rivera LD, Govoni G, Raneiri ML, et al. Identification and characterization of novel *Salmonella* mobile elements involved in the dissemination of genes linked to virulence and transmission. PloS one. 2012;7(7):e41247.
301. Seth-Smith HM. SPI-7: *Salmonella*'s Vi-encoding pathogenicity island. The journal of infection in developing countries. 2008;2(04):267-71.
302. Brüssow H, Hendrix RW. Phage genomics: small is beautiful. Cell. 2002;108(1):13-6.
303. Suttle CA. Viruses in the sea. Nature. 2005;437(7057):356-61.
304. Hutkins RW. Microbiology and technology of fermented foods: John Wiley & Sons; 2008.
305. Stern A, Sorek R. The phage‐host arms race: shaping the evolution of microbes. BioEssays : news and reviews in molecular, cellular and developmental biology. 2011;33(1):43-51.
306. Wright A. Mechanism of conversion of the *Salmonella* O antigen by bacteriophage ε 34. Journal of Bacteriology. 1971;105(3):927-36.
307. Kintz E, Davies MR, Hammarlöf DL, Canals R, Hinton JC, van der Woude MW. A BTP 1 prophage gene present in invasive non‐typhoidal S almonella determines composition and length of the O‐antigen of the lipopolysaccharide. Molecular microbiology. 2015;96(2):263-75.
308. Mirold S, Rabsch W, Rohde M, Stender S, Tsch√§pe H, R√∞ssmann H, et al. Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella* typhimurium strain. Proceedings of the National Academy of Sciences of the United States of America. 1999;96(17):9845-50.
309. Coombes BK, Wickham ME, Brown NF, Lemire S, Bossi L, Hsiao WW, et al. Genetic and molecular analysis of GogB, a phage-encoded type III-secreted substrate in *Salmonella enterica* serovar typhimurium with autonomous expression from its associated phage. JOurnal of Molecular Biology. 2005;348(4):817-30.
310. Figueroa-Bossi N, Bossi L. Inducible prophages contribute to *Salmonella* virulence in mice. Molecular microbiology. 1999;33(1):167-76.
311. Figueroa-Bossi N, Uzzau S, Maloriol D, Bossi L. Variable assortment of prophages provides a transferable repertoire of pathogenic determinants in *Salmonella*. Mol Microbiol. 2001;39(2):260-72.

312. Ho TD, Slauch JM. Characterization of grvA, an antivirulence gene on the Gifsy-2 phage in *Salmonella enterica* serovar Typhimurium. Journal of bacteriology. 2001;183(2):611-20.
313. Ackermann H-W. Tailed bacteriophages: the order Caudovirales. Advances in virus research. 1998;51:135-201.
314. Cornelis GR. The type III secretion injectisome. Nature Reviews Microbiology. 2006;4(11):811-25.
315. Strobel E, Schmieger H. Invitro Packaging of Exogenous DNA by *Salmonella* Phage-P22. J Gen Virol. 1979;45(Nov):291-9.
316. Earnshaw WC, Casjens SR. DNA packaging by the double-stranded DNA bacteriophages. Cell. 1980;21(2):319-31.
317. Abrescia NG, Cockburn JJ, Grimes JM, Sutton GC, Diprose JM, Butcher SJ, et al. Insights into assembly from structural analysis of bacteriophage PRD1. Nature. 2004;432(7013):68-74.
318. Susskind MM, Botstein D. Molecular genetics of bacteriophage P22. Microbiological reviews. 1978;42(2):385.
319. Reichardt LF. Control of bacteriophage lambda repressor synthesis after phage infection: the role of the N, cII, cIII and cro products. J Mol Biol. 1975;93(2):267-88.
320. Botstein D. Synthesis and maturation of phage P22 DNA: I. Identification of intermediates. J Mol Biol. 1968;34(3):621-41.
321. Botstein D, Waddell CH, King J. Mechanism of head assembly and DNA encapsulation in *Salmonella* phage P22: I. Genes, proteins, structures and DNA maturation. J Mol Biol. 1973;80(4):669IN13679-678IN8695.
322. Earnshaw W, Casjens S, Harrison SC. Assembly of the head of bacteriophage P22: x-ray diffraction from heads, proheads and related structures. J Mol Biol. 1976;104(2):387-410.
323. Israel J, Anderson TF, Levine M. In vitro morphogenesis of phage P22 from heads and base-plate parts. Proceedings of the National Academy of Sciences. 1967;57(2):284-91.
324. Levine M. Mutations in the temperate phage P22 and lysogeny in *Salmonella*. Virology. 1957;3(1):22-41.
325. Schmieger H. Phage P22-mutants with increased or decreased transduction abilities. Molecular and General Genetics MGG. 1972;119(1):75-88.
326. Smith HO, Levine M. A phage P22 gene controlling integration of prophage. Virology. 1967;31(2):207-16.
327. De Lappe N, Doran G, O'Connor J, O'Hare C, Cormican M. Characterization of bacteriophages used in the *Salmonella enterica* serovar Enteritidis phage-typing scheme. Journal of Medical Microbiology. 2009;58(1):86-93.
328. Allison GE, Angeles D, Tran-Dinh N, Verma NK. Complete genomic sequence of SfV, a serotype-converting temperate bacteriophage of Shigella flexneri. Journal of bacteriology. 2002;184(7):1974-87.
329. Mmolawa PT, Schmieger H, Heuzenroeder MW. Bacteriophage ST64B, a genetic mosaic of genes from diverse sources isolated from *Salmonella enterica* serovar typhimurium DT 64. Journal of bacteriology. 2003;185(21):6481-5.

330. Smith DL, Rooks DJ, Fogg PC, Darby AC, Thomson NR, McCarthy AJ, et al. Comparative genomics of Shiga toxin encoding bacteriophages. BMC genomics. 2012;13(1):1-10.
331. Lindberg AA, Hellerqvist CG, Bagbian-Motta G, Makela PH. Lipopolysaccharide modification accompanying antigenic conversion by phage P27. Microbiology. 1978;107(2):279-87.
332. Karambelkar S, Udupa S, Gowthami VN, Ramachandra SG, Swapna G, Nagaraja V. Emergence of a novel immune-evasion strategy from an ancestral protein fold in bacteriophage Mu. Nucleic acids research. 2020;48(10):5294-305.
333. Howe M, Bade E. Molecular biology of bacteriophage Mu. Science. 1975;190(4215):624-32.
334. Morgan GJ, Hatfull GF, Casjens S, Hendrix RW. Bacteriophage Mu genome sequence: analysis and comparison with Mu-like prophages in Haemophilus, Neisseria and Deinococcus. JOurnal of Molecular Biology. 2002;317(3):337-59.
335. Mirold S, Rabsch W, Rohde M, Stender S, Tschäpe H, Rüssmann H, et al. Isolation of a temperate bacteriophage encoding the type III effector protein SopE from an epidemic *Salmonella* typhimurium strain. Proceedings of the National Academy of Sciences. 1999;96(17):9845-50.
336. Nilsson AS, Haggard-Ljungquist E. The P2-like bacteriophages. The bacteriophages. 2006;1:365-90.
337. Bullas LR, Mostaghimi AR, Arensdorf JJ, Rajadas PT, Zuccarelli AJ. *Salmonella* phage PSP3, another member of the P2-like phage group. Virology. 1991;185(2):918-21.
338. Gao R, Naushad S, Moineau S, Levesque R, Goodridge L, Ogunremi D. Comparative genomic analysis of 142 bacteriophages infecting *Salmonella enterica* subsp. *enterica*. BMC Genomics. 2020;21:1-13.
339. Reen F, Boyd E, Porwollik S, Murphy B, Gilroy D, Fanning S, et al. Genomic comparisons of *Salmonella enterica* serovar Dublin, Agona, and Typhimurium strains recently isolated from milk filters and bovine samples from Ireland, using a *Salmonella* microarray. Applied and environmental microbiology. 2005;71(3):1616-25.
340. Bossi L, Figueroa-Bossi N. Prophage arsenal of *Salmonella enterica* serovar Typhimurium. Phages: American Society of Microbiology; 2005. p. 165-86.
341. Owen SV, Wenner N, Canals R, Makumi A, Hammarlöf DL, Gordon MA, et al. Characterization of the prophage repertoire of African *Salmonella* Typhimurium ST313 reveals high levels of spontaneous induction of novel phage BTP1. Frontiers in microbiology. 2017;8:235.
342. Calendar R, Lindqvist B, Sironi G, Clark AJ. Characterization of REP− mutants and their interaction with P2 phage. Virology. 1970;40(1):72-83.
343. Nilsson AS, Haggård-Ljungquist E. Evolution of P2-like phages and their impact on bacterial evolution. Research in microbiology. 2007;158(4):311-7.
344. Trinh JT, Zeng L. Phage-Phage Interactions. Biocommunication of Phages: Springer; 2020. p. 87-102.
345. Mitarai N. How pirate phage interferes with helper phage: Comparison of the two distinct strategies. Journal of Theoretical Biology. 2020;486:110096.
346. Dion MB, Oechslin F, Moineau S. Phage diversity, genomics and phylogeny. Nature Reviews Microbiology. 2020:1-14.

347. Tanaka K, Nishimori K, Makino S-I, Nishimori T, Kanno T, Ishihara R, et al. Molecular characterization of a prophage of *Salmonella enterica* serotype Typhimurium DT104. Journal of clinical microbiology. 2004;42(4):1807-12.
348. Vander Byl C, Kropinski AM. Sequence of the Genome of *Salmonella*Bacteriophage P22. Journal of bacteriology. 2000;182(22):6472-81.
349. Pedulla ML, Ford ME, Karthikeyan T, Houtz JM, Hendrix RW, Hatfull GF, et al. Corrected sequence of the bacteriophage P22 genome. Journal of bacteriology. 2003;185(4):1475-7.
350. Neal B, Brown P, Reeves P. Use of *Salmonella* phage P22 for transduction in *Escherichia coli*. Journal of bacteriology. 1993;175(21):7115-8.
351. Mmolawa PT, Schmieger H, Tucker CP, Heuzenroeder MW. Genomic structure of the *Salmonella enterica* serovar Typhimurium DT 64 bacteriophage ST64T: evidence for modular genetic architecture. Journal of bacteriology. 2003;185(11):3473-5.
352. Villafane R, Casjens S, Kropinski A. Sequence of *Salmonella enterica* serovar Anatum-specific bacteriophage Epsilon34. Unpublished results. 2005.
353. Casjens SR, Gilcrease EB, Winn-Stapley DA, Schicklmaier P, Schmieger H, Pedulla ML, et al. The generalized transducing *Salmonella* bacteriophage ES18: complete genome sequence and DNA packaging strategy. Journal of bacteriology. 2005;187(3):1091-104.
354. Kropinski AM, Kovalyova IV, Billington SJ, Patrick AN, Butts BD, Guichard JA, et al. The genome of ε 15, a serotype-converting, Group E1 *Salmonella enterica*-specific bacteriophage. Virology. 2007;369(2):234-44.
355. Sain A, Jayaprakash N. Draft genome sequence data of a T7like phage 3A_8767 isolated from wastewater of a butcher house near Palar river. Data in brief. 2020:105446.
356. Isaev A, Drobiazko A, Sierro N, Gordeeva J, Yosef I, Qimron U, et al. Phage T7 DNA mimic protein Ocr is a potent inhibitor of BREX defence. Nucleic acids research. 2020;48(10):5397-406.
357. Dobbins AT, George M, Basham DA, Ford ME, Houtz JM, Pedulla ML, et al. Complete genomic sequence of the virulent *Salmonella* bacteriophage SP6. Journal of bacteriology. 2004;186(7):1933-44.
358. Tu J, Park T, Morado DR, Hughes KT, Molineux IJ, Liu J. Dual host specificity of phage SP6 is facilitated by tailspike rotation. Virology. 2017;507:206-15.
359. Kuhn J, Suissa M, Chiswell D, Azriel A, Berman B, Shahar D, et al. A bacteriophage reagent for *Salmonella*: molecular studies on Felix 01. International journal of food microbiology. 2002;74(3):217-27.
360. Kropinski AM, Sulakvelidze A, Konczy P, Poppe C. *Salmonella* phages and prophages—genomics and practical aspects. *Salmonella*: Springer; 2007. p. 133-75.
361. Koskella B, Brockhurst MA. Bacteria–phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS microbiology reviews. 2014;38(5):916-31.
362. Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. Nature. 2010;464(7286):275.
363. Pal C, Maciá MD, Oliver A, Schachar I, Buckling A. Coevolution with viruses drives the evolution of bacterial mutation rates. Nature. 2007;450(7172):1079.

364. Gómez P, Buckling A. Bacteria-phage antagonistic coevolution in soil. Science. 2011;332(6025):106-9.
365. Hampton HG, Watson BN, Fineran PC. The arms race between bacteria and their phage foes. Nature. 2020;577(7790):327-36.
366. Bullas LR, Ryu J. *Salmonella* typhimurium LT2 strains which are r-m+ for all three chromosomally located systems of DNA restriction and modification. Journal of bacteriology. 1983;156(1):471-4.
367. Bullas L, Colson C, Neufeld B. Deoxyribonucleic acid restriction and modification systems in *Salmonella*: chromosomally located systems of different serotypes. Journal of bacteriology. 1980;141(1):275-92.
368. Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley EG. Characterization and evolution of *Salmonella* CRISPR-Cas systems. Pennsylvania State University State College; 2014.
369. Hofer B, Ruge M, Dreiseikelmann B. The superinfection exclusion gene (sieA) of bacteriophage P22: identification and overexpression of the gene and localization of the gene product. Journal of bacteriology. 1995;177(11):3080-6.
370. Goyal R, Chakravorty M. Abortive infection of the virulent phage 9NA in a fatty acid auxotroph of *Salmonella*typhimurium: Effect of fatty acid supplementation. Biochemical and biophysical research communications. 1989;161(2):923-30.
371. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak-Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. The EMBO Journal. 2015;34(2):169-83.
372. Gencay YE, Gambino M, From Prüssing T, Brøndsted L. The genera of bacteriophages and their receptors are the major determinants of host range. Environmental microbiology. 2019.
373. Andres D, Hanke C, Baxa U, Seul A, Barbirz S, Seckler R. Tailspike interactions with lipopolysaccharide effect DNA ejection from phage P22 particles in vitro. Journal of Biological Chemistry. 2010;285(47):36768-75.
374. Wang C, Nie T, Lin F, Connerton IF, Lu Z, Zhou S, et al. Resistance mechanisms adopted by a *Salmonella* Typhimurium mutant against bacteriophage. Virus research. 2019;273:197759.
375. Cota I, Sanchez-Romero MA, Hernandez SB, Pucciarelli MG, Garcia-Del Portillo F, Casadesus J. Epigenetic Control of *Salmonella enterica* O-Antigen Chain Length: A Tradeoff between Virulence and Bacteriophage Resistance. Plos Genet. 2015;11(11):e1005667.
376. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. Annual review of biochemistry. 2002;71(1):635-700.
377. Slauch JM, Mahan MJ, Michetti P, Neutra MR, Mekalanos JJ. Acetylation (O-factor 5) affects the structural and immunological properties of *Salmonella* typhimurium lipopolysaccharide O antigen. Infection and immunity. 1995;63(2):437-41.
378. Hauser E, Junker E, Helmuth R, Malorny B. Different mutations in the oafA gene lead to loss of O5 - antigen expression in *Salmonella enterica* serovar Typhimurium. Journal of applied microbiology. 2011;110(1):248-53.
379. Islam ST, Lam JS. Synthesis of bacterial polysaccharides via the Wzx/Wzy-dependent pathway. Canadian journal of microbiology. 2014;60(11):697-716.

380. Reeves PR, Cunneen MM, Liu B, Wang L. Genetics and evolution of the *Salmonella* galactose-initiated set of O antigens. PLoS One. 2013;8(7):e69306.
381. Kintz E, Davies MR, Hammarlöf DL, Canals R, Hinton JC, Woude MW. A BTP1 prophage gene present in invasive non‐typhoidal *Salmonella* determines composition and length of the O‐antigen of the lipopolysaccharide. Molecular microbiology. 2015;96(2):263-75.
382. Ho TD, Slauch JM. OmpC is the receptor for Gifsy-1 and Gifsy-2 bacteriophages of *Salmonella*. J Bacteriol. 2001;183(4):1495-8.
383. Cota I, Sánchez-Romero MA, Hernández SB, Pucciarelli MG, García-del Portillo F, Casadesús J. Epigenetic control of *Salmonella enterica* O-antigen chain length: a tradeoff between virulence and bacteriophage resistance. PLoS genetics. 2015;11(11):e1005667.
384. Luo F, Sun X, Qu Z, Zhang X. *Salmonella* typhimurium-induced M1 macrophage polarization is dependent on the bacterial O antigen. World Journal of Microbiology and Biotechnology. 2016;32(2):22.
385. Raetz CR, Whitfield C. Lipopolysaccharide endotoxins. Annu Rev Biochem. 2002;71:635-700.
386. Chow JC, Young DW, Golenbock DT, Christ WJ, Gusovsky F. Toll-like receptor-4 mediates lipopolysaccharide-induced signal transduction. Journal of Biological Chemistry. 1999;274(16):10689-92.
387. Shi J, Zhao Y, Wang Y, Gao W, Ding J, Li P, et al. Inflammatory caspases are innate immune receptors for intracellular LPS. Nature. 2014;514(7521):187.
388. Molineux I. Host-parasite interactions: recent developments in the genetics of abortive phage infections. The New biologist. 1991;3(3):230-6.
389. Bobonis J, Mitosch K, Mateus A, Kritikos G, Elfenbein JR, Savitski MM, et al. Phage proteins block and trigger retron toxin/antitoxin systems. BioRxiv. 2020.
390. Dy RL, Przybilski R, Semeijn K, Salmond GP, Fineran PC. A widespread bacteriophage abortive infection system functions through a Type IV toxin–antitoxin mechanism. Nucleic acids research. 2014;42(7):4590-605.
391. Tock MR, Dryden DT. The biology of restriction and anti-restriction. Current opinion in microbiology. 2005;8(4):466-72.
392. Seed KD. Battling phages: How bacteria defend against viral attack. PLoS pathogens. 2015;11(6):e1004847.
393. Blow MJ, Clark TA, Daum CG, Deutschbauer AM, Fomenkov A, Fries R, et al. The epigenomic landscape of prokaryotes. PLoS genetics. 2016;12(2):e1005854.
394. Thiaville JJ, Kellner SM, Yuan Y, Hutinet G, Thiaville PC, Jumpathong W, et al. Novel genomic island modifies DNA with 7-deazaguanine derivatives. Proceedings of the National Academy of Sciences. 2016;113(11):E1452-E9.
395. Atack JM, Guo C, Yang L, Zhou Y, Jennings MP. DNA sequence repeats identify numerous Type I restriction‐modification systems that are potential epigenetic regulators controlling phase‐variable regulons; phasevarions. The FASEB Journal. 2020;34(1):1038-51.
396. Xiong X, Wu G, Wei Y, Liu L, Zhang Y, Su R, et al. SspABCD–SspE is a phosphorothioation-sensing bacterial defence system with broad anti-phage activities. Nature Microbiology. 2020:1-12.

397. Wang L, Jiang S, Deng Z, Dedon PC, Chen S. DNA phosphorothioate modification—a new multi-functional epigenetic system in bacteria. FEMS Microbiology Reviews. 2019;43(2):109-22.
398. Roer L, Hendriksen RS, Leekitcharoenphon P, Lukjancenko O, Kaas RS, Hasman H, et al. Is the Evolution of *Salmonella enterica* subsp. *enterica* Linked to Restriction-Modification Systems? mSystems. 2016;1(3):e00009-16.
399. Wiedenheft B, Sternberg SH, Doudna JA. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012;482(7385):331-8.
400. Grissa I, Vergnaud G, Pourcel C. The CRISPRdb database and tools to display CRISPRs and to generate dictionaries of spacers and repeats. Bmc Bioinformatics. 2007;8(1):172.
401. Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley EG. Characterization and evolution of *Salmonella* CRISPR-Cas systems. Microbiol-Sgm. 2015;161:374-86.
402. Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, et al. CRISPR provides acquired resistance against viruses in prokaryotes. Science. 2007;315(5819):1709-12.
403. Mojica FJ, Diez-Villasenor C, Garcia-Martinez J, Soria E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. Journal of molecular evolution. 2005;60(2):174-82.
404. Barrangou R, Horvath P. A decade of discovery: CRISPR functions and applications. Nature microbiology. 2017;2(7):1-9.
405. Marraffini LA. CRISPR-Cas immunity against phages: its effects on the evolution and survival of bacterial pathogens. PLoS Pathog. 2013;9(12):e1003765.
406. Pawluk A, Davidson AR, Maxwell KL. Anti-CRISPR: discovery, mechanism and function. Nature Reviews Microbiology. 2018;16(1):12.
407. Common J, Walker‐Sünderhauf D, van Houte S, Westra ER. Diversity in CRISPR‐based immunity protects susceptible genotypes by restricting phage spread and evolution. Journal of Evolutionary Biology. 2020.
408. Mendoza SD, Nieweglowska ES, Govindarajan S, Leon LM, Berry JD, Tiwari A, et al. A bacteriophage nucleus-like compartment shields DNA from CRISPR nucleases. Nature. 2020;577(7789):244-8.
409. Malone LM, Warring SL, Jackson SA, Warnecke C, Gardner PP, Gumy LF, et al. A jumbo phage that forms a nucleus-like structure evades CRISPR–Cas DNA targeting but is vulnerable to type III RNA-based immunity. Nature Microbiology. 2020;5(1):48-55.
410. Borges AL. CRISPR vs. Anti-CRISPR: How bacterial viruses fight CRISPR-Cas immunity: UCSF; 2020.
411. Shariat N, Timme RE, Pettengill JB, Barrangou R, Dudley EG. Characterization and evolution of *Salmonella* CRISPR-Cas systems. Microbiology. 2015;161(2):374-86.
412. Westra ER, Buckling A, Fineran PC. CRISPR-Cas systems: beyond adaptive immunity. Nature reviews Microbiology. 2014;12(5):317-26.
413. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak-Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. The EMBO journal. 2015;34(2):169-83.

414. Hinton JCD. Functional Studies of Bacteriophage Exclusion [web page]. www.findaphd.com: www.findaphd.com; 2016 [This is a webpage requesting applications to work on the research topic, and implies that Non-typhoidal *Salmonella* possess BREX genes]. Available from: https://www.findaphd.com/search/projectdetails.aspx?PJID=69292.
415. Broniewski JM, Meaden S, Paterson S, Buckling A, Westra ER. The effect of phage genetic diversity on bacterial resistance evolution. The ISME Journal. 2020;14(3):828-36.
416. De Smet J, Hendrix H, Blasdel BG, Danis-Wlodarczyk K, Lavigne R. Pseudomonas predators: understanding and exploiting phage–host interactions. Nature Reviews Microbiology. 2017;15(9):517.
417. Doron S, Melamed S, Ofir G, Leavitt A, Lopatina A, Keren M, et al. Systematic discovery of antiphage defense systems in the microbial pangenome. Science. 2018;359(6379).
418. Kortright KE, Chan BK, Turner PE. High-throughput discovery of phage receptors using transposon insertion sequencing of bacteria. Proceedings of the National Academy of Sciences. 2020.
419. Cowley LA, Jenkins C, Sheppard SK. Phage predation shapes the population structure of Shiga-toxigenic *Escherichia coli* O157: H7 in the UK: an evolutionary perspective. Frontiers in genetics. 2019;10:763.
420. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. Proceedings of the National Academy of Sciences. 2000;97(12):6640-5.
421. Sharan SK, Thomason LC, Kuznetsov SG, Court DL. Recombineering: a homologous recombination-based method of genetic engineering. Nature protocols. 2009;4(2):206.
422. O'Callaghan D, Charbit A. High efficiency transformation of *Salmonella* typhimurium and *Salmonella* typhi by electroporation. Molecular and General Genetics MGG. 1990;223(1):156-8.
423. Tu Q, Yin J, Fu J, Herrmann J, Li Y, Yin Y, et al. Room temperature electrocompetent bacterial cells improve DNA transformation and recombineering efficiency. Scientific reports. 2016;6(1):1-8.
424. Kropinski AM, Mazzocco A, Waddell TE, Lingohr E, Johnson RP. Enumeration of bacteriophages by double agar overlay plaque assay. Bacteriophages: Springer; 2009. p. 69-76.
425. Rodwell EV, Wenner N, Pulford CV, Cai Y, Bowers-Barnard A, Beckett A, et al. Isolation and characterisation of bacteriophages with activity against invasive non-typhoidal *Salmonella* causing bloodstream infection in Malawi. bioRxiv. 2021.
426. Lemon J, Bolker B, Oom S, Klein E, Rowlingson B, Wickham H, et al. Package 'plotrix'. R Development Core Team, R: a. 2015.
427. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes De Novo Assembler. Current protocols in bioinformatics. 2020;70(1):e102.
428. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. Journal of computational biology. 2012;19(5):455-77.
429. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30(14):2068-9.
430. Guy L, Roat Kultima J, Andersson SG. genoPlotR: comparative gene and genome visualization in R. Bioinformatics. 2010;26(18):2334-5.

431. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics (Oxford, England). 2009;25(14):1754-60.
432. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. arXiv preprint arXiv:12073907. 2012.
433. Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, et al. The variant call format and VCFtools. Bioinformatics. 2011;27(15):2156-8.
434. Seemann T. Snippy: rapid haploid variant calling and core SNP phylogeny. 2015.
435. Wertheim JO, Sanderson MJ, Worobey M, Bjork A. Relaxed molecular clocks, the bias–variance trade-off, and the quality of phylogenetic inference. Systematic biology. 2010;59(1):1-8.
436. Liu K, Linder CR, Warnow T. RAxML and FastTree: comparing two methods for large-scale maximum likelihood phylogeny estimation. PloS one. 2011;6(11).
437. Price MN, Dehal PS, Arkin AP. FastTree 2–approximately maximum-likelihood trees for large alignments. PloS one. 2010;5(3).
438. Stamatakis A. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics. 2006;22(21):2688-90.
439. Tavaré S. Some probabilistic and statistical problems in the analysis of DNA sequences. Lectures on mathematics in the life sciences. 1986;17(2):57-86.
440. Posada D, Crandall KA. Modeltest: testing the model of DNA substitution. Bioinformatics (Oxford, England). 1998;14(9):817-8.
441. Sumner JG, Jarvis PD, Fernández-Sánchez J, Ferńandez-Sánchez J, Kaine BT, Woodhams MD, et al. Is the general time-reversible model bad for molecular phylogenetics? Systematic biology. 2012;61(6):1069-74.
442. Nguyen L-T, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular biology and evolution. 2015;32(1):268-74.
443. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. ViPTree: the viral proteomic tree server. Bioinformatics. 2017;33(15):2379-80.
444. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. Nucleic acids research. 2015;43(3):e15-e.
445. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Molecular biology and evolution. 2020;37(5):1530-4.
446. Tonkin-Hill G, Lees JA, Bentley SD, Frost SD, Corander J. RhierBAPS: An R implementation of the population clustering algorithm hierBAPS. Wellcome open research. 2018;3.
447. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis Comparison Tool. Bioinformatics. 2005;21(16):3422-3.
448. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. 1997;25(17):3389-402.

449. Finn R, Tate J, Mistry J, Coggill P, Sammut S, Hotz H-R, et al. The Pfam protein families database. Nucl Acids Res. 2008;36(suppl_1):D281-8.
450. Eddy SR. Accelerated profile HMM searches. PLoS Comput Biol. 2011;7(10):e1002195.
451. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947-8.
452. Arndt D, Grant JR, Marcu A, Sajed T, Pon A, Liang Y, et al. PHASTER: a better, faster version of the PHAST phage search tool. Nucleic acids research. 2016;44(W1):W16-W21.
453. Chen Y, Ye W, Zhang Y, Xu Y. High speed BLASTN: an accelerated MegaBLAST search tool. Nucleic acids research. 2015;43(16):7762-8.
454. Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21(16):3422-3.
455. Revell LJ, Revell MLJ. Package 'phytools'. 2019.
456. Mörters P, Peres Y. Brownian motion: Cambridge University Press; 2010.
457. Clavel J, Escarguel G, Merceron G. mvMORPH: an R package for fitting multivariate evolutionary models to morphometric data. Methods in Ecology and Evolution. 2015;6(11):1311-9.
458. Pagel M, Meade A, Barker D. Bayesian estimation of ancestral character states on phylogenies. Systematic biology. 2004;53(5):673-84.
459. Paradis E, Claude J, Strimmer K. APE: analyses of phylogenetics and evolution in R language. Bioinformatics. 2004;20(2):289-90.
460. Ishikawa SA, Zhukova A, Iwasaki W, Gascuel O. A fast likelihood method to reconstruct and visualize ancestral scenarios. Molecular biology and evolution. 2019;36(9):2069-85.
461. Bollback JP. SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC bioinformatics. 2006;7(1):88.
462. Revell LJ. phytools: an R package for phylogenetic comparative biology (and other things). Methods in Ecology and Evolution. 2012;3(2):217-23.
463. Letunic I, Bork P. Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. Bioinformatics. 2006;23(1):127-8.
464. Hunt M, Mather AE, Sánchez-Busó L, Page AJ, Parkhill J, Keane JA, et al. ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads. Microbial Genomics. 2017;3(10).
465. Didelot X, Croucher NJ, Bentley SD, Harris SR, Wilson DJ. Bayesian inference of ancestral dates on bacterial phylogenetic trees. Nucleic acids research. 2018;46(22):e134-e.
466. Kassambara A. ggpubr:"ggplot2" based publication ready plots. R package version 0.1. 6. 2017.
467. Lees JA, Galardini M, Bentley SD, Weiser JN, Corander J. pyseer: a comprehensive tool for microbial pangenome-wide association studies. Bioinformatics. 2018;34(24):4310-2.
468. Chen PE, Shapiro BJ. The advent of genome-wide association studies for bacteria. Current opinion in microbiology. 2015;25:17-24.
469. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: fast genome and metagenome distance estimation using MinHash. Genome biology. 2016;17(1):132.
470. Lees JA, Vehkala M, Valimaki N, Harris SR, Chewapreecha C, Croucher NJ, et al. Sequence element enrichment analysis to determine the genetic basis of bacterial phenotypes. Nat Commun. 2016;7:12797.

471. Jaillard M, Lima L, Tournoud M, Mahé P, Van Belkum A, Lacroix V, et al. A fast and agnostic method for bacterial genome-wide association studies: Bridging the gap between k-mers and genetic events. PLoS genetics. 2018;14(11):e1007758.
472. Langmead B, Salzberg S. Langmead. 2013. Bowtie2. Nature Methods. 2013;9:357-9.
473. Quinlan AR. BEDTools: the Swiss‐army tool for genome feature analysis. Current protocols in bioinformatics. 2014;47(1):11.2. 1-.2. 34.
474. Yu G, Smith DK, Zhu H, Guan Y, Lam TTY. ggtree: an R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 2017;8(1):28-36.
475. Yu G, Smith DK, Zhu H, Guan Y, Lam TT-Y, McInerny G. ggtree: anrpackage for visualization and annotation of phylogenetic trees with their covariates and other associated data. Methods in Ecology and Evolution. 2017;8(1):28-36.
476. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3—new capabilities and interfaces. Nucleic acids research. 2012;40(15):e115-e.
477. Team RC, Worldwide C. The R stats package. R Foundation for Statistical Computing, Vienna, Austria: Available from: http://www R-project org. 2002.
478. Revelle W, Revelle MW. Package 'psych'. The comprehensive R archive network. 2015.
479. Scholz F, Zhu A. kSamples-package: The Package kSamples Contains Several Nonparametric K-Sample. 2018.
480. Revelle W. An overview of the psych package. Dep Psychol Northwest Univ. 2011;3:1-25.
481. Shapiro SS, Francia R. An approximate analysis of variance test for normality. Journal of the American Statistical Association. 1972;67(337):215-6.
482. Massey Jr FJ. The Kolmogorov-Smirnov test for goodness of fit. Journal of the American statistical Association. 1951;46(253):68-78.
483. Öztürk Ö, Wolfe DA. An improved ranked set two‐sample mann‐whitney‐wilcoxon test. Canadian Journal of Statistics. 2000;28(1):123-35.
484. Wallace DL. Simplified beta-approximations to the Kruskal-Wallis H test. Journal of the American Statistical Association. 1959;54(285):225-30.
485. Anisimova M, Bielawski JP, Yang Z. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Molecular biology and evolution. 2001;18(8):1585-92.
486. Berger JO, Pericchi LR. The intrinsic Bayes factor for model selection and prediction. Journal of the American Statistical Association. 1996;91(433):109-22.
487. Kalyaanamoorthy S, Minh BQ, Wong TK, Von Haeseler A, Jermiin LS. ModelFinder: fast model selection for accurate phylogenetic estimates. Nature methods. 2017;14(6):587-9.
488. Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic acids research. 2008;36(suppl_2):W5-W9.
489. Hochhut B, Waldor MK. Site-specific integration of the conjugal Vibrio cholerae SXT element into prfC. Mol Microbiol. 1999;32(1):99-110.
490. Wozniak RA, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nature Reviews Microbiology. 2010;8(8):552.

491. Lewis JA, Hatfull GF. Control of directionality in integrase-mediated recombination: examination of recombination directionality factors (RDFs) including Xis and Cox proteins. Nucleic acids research. 2001;29(11):2205-16.
492. Williams SL, Schildbach JF. Examination of an inverted repeat within the F factor origin of transfer: context dependence of F TraI relaxase DNA specificity. Nucleic acids research. 2006;34(2):426-35.
493. Salyers AA, Shoemaker NB, Stevens AM, Li L-Y. Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. Microbiological reviews. 1995;59(4):579-90.
494. Mierzejewska J, Jagura-Burdzy G. Prokaryotic ParA–ParB–parS system links bacterial chromosome segregation with the cell cycle. Plasmid. 2012;67(1):1-14.
495. Wong JJ, Lu J, Glover JM. Relaxosome function and conjugation regulation in F‐like plasmids–a structural biology perspective. Molecular microbiology. 2012;85(4):602-17.
496. Firth N, Skurray R. Characterization of the F plasmid bifunctional conjugation gene, traG. Molecular and General Genetics MGG. 1992;232(1):145-53.
497. Guglielmini J, de la Cruz F, Rocha EP. Evolution of conjugation and type IV secretion systems. Molecular biology and evolution. 2013;30(2):315-31.
498. Li X, Xie Y, Liu M, Tai C, Sun J, Deng Z, et al. oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. Nucleic Acids Res [Internet]. 2018 May 4; 46(Web Server issue):[W229-W34 pp.]. Available from: https://www.ncbi.nlm.nih.gov/pubmed/29733379.
499. Garbinski LD, Rosen BP, Chen J. Pathways of arsenic uptake and efflux. Environment international. 2019;126:585-97.
500. Berset Y, Merulla D, Joublin A, Hatzimanikatis V, Van Der Meer JR. Mechanistic modeling of genetic circuits for ArsR arsenic regulation. ACS synthetic biology. 2017;6(5):862-74.
501. Debarbouille M, Martin-Verstraete I, Kunst F, Rapoport G. The Bacillus subtilis sigL gene encodes an equivalent of sigma 54 from gram-negative bacteria. Proceedings of the National Academy of Sciences. 1991;88(20):9092-6.
502. Abdeen A, Miki B. The pleiotropic effects of the bar gene and glufosinate on the Arabidopsis transcriptome. Plant biotechnology journal. 2009;7(3):266-82.
503. Thompson CJ, Movva NR, Tizard R, Crameri R, Davies JE, Lauwereys M, et al. Characterization of the herbicide‐resistance gene bar from Streptomyces hygroscopicus. The EMBO journal. 1987;6(9):2519-23.
504. Delmar JA, Su C-C, Edward WY. Structural mechanisms of heavy-metal extrusion by the Cus efflux system. Biometals. 2013;26(4):593-607.
505. Randall CP, Gupta A, Jackson N, Busse D, O'neill AJ. Silver resistance in Gram-negative bacteria: a dissection of endogenous and exogenous mechanisms. Journal of Antimicrobial Chemotherapy. 2015;70(4):1037-46.
506. Osman D, Waldron KJ, Denton H, Taylor CM, Grant AJ, Mastroeni P, et al. Copper homeostasis in *Salmonella* is atypical and copper-CueP is a major periplasmic metal complex. Journal of Biological Chemistry. 2010;285(33):25259-68.

507. Silver S. Bacterial silver resistance: molecular biology and uses and misuses of silver compounds. FEMS microbiology reviews. 2003;27(2-3):341-53.
508. Massani MB, Klumpp J, Widmer M, Speck C, Nisple M, Lehmann R, et al. Chromosomal Sil system contributes to silver resistance in E. coli ATCC 8739. BioMetals. 2018;31(6):1101-14.
509. Lee SM, Grass G, Rensing C, Barrett SR, Yates CJ, Stoyanov JV, et al. The Pco proteins are involved in periplasmic copper handling in *Escherichia coli*. Biochemical and biophysical research communications. 2002;295(3):616-20.
510. Dupont CL, Grass G, Rensing C. Copper toxicity and the origin of bacterial resistance—new insights and applications. Metallomics. 2011;3(11):1109-18.
511. Staehlin BM, Gibbons JG, Rokas A, O'Halloran TV, Slot JC. Evolution of a heavy metal homeostasis/resistance island reflects increasing copper stress in Enterobacteria. Genome biology and evolution. 2016;8(3):811-26.
512. Su CC, Long F, Yu EW. The Cus efflux system removes toxic ions via a methionine shuttle. Protein Science. 2011;20(1):6-18.
513. Gupta A, Matsui K, Lo J-F, Silver S. Molecular basis for resistance to silver cations in *Salmonella*. Nature medicine. 1999;5(2):183-8.
514. Husnik F, McCutcheon JP. Functional horizontal gene transfer from bacteria to eukaryotes. Nature Reviews Microbiology. 2018;16(2):67.
515. Bañuelos-Vazquez LA, Tejerizo GT, Brom S. Regulation of conjugative transfer of plasmids and integrative conjugative elements. Plasmid. 2017;91:82-9.
516. Delavat F, Miyazaki R, Carraro N, Pradervand N, van der Meer JR. The hidden life of integrative and conjugative elements. FEMS microbiology reviews. 2017;41(4):512-37.
517. Checa SK, Giri GF, Espariz M, Argüello JM, Soncini FC. Copper Handling in the *Salmonella* Cell Envelope and Its Impact on Virulence. Trends in Microbiology. 2021.
518. Outten FW, Huffman DL, Hale JA, O'Halloran TV. The independent cue and cusSystems confer copper tolerance during aerobic and anaerobic growth in*Escherichia coli*. Journal of Biological Chemistry. 2001;276(33):30670-7.
519. Macomber L, Imlay JA. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proceedings of the National Academy of Sciences. 2009;106(20):8344-9.
520. Tan G, Yang J, Li T, Zhao J, Sun S, Li X, et al. Anaerobic copper toxicity and iron-sulfur cluster biogenesis in *Escherichia coli*. Applied and environmental microbiology. 2017;83(16).
521. Harris ED. Copper as a cofactor and regulator of copper, zinc superoxide dismutase. The Journal of nutrition. 1992;122(suppl_3):636-40.
522. Knop M, Dang TQ, Jeschke G, Seebeck FP. Copper is a Cofactor of the Formylglycine - Generating Enzyme. ChemBioChem. 2017;18(2):161.
523. Ammendola S, Pasquali P, Pacello F, Rotilio G, Castor M, Libby SJ, et al. Regulatory and structural differences in the Cu, Zn-superoxide dismutases of *Salmonella enterica* and their significance for virulence. Journal of Biological Chemistry. 2008;283(20):13688-99.
524. Chabert V, Hologne M, Sénèque O, Crochet A, Walker O, Fromm KM. Model peptide studies of Ag+ binding sites from the silver resistance protein SilE. Chemical Communications. 2017;53(45):6105-8.

525. Achard ME, Tree JJ, Holden JA, Simpfendorfer KR, Wijburg OL, Strugnell RA, et al. The multi-copper-ion oxidase CueO of *Salmonella enterica* serovar Typhimurium is required for systemic virulence. Infection and immunity. 2010;78(5):2312-9.
526. Maillard J-Y, Hartemann P. Silver as an antimicrobial: facts and gaps in knowledge. Critical reviews in microbiology. 2013;39(4):373-83.
527. Martin P, DeMel S, Shi J, Gladysheva T, Gatti DL, Rosen BP, et al. Insights into the structure, solvation, and mechanism of ArsC arsenate reductase, a novel arsenic detoxification enzyme. Structure. 2001;9(11):1071-81.
528. Tisa LS, Rosen BP. Molecular characterization of an anion pump. The ArsB protein is the membrane anchor for the ArsA protein. Journal of Biological Chemistry. 1990;265(1):190-4.
529. Zhou T, Radaev S, Rosen BP, Gatti DL. Structure of the ArsA ATPase: the catalytic subunit of a heavy metal resistance pump. The EMBO journal. 2000;19(17):4838-45.
530. Ajees AA, Yang J, Rosen BP. The ArsD as (III) metallochaperone. Biometals. 2011;24(3):391-9.
531. Murphy JN, Saltikov CW. The ArsR repressor mediates arsenite-dependent regulation of arsenate respiration and detoxification operons of Shewanella sp. strain ANA-3. Journal of bacteriology. 2009;191(21):6722-31.
532. Arai N, Sekizuka T, Tamamura Y, Kusumoto M, Hinenoya A, Yamasaki S, et al. *Salmonella* genomic island 3 is an integrative and conjugative element and contributes to copper and arsenic tolerance of *Salmonella enterica*. Antimicrobial agents and chemotherapy. 2019:AAC. 00429-19.
533. Medardus JJ, Molla BZ, Nicol M, Morrow WM, Rajala-Schultz PJ, Kazwala R, et al. In-feed use of heavy metal micronutrients in U.S. swine production systems and its role in persistence of multidrug-resistant *Salmonella*e. Appl Environ Microbiol. 2014;80(7):2317-25.
534. Bawn M, Alikhan NF, Thilliez G, Kirkwood M, Wheeler NE, Petrovska L, et al. Evolution of *Salmonella enterica* serotype Typhimurium driven by anthropogenic selection and niche adaptation. PLoS Genet. 2020;16(6):e1008850.
535. Mather A, Reid S, Maskell D, Parkhill J, Fookes M, Harris S, et al. Distinguishable epidemics of multidrug-resistant *Salmonella* Typhimurium DT104 in different hosts. Science. 2013;341(6153):1514-7.
536. Tassinari E, Bawn M, Thilliez G, Charity O, Acton L, Kirkwood M, et al. Whole-genome epidemiology links phage-mediated acquisition of a virulence gene to the clonal expansion of a pandemic *Salmonella enterica* serovar Typhimurium clone. Microbial Genomics. 2020;6(11):mgen000456.
537. Cornell J, Neal K. Protracted outbreak of *Salmonella* typhimurium definitive phage type 170 food poisoning related to tripe,'pig bag', and chitterlings. Commun Dis Public Health. 1998;1(1):28-30.
538. Hooton SP, Timms AR, Moreton J, Wilson R, Connerton IF. Complete genome sequence of *Salmonella enterica* serovar Typhimurium U288. Genome announcements. 2013;1(4).
539. McLaren IM, Wray C. Epidemiology of *Salmonella* typhimurium infection in calves: persistence of *Salmonella*e on calf units. Vet Rec. 1991;129(21):461-2.

540. Barone L, Dal VA, Pellissier N, Vigano A, Romani C, Pontello M. Emergence of *Salmonella* Typhimurium monophasic serovar: determinants of antimicrobial resistance in porcine and human strains. Annali di igiene : medicina preventiva e di comunita. 2008;20(3):199-209.
541. Sun R-Y, Ke B-X, Fang L-X, Guo W-Y, Li X-P, Yu Y, et al. Global clonal spread of mcr-3-carrying MDR ST34 *Salmonella enterica* serotype Typhimurium and monophasic 1, 4,[5], 12: i:− variants from clinical isolates. Journal of Antimicrobial Chemotherapy. 2020;75(7):1756-65.
542. Pennycott TW, Park A, Mather HA. Isolation of different serovars of *Salmonella enterica* from wild birds in Great Britain between 1995 and 2003. Vet Rec. 2006;158(24):817-20.
543. Mohammed M, Cormican M. Whole genome sequencing provides possible explanations for the difference in phage susceptibility among two *Salmonella* Typhimurium phage types (DT8 and DT30) associated with a single foodborne outbreak. BMC research notes. 2015;8(1):1-14.
544. Soyer Y, Switt AM, Davis MA, Maurer J, McDonough PL, Schoonmaker-Bopp DJ, et al. *Salmonella enterica* Serotype 4,5,12:i:-, an Emerging *Salmonella* Serotype That Represents Multiple Distinct Clones. Journal of Clinical Microbiology. 2009;47(11):3546-56.
545. Bondy-Denomy J, Qian J, Westra ER, Buckling A, Guttman DS, Davidson AR, et al. Prophages mediate defense against phage infection through diverse mechanisms. The ISME journal. 2016;10(12):2854-66.
546. Song S, Wood TK. A primary physiological role of toxin/antitoxin systems is phage inhibition. Frontiers in Microbiology. 2020;11.
547. Dziewit L, Oscik K, Bartosik D, Radlinska M. Molecular characterization of a novel temperate Sinorhizobium bacteriophage, ΦLM21, encoding DNA methyltransferase with CcrM-like specificity. Journal of virology. 2014;88(22):13111-24.
548. Seed KD, Lazinski DW, Calderwood SB, Camilli A. A bacteriophage encodes its own CRISPR/Cas adaptive response to evade host innate immunity. Nature. 2013;494(7438):489-91.
549. Arndt D, Marcu A, Liang Y, Wishart DS. PHAST, PHASTER and PHASTEST: tools for finding prophage in bacterial genomes. Briefings in Bioinformatics. 2019;20(4):1560-7.
550. Joy JB, Liang RH, McCloskey RM, Nguyen T, Poon AF. Ancestral reconstruction. PLoS computational biology. 2016;12(7):e1004763.
551. Smaers JB, Mongle CS, Kandler A. A multiple variance Brownian motion framework for estimating variable rates and inferring ancestral states. Biological Journal of the Linnean Society. 2016;118(1):78-94.
552. Smaers J, Mongle C. On the accuracy and theoretical underpinnings of the multiple variance Brownian motion approach for estimating variable rates and inferring ancestral states. Biological Journal of the Linnean Society. 2017;121(1):229-38.
553. Anderson ES, Ward LR, Saxe MJ, de Sa JD. Bacteriophage-typing designations of *Salmonella* Typhimurium. J Hyg (Lond). 1977;78(2):297-300.
554. Bollback JP. SIMMAP: stochastic character mapping of discrete traits on phylogenies. BMC bioinformatics. 2006;7(1):1-7.
555. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Research. 2019;47(W1):W256-W9.

556. Sun Q, Lan R, Wang Y, Wang J, Wang Y, Li P, et al. Isolation and genomic characterization of SfI, a serotype-converting bacteriophage of Shigella flexneri. BMC microbiology. 2013;13(1):1-8.
557. George DT, Stephenson DP, Tran E, Morona R, Verma NK. Complete genome sequence of SfII, a serotype-converting bacteriophage of the highly prevalent Shigella flexneri serotype 2a. Genome announcements. 2013;1(5).
558. Jakhetia R, Talukder KA, Verma NK. Isolation, characterization and comparative genomics of bacteriophage SfIV: a novel serotype converting phage from Shigella flexneri. BMC genomics. 2013;14(1):1-8.
559. Zhang Y, Yamaguchi Y, Inouye M. Characterization of YafO, an *Escherichia coli* toxin. Journal of Biological Chemistry. 2009;284(38):25522-31.
560. Bindal G, Krishnamurthi R, Seshasayee ASN, Rath D. CRISPR-Cas-mediated gene silencing reveals RacR to be a negative regulator of YdaS and YdaT toxins in *Escherichia coli* K-12. Msphere. 2017;2(6).
561. Krishnamurthi R, Ghosh S, Khedkar S, Seshasayee ASN. Repression of YdaS toxin is mediated by transcriptional repressor RacR in the cryptic rac prophage of *Escherichia coli* K-12. Msphere. 2017;2(6).
562. Jobling MG. Ectopic expression of the ydaS and ydaT genes of the cryptic prophage Rac of *Escherichia coli* K-12 may be toxic but do they really encode toxins?: a case for using genetic context to understand function. Msphere. 2018;3(2).
563. Singletary LA, Gibson JL, Tanner EJ, McKenzie GJ, Lee PL, Gonzalez C, et al. An SOS-regulated type 2 toxin-antitoxin system. Journal of bacteriology. 2009;191(24):7456-65.
564. Gerdes K, Christensen SK, Løbner-Olesen A. Prokaryotic toxin–antitoxin stress response loci. Nature Reviews Microbiology. 2005;3(5):371-82.
565. Van Melderen L. Toxin–antitoxin systems: why so many, what for? Current opinion in microbiology. 2010;13(6):781-5.
566. Yamaguchi Y, Park J-H, Inouye M. Toxin-antitoxin systems in bacteria and archaea. Annual review of genetics. 2011;45:61-79.
567. Unterholzner SJ, Poppenberger B, Rozhon W. Toxin–antitoxin systems: biology, identification, and application. Mobile genetic elements. 2013;3(5):e26219.
568. Yamaguchi Y, Inouye M. Regulation of growth and death in *Escherichia coli* by toxin–antitoxin systems. Nature Reviews Microbiology. 2011;9(11):779-90.
569. Goeders N, Van Melderen L. Toxin-antitoxin systems as multilevel interaction systems. Toxins. 2014;6(1):304-24.
570. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. The annals of mathematical statistics. 1947:50-60.
571. Wilcoxon F. Individual comparisons of grouped data by ranking methods. Journal of economic entomology. 1946;39(2):269-70.
572. Vasu K, Nagaraja V. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. Microbiology and Molecular Biology Reviews. 2013;77(1):53-72.

573. Atack JM, Yang Y, Seib KL, Zhou Y, Jennings MP. A survey of Type III restriction-modification systems reveals numerous, novel epigenetic regulators controlling phase-variable regulons; phasevarions. Nucleic acids research. 2018;46(7):3532-42.
574. Harms A, Brodersen DE, Mitarai N, Gerdes K. Toxins, targets, and triggers: an overview of toxin-antitoxin biology. Molecular cell. 2018;70(5):768-84.
575. Lobato-Márquez D, Moreno-Córdoba I, Figueroa V, Díaz-Orejas R, García-del Portillo F. Distinct type I and type II toxin-antitoxin modules control *Salmonella* lifestyle inside eukaryotic cells. Scientific reports. 2015;5(1):1-10.
576. Kado CI, Liu ST. Rapid procedure for detection and isolation of large and small plasmids. J Bacteriol. 1981;145(3):1365-73.
577. Goldfarb T, Sberro H, Weinstock E, Cohen O, Doron S, Charpak‐Amikam Y, et al. BREX is a novel phage resistance system widespread in microbial genomes. The EMBO journal. 2015;34(2):169-83.
578. Woolf B. The log likelihood ratio test (the G‐test). Annals of human genetics. 1957;21(4):397-409.
579. Wang L, Andrianopoulos K, Liu D, Popoff MY, Reeves PR. Extensive variation in the O-antigen gene cluster within one *Salmonella enterica* serogroup reveals an unexpected complex history. Journal of bacteriology. 2002;184(6):1669-77.
580. Puymège A, Bertin S, Chuzeville S, Guédon G, Payot S. Conjugative transfer and cis-mobilization of a genomic island by an integrative and conjugative element of Streptococcus agalactiae. Journal of bacteriology. 2013;195(6):1142-51.
581. Guédon G, Libante V, Coluzzi C, Payot S, Leblond-Bourget N. The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. Genes. 2017;8(11):337.
582. Piao HH, Tam VTM, Na HS, Kim HJ, Ryu PY, Kim SY, et al. Immunological responses induced by asd and wzy/asd mutant strains of *Salmonella enterica* serovar Typhimurium in BALB/c mice. The Journal of Microbiology. 2010;48(4):486-95.
583. Kong Q, Yang J, Liu Q, Alamuri P, Roland KL, Curtiss R. Effect of deletion of genes involved in lipopolysaccharide core and O-antigen synthesis on virulence and immunogenicity of *Salmonella enterica* serovar Typhimurium. Infection and immunity. 2011;79(10):4227-39.
584. Oliveira J, Mahony J, Hanemaaijer L, Kouwen TR, Neve H, MacSharry J, et al. Detecting Lactococcus lactis prophages by mitomycin C-mediated induction coupled to flow cytometry analysis. Frontiers in microbiology. 2017;8:1343.
585. Cadieux B, Colavecchio A, Jeukens J, Freschi L, Emond-Rheault J-G, Kukavica-Ibrulj I, et al. Prophage induction reduces Shiga toxin producing *Escherichia coli* (STEC) and *Salmonella enterica* on tomatoes and spinach: A model study. Food Control. 2018;89:250-9.
586. Garcia-Russell N, Elrod B, Dominguez K. Stress-induced prophage DNA replication in *Salmonella enterica* serovar Typhimurium. Infection, Genetics and Evolution. 2009;9(5):889-95.
587. Evans M, Salmon R, Nehaul L, Mably S, Wafford L, Nolan-Farrell M, et al. An outbreak of *Salmonella* typhimurium DT170 associated with kebab meat and yoghurt relish. Epidemiology & Infection. 1999;122(3):377-83.

588. Chow LK, Ghaly TM, Gillings MR. A survey of sub-inhibitory concentrations of antibiotics in the environment. Journal of Environmental Sciences. 2021;99:21-7.

Appendix I: The Anderson Phage Typing Scheme

As provided by Public Health England

Appendix I.1 Terminology used for the outcome of phage infection of bacterial isolates to be tested

Phage typing terminology:

Individual plaque sizes:

L = large, above 1-5 mm.,

N = normal, about 1-0 mm.,

S = small, 1-0*1 mm.

m = minute, less than 0a I mm. only with hand lens,

u = micro plaques, barely visable

+ to + + + = increasing numbers of discrete plaques.

Degrees of confluent lysis:

CL = confluent lysis,

SCL = semi-confluent lysis,

<CL

< SCL intermediate degrees of lysis,

OL = confluent lysis with a heavy central opacity due to secondary

(Lysogenized)growth.

Appendix I.2 Outcomes and Phage Type Designations when Challenging Bacterial Isolates with 30 Typing Phage Preparations

Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35	Pool O
1	CL	CL	CL	CL	CL	CL	CL	-	+++	CL	CL	CL	CL	CL	CL	CL	CL	CL	+++	CL	CL	CL	CL	CL	CL	CL	-/+	CL	+++	CL	CL
2	-	CL	CL	CL	CL	CL	-	-	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL
2a	-	<CL	CL	OL	CL	CL	-	-	CL	CL	-/++	-/++	CL	CL	CL	CL	-	CL	CL	OL	OL	CL	CL	CL	CL	CL	-/+	CL	CL	CL	CL
3	-	+++	CL	CL	CL	CL	-	-	CL	CL	-	-	CL	SCL	CL	CL	-	SCL	CL	SCL	+++	CL	CL	+++	+++	CL	-	CL	+++	CL	CL
3	-	+++	CL	CL	CL	CL	-	-	CL	CL	-	-	CL	SCL	CL	CL	-	SCL	CL	SCL	+++	CL	CL	+++	+++	CL	-	CL	+++	CL	CL
4	-	-	-	CL	CL	CL	-	-	SCL	OL	++	++	-	CL	CL	-	-	CL	CL	±	CL	CL	-	SCL	+	CL	-	CL	CL	CL	CL
5	-	-	-	+	OL	<CL	-	-	-	-/+++	-	-	-	-/+++	+/+++	-	-	+/+++	-	++	++	-	-	-	-	-	-	-/+	-	-	SCL
6	-	-	-	-	+	OL	-	-	-/+	-/+	-	-	-	-	-/<CL	-/+	-	-	-	-	-	-	-	-	-	-/+++	-	OL	-	-/+++	CL
7	-	-	-	-	-	-	CL	±	-	-	-	-	-	-	-	-	CL	-	SCL	-	-	-	-	-	-	-	-	-	SCL	-	CL
7a	-	-	-	-	-	-	CL	±	-	-	-	-	-	-	-	-	CL	-	SCL	-	-	-	-	-	-	-	-	SCL	SCL	-	CL
8	-	-	-	-	-	-	-	CL	SCL	++/SCL	-	-	-	-	+++	-	-	-	SCL	-	SCL	SCL	-	±	±	-	-	CL	CL	-	CL
9	-	-	-	-	-	-	-	CL	SCL	CL	CL	CL	-	-	SCL	-	-	-	CL	±	CL	CL	-	±	+	-	-	CL	CL	-	CL
10	-	-	-	-	-	-	-	-	SCL	CL	CL	CL	-	-	+++	-	-	-	CL	±	CL	CL	-	±	±	-	-	CL	CL	-	CL
11	-	-	-	-	-	-	-	-	-	SCL	+/+++	++	-	-	-	SCL	CL	OL	+++	-	-	-	-	-	-	-	-	-	CL	-	CL
12	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	-	-	-	±	-	-	-	-	-	-	-	-	-	-	-	CL
12a	-	-	-	-	-	-	-	±	-	-	CL	CL	-	-	-	-	CL	-	-	++	-	-	-	-	-	OL	-	-	-	OL	CL
13	-	-	-	-	-	-	-	-	-	OL	-	CL	OL	-	-/+++	-	-/+++	-	-	-	-	-	-	-	SCL	OL	-	OL	-	OL	CL
14	-	-	-	-	-	-	-	-	-	<SCL	+	++	OL	-	-/++	-	-	-	-	-	-	-	-	-	SCL	OL	-	OL	-	OL	CL
15	-	-	-	-	-	-	-	-	-	+++	CL	SCL	-	OL	-	SCL	CL	OL	SCL	-	-	-	-	-	-	+	±	-	OL	+/SCL	CL
15a	-	-	-	-	-	-	-	-	-	+++	CL	CL	-	OL	-	OL	-	OL	OL	-	-	-	-	-	-	+	±	-	OL	+/SCL	CL
16	-	-	-	-	-	-	-	-	-	SCL	-	-	OL	-	OL	-/+++	-	-	-	-	-	-	-	+++	+++	OL	-	OL	-	OL	CL
17	-	-	-	-	-	-	-	-	-	+++	-	-	-	OL	-	OL	-	OL	OL	-	-	-	-	-	++	±	-	-	OL	-/+++	CL
18	-	-	-	-	-	-	-	-	-	+++	-	-	-	OL	-	OL	CL	OL	SCL	-	-	-	-	-	++	±	-	-	OL	+/SCL	CL
19	-	-	-	-	-	-	-	-	-	+++	-	-	-	OL	-	-	-	OL	SCL	-	-	-	-	-	++	±	-	-	OL	+++	CL
20	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	-	-	-	OL	-	-	-	-	-	-	-	-	-	OL	-	CL
23	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL	CL	+++	+++	-	-	+/+++	±	OL	CL
24	-	-	CL	-	-	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	SCL	-	SCL	-	-	-	-	-	CL
25	-	-	-	-	-	-	-	-	-	-	-	-	-	-	±	-	-	-	-	-	-	-	++	OL	OL	-	-	+	-	OL	CL
26	-	-	-	-	-	-	-	-	-	-	-	-	+/SCL	-	-	-	-	-	-	-	-	-	-/+++	-/+++	OL	-	-	-	-	OL	-
27	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-/+++	-	-	-	-	-	OL	-	-	-	+	CL
28	-	-	-	-	-	-	-	OL	-	+++	SCL	SCL	-	OL	-	OL	CL	OL	SCL	-	-	-	-	-	-	++	OL	-	SCL	SCL	CL
29	-	-	-	-	-	-	-	-	±	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-/+++	-	SCL/CL	-	-/+++	SCL
30	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
31	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+++	-	<OL	-	-	-	-/SCL	-	OL	CL
32	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	±	-	-	-	-	-	-	-	-	-	OL	-	CL
35	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	OL	CL
36	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	<CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	SCL	CL	CL	OL	CL	OL	CL
37	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL
38	+++	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	SCL
39	<SCL	CL	SCL	CL	CL	CL	CL	-	<CL	CL	CL	CL	CL	CL	CL	SCL	-	CL	CL	CL	CL	CL	CL	CL	SCL	CL	+	CL	<CL	OL	CL
40	SCL	<CL	<CL	CL	CL	CL	CL	-	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL
41	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	-	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL
41a	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	-	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	OL	CL	CL	CL	CL
41b	+++	++	<CL	CL	CL	+++	<SCL	-	<SCL	CL	-	-	CL	CL	CL	CL	+++	<CL	+++	OL	OL	SCL	<CL	CL	++	<CL	-	CL	CL	OL	-
42	<SCL	CL	SCL	-/+++	-/+++	-/+++	SCL	-	-	CL	+++	+++	OL	OL	+/+++	CL	-/SCL	-/SCL	-/++	OL	CL	-/++	<CL	++	<CL	SCL	-	OL	-	OL	CL
43	CL	CL	CL	-	-	-	CL	-	CL	CL	-	-	-	+	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	-	-	CL	CL	-
44	+/+++	+++	+++	-	-	-	SCL	-	+++	+/+++	-	-	CL	-	+++	+++	+++	-	SCL	OL	SCL	OL	CL	+	+/+++	OL	-	OL	SCL	CL	CL
45	+	OL	OL	OL	OL	OL	+++	-	+	OL	-	-	OL	OL	SCL	-	-	OL	OL	OL	OL	SCL	OL	OL	SCL	OL	-	OL	OL	OL	CL
46	-	CL	CL	CL	CL	CL	-	-	CL	CL	CL	CL	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	+++
46a	-	CL	CL	CL	CL	CL	-	-	CL	CL	-	-	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL	CL
47	-	OL	+++	SCL	SCL	SCL	-	-	+++	SCL	CL	CL	CL	+++	SCL	CL	-	SCL	SCL	CL	CL	CL	SCL	+++	CL	-	-	CL	+++	-	CL
49	-	+++	++	-	-	SCL	-	-	SCL	+++	+	++	CL	±	CL	SCL	-	±	CL	CL	+++	CL	CL	±	+	OL	±	CL	+++	OL	CL
49a	-	+++	++	-	-	SCL	-	-	SCL	+++	+	++	CL	±	CL	SCL	-	±	CL	CL	+++	CL	CL	±	+	OL	OL	CL	+++	OL	CL
49b	-	+++	++	-	-	SCL	-	-	SCL	+++	-	-	CL	±	CL	SCL	-	±	CL	CL	+++	CL	CL	±	+	OL	±	CL	+++	OL	CL
50	-	CL	-	+++	++	+++	++	-	+	++	-	-	+++	+++	+++	+++	-	+++	++	CL	CL	++	++	++	-	+++	-	CL	+	SCL	CL

51	-	+++	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	-	-	-	-	-	-	CL
52	-	-	-	CL	CL	+++	-	-	CL	+++	-	-	-	CL	CL	-	-	SCL	SCL	-	SCL	CL	-	+++	-	CL	CL	CL	CL	CL	+++
52a	-	-	-	CL	CL	+++	-	-	CL	+++	-	-	-	CL	CL	-	-	SCL	SCL	-	SCL	CL	-	+++	-	CL	+	CL	CL	CL	+++
53	-	-	-	+++	+++	SCL	++	-	-	OL	-	-	+++	SCL	OL	-	++	SCL	-	-	SCL	SCL	-	OL	++	OL	-	OL	-	OL	CL
54	-	-	-	CL	CL	CL	-	-	-	CL	±	±	-	CL	CL	-	-	CL	SCL	-	-	-	-	CL	-	-	-	-	-	-	CL

Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35	Pool O
55	±	-	-	+++	SCL	+/+++	-	-	-	<CL	-/<SCL	-/<SCL	-/<SCL	+++	<SCL	-	-	<SCL	±	-	<SCL	+++	±	<CL	++	CL	-	CL	-	CL	CL
56	-	-	-	+++	SCL	-	-	-	±	<SCL	-	-	-	+++	-	++/SCL	-	<SCL	-	-	CL	-	-	-	+	-/+++	+	-	-	-/+++	CL
57	-	-	-	+	+++	SCL	-	-	-	±	-	-	+/+++	-	+++	+	-	+++	±	+++	±	±	+++	-	-	+++	-	OL	-	OL	CL
58	±	-	±	-	-	-	SCL	-	-	OL	CL	CL	OL	-	-/+++	-	CL	-	±	±	±	+++	+++	-	+++	OL	-	CL	-	OL	CL
59	-	-	-	-	-	-	CL	-	-	-	±	±	-	-	-	-	-/+++	-	+++	-	-	-	-	-	-	-	-	-	+++	-	<CL
60	-	-	-	-	-	-	-	CL	-	CL	+++	+++	-	-	-/+++	-	+++	-	-	-	+++	-	-	-	-	-	-	OL	-/+++	-	CL
61	-	-	-	-	-	-	-	CL	-	OL	+++	+++	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+/OL	-	-	CL
62	-	-	-	-	-	-	+	SCL	-	CL	-	±	CL	-	CL	±	CL	++	++	+++	++	-	++	CL	CL	CL	CL	CL	-	CL	CL
63	-	-	-	-	-	-	-	SCL	-	+++	-	-	-	SCL	-	OL	CL	SCL	SCL	-	+	-	-	-	-	+++	CL	-	CL	SCL	CL
64	-	-	-	-	-	-	-	CL	-	CL	-	-	±	-	±	-	-	-	-	-	-/+++	-/+++	-	-	-	-	-	CL	-	-	CL
65	-	-	-	-	-	-	-	SCL	-	+OL	-	-	-	-	-	+++	-/+++	<SCL	<SCL	-	OL	-	-	-	-	-	-	-	+++	-	CL
66	-	-	-	-	-	-	-	-	SCL	OL	-/+++	-/+	-	-	SCL	-	-	-	CL	-	CL	CL	-	±	±	-	-	CL	CL	-	CL
66a	-	-	-	-	-	-	-	-	SCL	OL	-	-	-	-	SCL	-	-	-	CL	-	CL	CL	-	±	±	-	-	CL	-	-	CL
67	-	-	-	-	-	-	-	-	CL	-	CL	CL	-	-	+++	-	-	-	CL	-	CL	CL	-	±	±	-	-	CL	CL	-	CL
68	-	-	-	-	-	-	-	-	CL	-	-	-	-	-	CL	-	-	-	CL	±	CL	CL	-	±	±	CL	-	CL	CL	CL	CL
69	-	-	-	-	++	-	-	-	OL	-	-	-	-/++	-	SCL	-	-	-	OL	-	CL	CL	-	-/+	-/+	-	-	CL	-	-	CL
70	-	-	-	-	-	-	-	-	-	SCL	CL	CL	CL	-	+++	+	CL	+++	-	-	-	-	-	-	++	SCL	-	CL	-	OL	CL
71	-	-	-	-	-	-	-	-	-	CL	CL	CL	CL	-	++	-	-	+++	-	+++	+	-	-	CL	CL	-	-	OL	-	OL	CL
72	-	-	-	-	-	-	-	-	-	OL	OL	OL	OL	-	-	-	-	-	-	-	±	-	-	-	+++	SCL	-	OL	-	OL	CL
73	-	-	-	-	-	-	-	-	-	+++	CL	CL	CL	-	-	-	-	-	-	-	-	-	-	++	CL	-	OL	-	SCL	CL
73a	-	-	-	-	-	-	-	CL	-	SCL	CL	CL	SCL	-	-	±	-	-	-	++	-	-	+	-	-	<SCL	-	<SCL	-	<SCL	CL
74	-	-	-	-	-	-	-	-	-	CL	CL	CL	-	-	-	CL	CL	CL	CL	±	CL	-	-	-	-	-	-	-	CL	-	CL
75	-	-	-	-	-	-	-	-	-	OL	OL	OL	-	-	-	CL	-	CL	CL	-	-	-	-	-	-	-	-	-	CL	-	CL
76	-	-	-	-	-	-	-	-	-	CL	++	++	-	±	-	-	-	CL	SCL	-	CL	-	-	-	-	-	-	-	+++	-	CL
77	-	-	-	-	-	-	-	-	-	OL	OL	OL	-	-	-	-	-	+	+	+++	-	-	-	-	-	-	-	+	+	-	CL
78	-	-	-	-	-	-	-	-	-	+++	+++	+++	-	-	-	-	-	-	+	-	+++	+++	-	-	-	-	-	SCL	-	-	CL
79	-	-	-	-	-	-	-	-	-	OL	OL	OL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
80	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	-	+++	CL	OL	SCL	-	SCL	-	-	-	-	-	-	-	SCL	-	CL
81	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	-	+++	-	OL	+++	-	SCL	-	-	-	-	-	-	-	SCL	-	CL
82	-	+	-	-	-	-	-	-	-	OL	-	+++	OL	SCL	-	-	-	-/++	+	+++	-	-	-	-	+++	OL	-	-/+	-/+	OL	CL
83	-	-	-	-	-	-	-	-	±	SCL	±	++	OL	-	+++	-	CL	-	-	+++	-	-	-	-	+++	OL	-	OL	-	OL	CL
84	-	±	-	-	-	-	-	-	-	OL	-	SCL	OL	-	-	+	-	-/+++	-	+++	-	-	+++	-	CL	CL	CL	+++	-	OL	CL
85	-	-	-	-	-	-	-	-	-	OL	++	+++	OL	-	-	-	-	-	-	-	-	-	-	-	SCL	SCL	OL	OL	-	OL	CL
86	-	-	-	-	-	-	-	-	-	OL	-	+++	OL	-	-	-	-	-	-	-	-	-	-	-	SCL	OL	-	OL	-	OL	CL
87	-	-	-	-	-	-	-	-	-	+++	-	CL	-	CL	-	CL	CL	CL	CL	-	-	-	-	-	-	±	±	-	OL	+	CL
88	-	+	-	-	-	-	-	-	-	SCL	-	-	OL	-	OL	+	-	+++	-	+++	±	-	+++	SCL	SCL	OL	-	OL	-	OL	CL
89	-	-	-	-	-	-	-	-	-	OL	-	-	OL	-	OL	-	OL	-	-	-	+	-	-	SCL	SCL	OL	-	OL	-	OL	CL
90	-	±	±	-	-	-	-	-	-	SCL	-	±	CL	-	+OL	±	-	-/+++	-	-/+++	±	-	-/++	SCL	CL	CL	-	<SCL	-	CL	CL
91	-	-	-	-	-	-	-	-	-	OL	-	-	-/+++	-	+	-	-	-	-	+++	+++	-	-	-	-	-	-	CL	-	-	CL
92	-	-	-	-	-	-	-	-	-	+++	-	-	SCL	-	++	-	-	-	-	-	-	-	-	-	++	SCL	±	OL	-	OL	+++
93	-	-	-	-	-	-	-	-	-	OL	-	-	OL	-	-	-	-	-	-	-	-	-	-	-	SCL	OL	-	OL	-	OL	CL
94	-	±	-	-	-	-	-	-	-	SCL	-	-	SCL	SCL	-	-	+	-	+	-	-	-	-	-	-/++	SCL	-	-	-	-/SCL	CL
95	-	-	-	-	-	-	-	-	-	SCL	-	-	-	-	-	SCL	CL	CL	SCL	-	-	-	-	-	-	-	-	-	+++	-	CL
96	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	-	OL	-	OL	OL	-	-	-	-	-	-	-	-	-	OL	-	CL
97	-	-	-	-	-	-	-	-	-	+++	-	-	-	-	-	+	SCL	+	-	+++	-/++	-	-	-	-	-	-	-	-	-	CL
98	-	-	-	-	-	-	-	-	-	OL	-	+	-	-	-	-	-	-	-	SCL	-	-	-	-	-	+	-	-	+	±	CL
99	-	-	-	-	-	-	-	-	-	+++	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
100	-	-	-	-	-	-	-	-	-	-	CL	CL	SCL	-	+	-	-	+++	+++	-	-	-	-	-	-	-	-	+	-	+	CL
101	-	-	-	-	-	-	-	-	-	±	CL	CL	-	-	+	+	CL	+++	++	+++	-	-	-	-	-	OL	-	+++	+	OL	CL
102	-	-	-	-	-	-	-	-	-	+	CL	CL	-	-	+++	±	-	+++	±	++	-	-	-	-	-	OL	-	-/+++	±	OL	CL
103	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	+++	-	-	-	-/++	-	CL	+	-	-	-	-	-	OL	-	-	CL
104	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
104b	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL	-	±	-	-	-	-	-	-	±	±	-	-	+	CL
104c	-	-	-	-	-	-	-	CL	-	-	CL	CL	-	-	-	-	CL	-	+	-	-	-	-	-	-	+	OL	-	+	+	CL
105	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	-/++	-	-	OL	-	-	-	-	-	-	+	-	-	-	-	CL
106	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	-	-	-	OL	-	-	-	-	-	-	-	-	-	OL	+	CL
107	-	-	-	-	-	-	-	-	-	-	SCL	SCL	-	-	-	++	-	-	-	-	+++	-	-	-	-	-	-	-	-	-	CL
108	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	-	-	-	±	+++	-	-	-	-	-	+++	-	-	-	+++	CL
109	-	-	-	-	-	-	-	-	-	-	OL	OL	-	-	-	-/+++	-	-	+	-	-	-	-	-	-	-	-	-	-	-	CL
110	-	-	-	-	-	-	-	-	-	-	-	CL	-	-	-	-	CL	-	-	-	-	-	-	-	-	±	-	-	-	+	CL
110a	-	-	-	-	-	-	-	CL	-	-	-	CL	-	-	-	-	<CL	-	-	-	-	-	-	-	-	±	-	<SCL	-	+	CL
110b	-	-	-	-	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	±	-	-	-	+	CL
111	-	-	-	-	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
112	-	±	±	-	-	-	±	-	-	-	-	-	+++	SCL	-	-	-	-	<SCL	+OL	+++	-	-	-/+++	±	OL	+	+	CL	OL	CL
113	±	±	-	-	-	-	-	-	-	±	-	-	OL	-	SCL	-	-/+++	-	-	-	SCL	+++	-/+	-/++	-	SCL	-	OL	-	-	CL
114	-	-	-	-	-	-	-	-	-	-	-	-	<SCL	-	-	±	-	-	-	-	-	-	-	-/++	-	-	-	<SCL	-	-	CL
115	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCL	-	-	-	+	-	<CL	±	-	-	-	-	-	CL	-/±	-	CL
116	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	OL	-	+	+	-	-	-	-	-	-	-	-	-	+	-	CL
117	-	+	-	-	-	-	-	-	-	-	-	-	-	-	-	SCL	-	-	-	+++	+++	-	-	-	-	-	-	+++	-	-	<CL

Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35	Pool O
118	-	-	-	-	-	-	-	-	-	±	-	-	-	-	-	+++	SCL	SCL	SCL	-	-	-	-	-	-	-	-	-	SCL	-	<SCL
119	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	+++	OL	±	-	-	-	-	-	-	+	+	-	-	+	+
120	-	-	-	-	-	-	-	-	±	-	-	-	-	-	-	-	CL	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
121	-	+	-	-	-	-	-	-	-	++	-	-	-	-	-	±	-	+++	++	<CL	-	-	-	-	-	±	-	-	<SCL	++	CL
122	-	-	-	-	-	-	-	-	±	-	-	-	-	-	±	-	-	-	++	-	+++	SCL	-	-	-	-	-	CL	-	-	CL
123	-	-	-	-	-	-	-	-	+	-	±	±	-	-	++	-	-	-	+	-	-	+++	-	-	-	OL	-	OL	-	OL	CL
124	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	OL	-	-	-	CL
125	<CL	SCL	<CL	CL	CL	CL	CL	CL	+++	CL	-	-	CL	CL	+++	<CL	<CL	CL	+++	CL	SCL	<CL	CL	CL	<CL	CL	+/CL	CL	SCL	CL	CL
126	<SCL	<SCL	<SCL	<SCL	SCL	-	<SCL	-	<SCL	SCL	<SCL	<SCL	CL	CL	<SCL	<SCL	SCL	SCL	<SCL	<SCL	SCL	CL	<SCL	<SCL	<SCL	CL	±	<SCL	OL	<SCL
126a	SCL	<OL	SCL	CL	CL	-	<SCL	<SCL	<CL	CL	-	-	CL	CL	<CL	<OL	<CL	<CL	CL	CL	CL	OL	CL	CL	<CL	SCL	OL	CL	CL	OL	CL
127	++	SCL	++	SCL	CL	OL	+++	-	CL	CL	CL	CL	<SCL	CL	CL	<SCL	-	CL	CL	CL	+++	CL	<SCL	<SCL	++/SCL	CL	-	OL	++	SCL	CL
128	SCL	-	<CL	-	-	OL	OL	-	SCL	CL	-	-	CL	-	SCL	SCL	SCL	-	<CL	-	+++	-	CL	±	++	CL	+	CL	++	OL	-
129	<SCL	SCL	SCL	-/±	-/±	-/++	<OL	-	-	<CL	-	+/CL	<OL	+/CL	±/+++	+/SCL	SCL	++/OL	+/SCL	<OL	-	-	<OL	-	<SCL	<CL	-	<SCL	-	<OL	CL
129a	<SCL	SCL	SCL	±	±	++	<OL	CL	-	<CL	-	+/CL	<OL	+/CL	±/+++	SCL	SCL	<OL	<SCL	<OL	-	-	<OL	-	<SCL	<CL	OL	<SCL	-	<CL	CL
130	+++	OL	+++	-	-	-	SCL	-	-	-/++	+++	+++	OL	<SCL	-	SCL	-/+	-/±	-	+++	+++	-	+++	-	+++	-	-	-/+++	-	-/+++	CL
131	+++	<SCL	+++	-/++	-/++	-/±	SCL	-	-	++	-	-	SCL	SCL	+++	<SCL	+++	+++	-	<SCL	<SCL	-	+++	++	+++	++	-/±	SCL	-	+++	CL
132	-	CL	CL	SCL	SCL	<SCL	-	-	±	CL	+/CL	++/CL	CL	CL	CL	CL	-	CL	+	CL	CL	++	CL	++	CL	<OL	-	CL	+	CL	CL
133	-	SCL	<SCL	<SCL	SCL	OL	<OL	-	±	OL	<CL	CL	OL	OL	<OL	±	-	<OL	SCL	<OL	SCL	<OL	SCL	++	+	<OL	±	OL	<SCL	OL	CL
134	+	++	+	-	-	-	+++	-	-	-	-/++	CL	-	SCL	SCL	+	+++	-/±	-	+++	++	-	+++	-	<CL	-/++	-	SCL	-	<CL	CL
135	-	+/+++	±/++	<SCL	<SCL	<CL	-	-	++/SC	++/SCL	-/CL	-/CL	SCL	+++	CL	CL	-	+++	+/SCL	+/SCL	+/++	+/CL	+/+++	+/+++	±/++	CL	±	CL	+/+++	<CL	CL
135a	-	5	±	++	+	++	-	-	++	±	-	-	++	-	+++	++	-	+	+	±	3	++	2	+++	2	4	-	CL	5	++	
136	-	-	-	OL	OL	OL	-	-	-	OL	+/<OL	+/OL	-	OL	OL	-	-	CL	+	-	-	-	-	<CL	-	-	-	-	-	-	CL
137	-	-/++	-	CL	CL	CL	-	-	-	CL	-/CL	-/CL	-	CL	CL	++/SC	CL	CL	+	CL	SCL	-	CL	±	-	-	-	-/+++	-	++/SC	CL
138	-	-	-	SCL	<OL	OL	-	-	-	SCL	±	SCL	-	SCL	OL	SCL	CL	SCL	-	<SCL	<CL	-	-	<SCL	-	-	-	-	-	<SCL	CL
139	-	-	-	OL	CL	CL	-	-	±	OL	-/++	-/SCL	-	OL	OL	+	-	CL	++	±	-	++	-	CL	-	CL	-	OL	SCL	+OL	CL
140	-	-	-	<SCL	<SC	<SCL	-	-	-	<SC	-	-	SCL	<CL	<OL	<SCL	-	<SCL	+	-	+	-	-	-	±	CL	-	OL	<SC	OL	CL
141	-	-	-	+++	+++	-	-	-	<SCL	±	±	+	-	+++	<SCL	-	-	+++	<SCL	-	<SCL	<SCL	-	++	+	CL	+	CL	<SCL	CL	<SCL
141a	-	-	-	+++	+++	-	-	-	<SCL	±	±	+	-	+++	<SCL	-	-	+++	<SCL	-	<SCL	<SCL	-	++	+	CL	OL	CL	<SCL	CL	<SCL
142	-	-	±	++	+++	SCL	-	-	+	+++	-	-	SCL	SCL	SCL	-	-	+++	+	-	+++	+++	++	+++	++	-	-	CL	+	-	CL
143	-	±/++	-/±	++	++	+++	-	-	++	±	-/+<SCL	-/+<SCL	<OL	+	+++	-/±	-	++	++	++	++	+/+++	+/+++	+	++	++	-/+	SCL	-/±	+++	CL
144	-	-	++	-	-	OL	-	-	-	-	-	-	CL	-	OL	-	-	-	-	-	-	-	SCL	-	++	-	-	SCL	-	±	CL
145	-	-	-	+	<OL	-	-	-	-	-	-	-	-	-	<OL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
146	-	±/SC	-	-	-	-	++	-	-	<SCL	-	-	OL	OL	-	-	-	++	+	++	-	-	±	++	CL	+	±	±	OL	CL
146a	+++	++	++	±	±	±	<OL	-	+	OL	-	-	OL	OL	<OL	-	-	±	+	±	+	+	++	-	+	OL	±	SCL	±	OL	CL
147	±	-	±	-	-	-	+++	-	-	±/SCL	CL	CL	+/SCL	-	-	-	-	-	+	<SCL	++	-	++	-	-	-	+	-	-	-	CL
148	-	-	-	-	-	-	-	<SCL	-	+++	-	-	OL	-	-	+	+++	-	+	+++	-	-	-	-	<SCL	OL	OL	OL	±	OL	CL
149	-	-	-	-	-	-	-	<CL	-	-	<CL	CL	-	-	-	-	CL	-	-	±	-	-	-	-	-	SCL	-	-	-	OL	CL
150	-	-	-	-	-	-	-	CL	-	-	+++	SCL	-	-	-	-	-	-	++	-	-	-	-	-	-	-	-	-	+	-	CL
151	-	-	-	-	-	-	-	+++	-	+	++	-	-	-	-	-	<SCL	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
152	-	-	-	-	-	-	-	<SCL	-	-	-	-	-	-	-	++	++	-	-	-	<SCL	-	-	-	-	-	-	-	-	-	CL
153	-	-	-	-	-	-	-	-	<SCL	±	-	-	-	-	SCL	-	-	-	<SCL	-	-	-	<CL	±	-	-	-	OL	-	-	CL
154	±	±	+SC	-	-	-	+/SCL	-	-	<SCL	CL	OL	-	-	<SCL	-	-	-	-	+/SCL	+/SCL	-	-	+/+++	SCL	<SCL	<SCL	OL	-	OL	CL
155	-	-	-	-	-	-	++	-	-	<SCL	CL	CL	OL	-	++	-	CL	++	-	++	-	-	-	-	<SCL	OL	-	<OL	-	OL	CL
156	-	-	-	-	-	-	-	-	-	SCL	CL	CL	OL	-	OL	-	CL	-	-	++	++	-	-	OL	+OL	OL	±	OL	-	OL	CL
157	-	-	-	-	-	-	-	-	-	<SCL	<SCL	<SCL	-	<SCL	-	-	-	<SCL	±	-	-	-	-	-	-	-	-	-	-	±	<SCL
158	-	-	-	-	-	-	-	-	+	SCL	++	SCL	-	-	±	-	-	-	+	-	-	+++	-	-	-	CL	-	-	CL	-	CL
159	-	<OL	-	-	-	-	-	-	-	OL	-	<SCL	-	-	-	<SCL	-	<OL	<SCL	OL	-	-	+	-	-	<SCL	-	-	-	<SCL	CL
160	-	-	-	-	-	-	-	-	-	OL	-	<OL	-	-	-	<SCL	SCL	OL	SCL	-	SCL	-	-	-	-	-	-	-	<SCL	-	CL
161	-	-	-	-	-	-	-	-	-	<SCL	-	-	<SCL	-	-	-	<SCL	-	-	-	-	-	-	-	++	<SCL	-	-	±	<SCL	CL
162	-	CL	-	-	-	-	-	-	-	SCL	-	-	-	±	-	CL	-	-	-	CL	CL	-	-	-	-	-	-	OL	-	-	CL

163	-	OL	-	-	-	-	-	±	-	OL	-	-	OL	+	-	+/SCL	CL	OL	-	OL	OL	-	-	-	-	-	-	OL	-	+	CL
164	-	-	-	-	-	-	-	-	-	SCL	-	-	±	-	-	-	-	-	+SCL	±	-	-	-	-	-	-	-	+	±	±	CL
165	-	-	-	-	-	-	-	-	-	+CL	+	+	-	+	-	-	-	-	+	-	+SCL	-	-	-	-	-	-	-	-	-	CL
166	-	-	-	-	-	-	-	-	-	SCL	-	-	OL	-	-	-	-	-	+	-	-	-	-	-	+/SCL	-	-	-	-	-	CL
167	-	-/+++	-	-	-	-	-	-	-	+	++/SCL	++/SCL	<OL	-	++/SCL	+	-	+SCL	-	++/SC	-	-	-	±/++	-	CL	-	-/SC	±	OL	CL
168	-	+	-	-	-	-	-	-	-	-	CL	CL	+++	-	±	-	CL	++	-	++	-	-	-	-	-	OL	-	+++	-	OL	CL
168a	-	+	-	-	-	-	-	-	-	±	CL	CL	+++	-	-	±	-	+	±	+	-	-	-	-	-	CL	-	+++	-	OL	CL
169	-	-	-	-	-	-	-	-	-	-	SCL	CL	OL	-	-	-	-	-	-	-	-	-	-	-	+SCL	-	-	-/SC	-	+SCL	CL
170	-	+	-	-	-	-	-	-	-	-	CL	CL	+++	-	-	±	-	-	-	±	-	-	-	-	-	++/SC	-	-	-	+OL	CL
170a	-	+	-	-	-	-	-	-	-	-	-	-	+++	-	-	±	-	-	-	±	-	-	-	-	-	++/SC	-	-	-	+OL	CL
171	-	-	-	-	-	-	-	-	-	-	CL	CL	++	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
172	-	-	-	-	-	-	-	-	-	-	<SCL	<SCL	-	-	-	-	-	-	SCL	-	-	-	-	-	-	-	-	<SCL	<SCL	+/+SC	CL
173	-	-	-	-	-	-	-	-	-	±	+	SCL	+	-	±	-	-	-	-	-	-	-	-	-	+	-	OL	-	+	CL
174	-	++	-	-	++	-	+	-	-	++	-	-	SCL	<OL	+++	-	-	++	-	+	++	-	-	-	+	-	+++	-	-	SCL	CL
175	-	-	-	-	-	-	-	-	-	-	-	-	<SCL	-	<SCL	-	+	-	-	-	-	-	-	-	-	+	-	OL	-	+	CL
176	-	+	-	-	-	++	-	-	-	±	-	-	OL	-	++	-	-/++	-	-	++	+	-	+	<SCL	SCL	<SCL	-	OL	-	+OL	CL
177	-	-	-	-	-	-	-	-	-	+	-	-	<SCL	-	-	-	CL	-	-	-	-	-	-	-	-	<SCL	-	+/SC	-	+SCL	-
178	-	-	-	-	-	-	-	-	-	-	±	-/++	OL	-	-	-	-	-	-	-	+	-	++	-	-	-	-	-	±	-	CL
179	-	-	-	-	-	-	-	±	-	-	±	±	-	-	-	-	CL	-	-	+	-	-	-	-	-	SCL	-	-	-	OL	CL
180	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCL	-	±	-	+++	-	-	-	-	-	-	-	-	+	CL
181	-	-	-	-	-	-	-	-	-	-	+/CL	+/CL	-	-	-	-	-	-	SCL	-	-	-	-	-	-	-	-	OL	SCL	-	OL

Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35	Pool O
182	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	±	-	-	CL	-	±	±	-	-	±	-	-/+++	CL
183*	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	+++	-	<OL	-	-	-	SCL	-	OL	CL
184	-	-	-	-	-	-	-	-	-	-	-	-	++	-	+	-	-	-	-	-	++	-	-	+	-	OL	<OL	OL	-	OL	<SCL
185	-	-	-	-	-	-	-	-	-	-	±	±	-	-	-	-	-	-	-	-	-	-	-	-	-	<SCL	-	-	-	+SCL	CL
186	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-/++	-	-	-	-	-	<SCL	-	-	-	OL	CL
187	-	-	-	-	-	-	-	-	-	-	-	-	-	-	+	-	-	-	-	-	-	-	-	-	-	-	-	SCL	<CL	-	CL
188	-	-	-	-	-	-	-	-	-	+++	-	-	-	-	-	-	-	+++	-	-	-	-	-	-	-	-	-	-	-	-	-
189	-	CL	<SC	-	-	-	-	-	CL	SCL	CL	CL	CL	-	CL	+SCL	-	±	CL	CL	-	CL	CL	±	SCL	CL	+	CL	CL	OL	±
190	-	OL	-	-	-	-	-	-	-	OL	+	++/+SCL	-	-	-	OL	-	OL	+	OL	+	-	-	-	-	±	-	-	-	±	
191	SCL	CL	-	SC	CL	+++	CL	-	+++	CL	-/++	-/+++	CL	CL	SCL	CL	SCL	CL	+++	CL	CL	SCL	-	CL	SCL	CL	+	OL	++	OL	SCL
191a	<CL	++	-	OL	CL	+++	CL	-	+++	+++	-	-	CL	CL	CL	CL	-	++	±	OL	OL	++	-	+	<SCL	++	2 s	+++	++	OL	-
192	-	-	-	-	-	-	-	-	-	-	-	+	OL	-	-	-	-	-	-	-	-	-	-	-	-/±	-	-	-	-	-	CL
193	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
194	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
195	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
196	-	-	-	OL	OL	OL	-	-	SCL	OL	OL	OL	-	OL	OL	CL	SCL	OL	SCL	-	+OL	CL	-	CL	-	++	+	SCL	SCL	+OL	CL
197	SCL	±	SCL	CL	OL	CL	CL	-	CL	CL	+	+	CL	CL	+CL	+CL	-	CL	CL	CL	+	CL	CL	CL	++/SCL	CL	-	CL	CL	SCL	CL
198	SCL	-	SCL	-	-	-	SCL	-	-	SCL	-	-	-	-	±	+	-	-	-	-	-	-	CL	-	-	CL	-	CL	-	OL	CL
199	-	±	-	-	-	-	+++	±	-	SCL	-	+	OL	-/+++	SCL	±	CL	<SCL	-	++	+SCL	-	++	SCL	<SCL	CL	+	OL	-	OL	CL
200	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
201	-	-	-	-	-	-	CL	-	-	CL	CL	-	-	-	-	-	-	-	+	±	-	-	-	-	-	++	-	-	±	OL	CL
202	-	-	+	+CL	<SCL	-	-	-	±	-	-	-	SCL	+	SCL	-	-	++	+	CL	CL	±	-	-	-	CL	-	<SCL	-	±	CL
203	-	-	-	-	-	-	-	-	+	-	-	-	-	<OL	-	±	-	-	±	-	-	-	-	-	-	±	-	-	-	+	CL
204	-	±	-	-	-	-	-	-	±	-	-	-	<SCL	-	++	±	-	-	-	-	-	-	-	-	-	++	-	<SCL	-	SCL	CL
204a	-	-	-	-	-	-	-	-	-	-	-	-	+++	-	-	-	-	-	-	-	-	-	-	-	-	±	-	-	-	++/SCL	CL
204b	-	±	-	-	-	-	-	-	±	-	-	-	<SCL	-	±	±	-	-	-	-	±	2s	-	-	±	±	-	CL	-	+	CL
204c	-	+	±	-	-	-	-	-	-	+	-	-	OL	-	+	-	-	-	±	±	-	±	-/++	-	±	SCL	+/SCL	+/+++	±	OL	CL
204d	-	-	-	-	-	-	-	-	-	±	-	-	OL	-	-	-	-	-	-	-	-	-	-	-	2s	++	-	CL	-	++/SCL	SCL
204e	-	±	±	-	-	-	-	-	-	-	-	-	-	+	+	-	-	-	±	±	±	±	±	-	-	+	-	CL	-	++	CL
205	+	++	±	++	+++	+++	+++	±	-	<CL	-	OL	OL	SCL	OL	+	<CL	+++	-	+++	SCL	++	<SCL	CL	<CL	OL	±	OL	-	OL	CL
206	+/++	+++	+/++	++	+++	<OL	+/++	-	±	+++	-	-	+++	<SCL	<CL	++	±	+++	++	+++	+++	+++	+++	+/++	±	+++	±	CL	±	SCL	CL
207	-	-	-	-	-	-	-	-	-	-	-	±	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
208	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
209	-	-	-	-	-	-	-	-	-	-	±	±	+++	-	-	-	-	-	+	-	+	-	-	-	-	-	-	-	+	-	CL
U14	-	-	-	-	-	-	-	-	-	+++	±	+	OL	-	-	-	-	-	-	-/++	-	-	++	-	+++	-	-	-/+++	-	±	++
U186	-	-	-	-	-	-	-	-	±	-	-	-	-/++	-	++	-	-	-	±	±	-	±	-	-	-	<SCL	-	<CL	-	+OL	CL
U274	-	-	-	-	-	-	-	-	-	SCL	-	-	-	-	-	-	-	SCL	+/SCL	-	-	-	-	-	±	±	-	-	+/SCL	±	CL
U275	-	-	-	-	-	-	-	-	-	SCL	-	-	-/+++	±	-	-	-	-	-	±	SCL	-	-	-	-	-	±	SCL	-	CL	
U276	-	-	-	-	-	-	-	-	-	++/SCL	-	-	OL	-	-	-	-	-	+	+	-	-	-	-	++	<OL	-	-	-	OL	CL
U277	±	++	+	-	-	-	+++	±	-	SCL	-	SCL	OL	<OL	-	±	SCL	SCL	±	SCL	-	-	++	-	<CL	OL	±	+	-	OL	CL
U278	±	OL	±	-	-	-	SCL	-	-	OL	-/SCL	-/SCL	OL	OL	-	-	-	OL	+/SCL	OL	±	-	+/SCL	-	+/SCL	SCL	±	+	+/SCL	OL	CL
U279	±	<OL	+	+/SCL	+/SCL	SCL	+	-	+	OL	CL	CL	OL	OL	OL	-	-	OL	OL	OL	+OL	+/SCL	+OL	+/SCL	±	OL	±	OL	SCL	OL	CL
U280	-	±	-	-	-	-	-	-	±	OL	CL	CL	-	-	-	++	-	++/OL	±	±	+++	-	-	-	-	-	±	+	-	CL	
U281	±	<CL	+/SCL	-	-	-	++	-	++	<SCL	±	-/++	SCL	+	±/SCL	±	-	+	<SCL	<SCL	+SCL	+/SCL	+/SCL	±	-	<SCL	±	SCL	+SCL	+OL	CL
U282	-	-	-	-	-	-	-	-	-	OL	CL	CL	-	-	SCL	-	CL	++	-	-	+	++	-	OL	-	±	±	+++	-	+	CL
U284	-	OL	-	-	-	-	-	-	-	<OL	-	-	-	2n	-	2n	-	<OL	OL	OL	-	-	-	-	-	+	-	-	<OL	SCL	CL
U285	-	-	-	-	-	-	-	-	-	SCL	-	+/SCL	+++	-	-	-	-	-	+/SCL	-	-	-	-	-	-	+	2s	-	+/SCL	SCL	CL
U286	±	+/SCL	±	++	++	-	±	-	-	+/SCL	-	-	SCL	OL	+	±/SCL	-	+/SCL	+	-	±	±	+	++	±	<CL	-	<CL	-	OL	CL
U287	SCL	CL	CL	OL	<OL	OL	SCL	-	CL	CL	++	+++	OL	OL	OL	CL	±	<OL	<OL	OL	<OL	CL	CL	SCL	SCL	CL	±	CL	CL	CL	CL
U287v	+++	<CL	SCL	OL	OL	++	+++	-	+++	CL	SCL	SCL	CL	CL	CL	CL	±	CL	++	CL	CL	+++	<OL	SCL	SCL	CL	-	CL	OL	OL	CL
U288	-	-	-	-	-	-	-	-	-	±	-	-	-	<SCL	-	±	±	±	+/SCL	-	-	-	-	-	-	±	±	-	+/SCL	+/SCL	CL
U288v	-	-	-	-	-	-	-	±	-	±	±	±	-	SCL	-	++	CL	±	+/<SCL	-	-	-	-	-	-	±	±	-	++	+/SCL	CL
U289	-	-	-	-	-	-	-	-	-	OL	+/SCL	<SCL	-	OL	-	+	-	OL	+	-	±	-	-	-	-	±	±	-	OL	±	CL
U290	-	±	-	±	±	±	-	-	-	±	-	-	<SCL	<SCL	+++	±	-	±	±	±	-	±	±	±	-	SCL	-	CL	-	OL	CL
U291	-	-	-	CL	CL	OL	-	-	-	CL	-	-	CL	CL	OL	-	SCL	CL	+	-	-	-	-	<CL	<CL	CL	-	-	-	-/OL	CL
U292	-	-	-	-	-	-	-	-	-	SCL	-	-	OL	-	-	±	-	-	-	±	-	-	-	-	±	+/SCL	-	±	-	SCL	CL
U293	±	±	±	++	+	+++	++	±	+++	+	±	++	+	±	SCL	<OL	<OL	+	+	±	++	++	-	-	±	±	-	<CL	++	+++	CL
U294	-	-	-	-	-	-	-	-	-	OL	-	±	-	±	±	+	±	OL	OL	-	+/+++	-	-	-	±	-	-	-	<OL	-	CL
U295	-	-	-	-	-	-	-	±	+++	±	-	-	+++	-	+/SCL	-	<CL	-	-	-	-	SCL	-	++	+	±	-	CL	-	+/+++	CL
U296	±	OL	+/SCL	OL	SCL	<SCL	++	-	+	OL	CL	CL	OL	OL	OL	±	±	OL	OL	OL	OL	SCL	OL	±	±	OL	±	OL	<OL	OL	CL
U297	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	-	SCL	-	SCL	+/SCL	-	-	-	-	-	-	-	-	-	-	-	CL
U298	SCL	SCL	SCL	CL	CL	CL	SCL	+	SCL	OL	+/+++	+++	CL	CL	OL	CL	CL	CL	SCL	CL	SCL	<SCL	CL	<CL	SCL	CL	OL	OL	SCL	OL	CL
U300	-	-	-	-	-	-	-	-	-	-	CL	CL	-	-	-	OL	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
U301	±	±	+	-	-	-	SCL	-	-	<OL	-	SCL	OL	OL	-	-	-	±	±	±	-	-	±	-	SCL	<OL	±	±	++	OL	CL
U302	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
U303	-	-	-	-	-	-	-	-	-	±	-	-	-	-/+++	+/+++	-	-	-	-	-	-	-	-	-	OL	-	-	-	-	OL	CL
U304	-	-	-	-	-	-	-	-	-	++	+++	<CL	-	-	-	-	-	-	+++	-	-	-	-	-	-	-	-	SCL	-	CL	
U305	CL	-	CL	-	-	-	CL	-	-	-	SCL	CL	CL	-	-	-	-	-	-	-	-	CL	-	CL	-	CL	-	-	-	CL	
U306	±	±	±	-	-	-	++	±	-	SCL	-	-	OL	++	-	-	CL	SCL	-	++	-	-	++	-	SCL	1n	-	2n	-	3n	CL
U307	-	+++	+++	-	-	-	-	-	SCL	SCL	1n	1n	-	-	SCL	CL	-	+++	SCL	+++	++	CL	+++	±	±	-	-	CL	SCL	-	CL
U308	-	-	-	-	-	-	-	-	-	-	++/SCL	++/SCL	-	<OL	-	±	±	+	+/OL	-	-	-	-	-	-	+/<OL	2n	-	SCL	SCL	CL

Phage Type	1	2	3	4	5	6	7	8	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	32	35	Pool O
U308a	-	-	-	-	-	-	-	-	-	-	-	-	-	+++	-	±	±	+	+/OL	-	-	-	-	-	-	+/<OL	2n	-	SCL	SCL	CL
U309	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
U310	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
U311	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	CL
U312	-	-	-	-	-	-	-	2s	-	±	-	-	+++	++	-	-	<CL	-	-	-	-	-	-	-	±	±	-	-	-	+	-
U313	<CL	<CL	CL	OL	CL	CL	CL	-	-	CL	-	-	CL	CL	CL	CL	CL	CL	+	CL	CL	-	CL	<CL	CL	-	-	CL	-	-	CL
U314	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	SCL	-	-	-	OL	CL
U315	-	-	-	-	-	-	-	-	-	+OL	-	SCL	OL	OL	-	-	-	++	±	-	-	-	-	-	SCL	SCL	SCL	-	-	+OL	CL
U316	-	-	-	SCL	SCL	+/+++	-	-	-	SCL	-	-	1s	SCL	CL	-	-	SCL	-	-	SCL	-	-	++	-	SCL	+	-	-	CL	CL
U317	+	-	+	-	-	-	SCL	-	-	-	-	-	OL	-	-	-	-	-	-	-	SCL	±	CL	-	OL	-	-	-	-	-	
U318	-	-	SCL	CL	CL	CL	-	-	CL	CL	+	+	CL	CL	CL	CL	-	CL	CL	CL	CL	CL	CL	<CL	CL	CL	-	CL	CL	OL	CL
U319	-	-	-	-	-	-	-	±	-	-	-	-	-	++	-	+/SCL	SCL	-	-	-	-	-	-	-	-	-	±	-	-/±	-	CL
U320	++	+	+	OL	SCL	CL	+++	-	SCL	SCL	-/+	-/++	+++	+CL	CL	<OL	++	<SCL	<SCL	<CL	SCL	SCL	+++	+	+	<CL	-	CL	CL	OL	CL
U321	3s	-	±	<SCL	+	++	+	-	±	<OL	-	-	++	++	SCL	±	2s	±	4ns	+	±	±	±	SCL	±	1s	-	CL	-	±	CL
U322	-	-	-	-	-	-	-	-	-	-/±	-	-	-	+++	-	-	-	-	±	-	-	-	-	-	-	+	-	-	±	CL	CL
U323	-	-	-	SCL	++	CL	-	-	-	±	-	-	-	-/±	+++	++	+++	-/+	-	++	-/±	-	-	SCL	-	-/±	-	-	-	+	CL
U324	-	-	-	-	-	-	-	-	-	± sm	-	-	-	±	OL	-	-	-	-	-	-	-	-	OL	-	-/±	-	-	-	-	CL
U325	-	-	-	-	-	-	-	-	-	-	-	-	-	-	++ sm	-	-	-	-	-	-	-	-	-	-	-	-	-	± sm	-	CL
U326	-	-	-	-	-	-	-	<CL	-	± ns	-	-	-	-	-	-	+++sm	-	<SCL	+ ns	-	-	-	-	-	-	-	-	CL	-	CL
U327	-	-/±	-	-	-	-	-	-	-	SCL	-	-	++ sm	SCL	++ sm	± sm	-	-	-	± sm	-	-	-	± sm	± sm	SCL	-	-	± sm	SCL	CL

Additional Phages							
Phage type	1	2	3	10	10 var 2	10 var 3	18

193	SCL	SCL	SCL	-/OL	-/OL	-/OL	-
193a	SCL	SCL	SCL	-/+	-/+	-/OL	+/++
194	-	SCL	-	-/OL	-/OL	-/OL	-
195	-	±	SCL	-/OL	-/OL	-/OL	-
208	-/++	-/++	-/++	-/OL	-/OL	-/OL	OL
U302*	-	-	-	OL	OL	OL	-
U310*	-	-	-	-/±	OL	±	-
U311*	-	-	-	-	-	<OL	-

* = Provisional phage type.

Appendix II: Primers Used in this Study

Description	Sequence	Location	size	Tm (^{0}C)	genotype
Forward primer for replacing *bar* with *cat*	tgctgtcgcgaattc cggtagtttgtaaaa caccattactggagc aattcgGTGTA GGCTGGAGC TGCTTCG	*bar*	537bp	77	S04698-09 Δ *bar::cat*
Reverse primer for replacing bar with cat	aagcgattgaaataa acataaaggggca gacgccccttattc ctgccCATATG AATATCCTC CTTAGT	*bar*	537bp	74	S04698-09 Δ *bar::cat*
Forward primer for replacing *copA* with *aphII*	cgattactaaggtca atgaattgatgacca atcataaaggagtttt tactGTGTAGG CTGGAGCTG CTTCG	*copA*	2502bp	71	SL1344 Δ *copA ::aph*
Reverse primer for replacing *copA* with *aphII*	tgatggcttaataata tatcaggcagctgat gctgcctgatatagc gtttCATATGA ATATCCTCC TTAGT	*copA*	2502bp	70	SL1344 Δ *copA ::aph*
Forward primer for polymerising right hand of SGI-4	aggctgcgcagggt attat	*traI*	2Kb in 2.7Kb out	59	NA
Reverse primer for polymerising right hand of SGI-4	gtccctcaagtaagg gaac	Phe-tRNA	2.5Kb in	54	NA
Reverse primer for polymerising right hand of SGI-4 and insert site	aatgatcggatcttct gatgga	*xerD*	2Kb out	55	NA
Forward primer for insertion of *aphII* upstream of *iciA*	aggattagcctctcta tgaatgccggatggc ggcgtgaacgtctta tccgCTATCAT TGGgtgtaggctg gagctgcttcg	upstream of *iciA*	1bp (1.2 Kb insert)	79	4/74::*aphII*

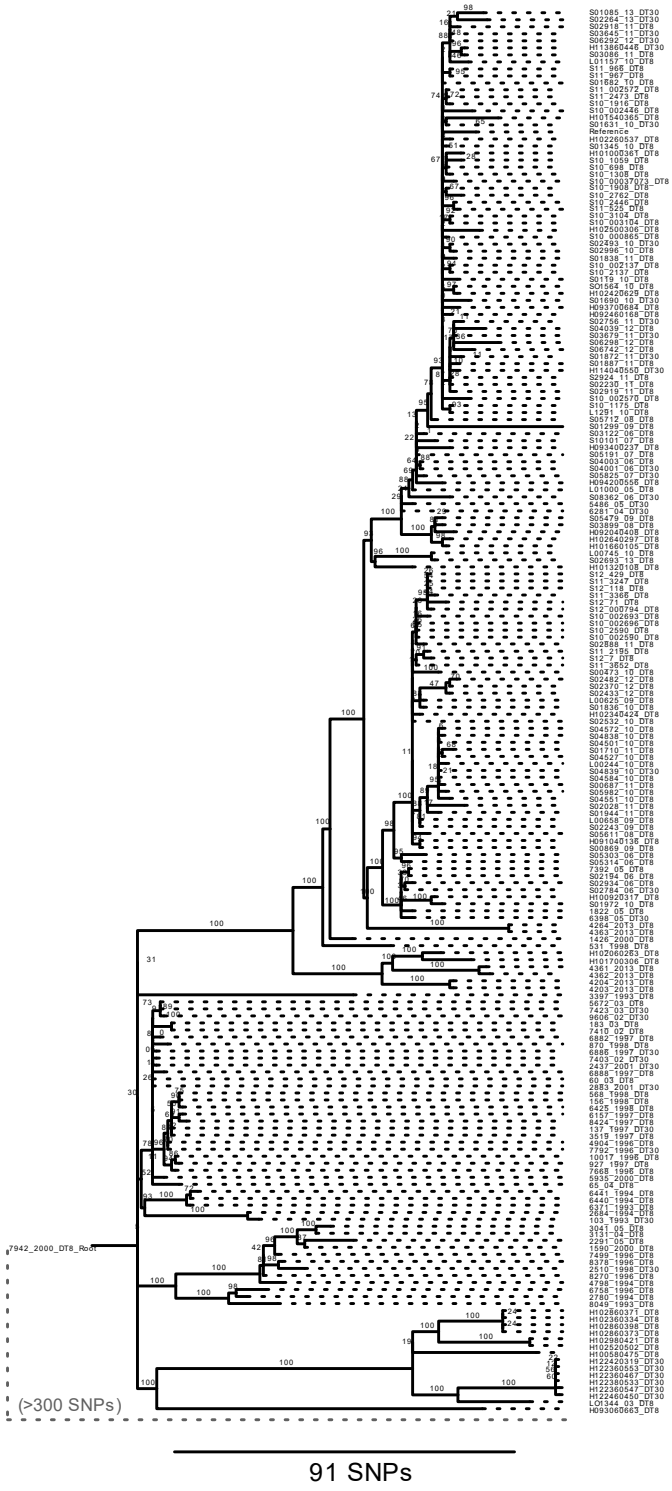
Reverse primer for insertion of *aphII* upstream of *iciA*	tttgtgatgaacttcaagaagtgacgaggaataatgggtcatgcaggccatatgaatatcctccttagt	upstream of iciA	1bp (1.2 Kb insert)	82	4/74::*aphII*
Forard primer for insertion of *cat* in mTmII	caagctggtaccgccaggactacacacagcataaagttgtggtgccgggtGTGTAGGCTGGAGCTGCTTCG	Right hand end of mTmII	0bp (1Kb insert)	80	S04698-09 mTmII::*cat*
Reverse primer for insertion of *cat* in mTmII	tgtggatacccacccgcctggtgtgcaaccttcgccaggcaccgggaggcCATATGAATATCCTCCTTAGT	Right hand end of mTmII	0bp (1Kb insert)	81	S04698-09 mTmII::*cat*
Forward primer for polymerising *wzy-wjx* locus	tgcgactatcaggttaccgt	*nucA*	3.2kb(4.2kb with *cat* insert)	59	
Reverse primer for polymerising *wzy-wjx* locus	gttagcgtgcggtcaagatc	*thrS*	3.2kb(4.2kb with *cat* insert)	59	
Forward primer for knockout of *wzy-wjx locus*	atgaatttattgacagtgagttgtgtatccatctagccaaccattgctaaGTGTAGGCTGGAGCTGCTTCG	*wzy-wjx*	2.7kb	73	L01157-10Δ *wzy-wjx*
Reverse primer for knockout of *wzy-wjx locus*	gaaagataacggagaaaccctgtcaagggtcttgatttgctatagagtgaCATATGAATATCCTCCTTAGT	*wzy-wjx*	2.7kb	71	L01157-10Δ *wzy-wjx*
Forward primer for polymerising absence of *wzy-wjx* locus	acggagaaaccctgtcaagg	intergenic	214bp	59	NA
Reverse primer for polymerising absence of *wzy-wjx* locus	agtggtgacccatacgcaag	intergenic	214bp	59	NA
Forward primer for polymerising middle of *wzy*	gcctgaagattttggcgcat	*wzy*	207bp	61	NA
Reverse primer for polymerising middle of *wzy*	tgcgctgactttgtttcctg	*wzy*	207bp	61	NA

Forward primer for polymerising housekeeping gene *rpoD*	gtcaacagtatgcgcgtgat	*rpoD*	154bp	64	NA
Reverse primer for polymerising housekeeping gene *rpoD*	gatagcggcattgaaccagg	*rpoD*	154bp	67	NA
Forward primer for polymerising middle of *cat*	acaaacggcatgatgaacct	*cat*	433bp	57	NA
Reverse primer for polymerising middle of *cat*	gcacaagttttatccggcct	*cat*	433bp	58	NA
Forward primer for insertion of *aphII* upstream of *iciA*	aggattagcctctctatgaatgccggatggcggcgtgaacgtcttatccggagtaagagcGTGTAGGCTGGAGCTGCTTCG	intergenic (*iciA* downstream)	1bp (knock-in) insert = 1.5kb	80	L01157-10::*aphII*
Reverse primer for insertion of *aphII* upstream of *iciA*	tttgtgatgaacttcaaagaagtgacgagggaataatgggtcatgcaggCATATGAATATCCTCCTTAGT	intergenic (*iciA* upstream)	1bp (knock-in) insert = 1.5kb	70	L01157-10::*aphII*
Forward primer for insertion of *cat* into *wjx*	tgccactgtatataaacacagtaacaatcatgtgtaattatcaaaagcatGTGTAGGCTGGAGCTGCTTCG	*wjx*	1bp (knock-in) insert = 1kb	71	L01157-10Δ*wjx*::*cat*
Reverse primer for insertion of *cat* into *wjx*	atgatgaatcttcgaatgccactgtatataaacacagtaacaatcatgtgGTGTAGGCTGGAGCTGCTTCG	*wjx*	1bp (knock-in) insert = 1kb	73	L01157-10Δ*wjx*::*cat*

N.B: Nucleotide coding sequences are in 5' to 3' direction

Appendix III: DT8 Epidemic Phylogenetic Tree Showing Bootstrap Values

Appendix III. Phylogenetic reconstruction of 196 DT8 and DT30 phage typed *S.* Typhimurium. Maximum



likelihood phylogenetic tree constructed from 2,297 core genome SNP sites using 162 DT8 strains, 34

DT30 strains, and rooted on a non-epidemic-clade DT8 strain. Bootstrap values are displayed for each bifurcating node.

Appendix IV: Likelihood Ratio Test and R^2 Calculation-Containing Scripts for Ancestral State Reconstruction

```
#this script is for ancestral reconstruction using phytools & simmap

library(phytools)

#set correct working directory

setwd("\\\\nbi-cfs2\\ifrprojects\\Rob-
Kingsley\\Oliver\\Genome_variation_S.typhimurium_phage_heavy_metals\\DT8_DT30_swit
ching_analysis\\ancestral_state_reconstruction")

#read finalised dt8 tree

dt8 <- read.tree("DT8_tree_drop_tips.newick")

#plot tree

plotTree(dt8, ftype="i", fsize = 0.4, type = "fan")

#read phage type data

pt <- read.csv("Strain_phagetype2.csv", row.names=1)

#make into  matrix

ptd <- as.matrix(pt)[,1]

#have a look at the tree

plotTree(dt8, fsize = 0.4, type = "fan")

cols<-setNames(palette()[1:length(unique(ptd))],sort(unique(ptd)))

tiplabels(pie=to.matrix(ptd,sort(unique(ptd))),piecol=cols,cex=0.2)

add.simmap.legend(colors=cols,prompt=FALSE,x=0.9*par()$usr[1],
                  y=-max(nodeHeights(dt8)),fsize=0.6)


#change the colours to something I want in hexadecimal format

cols <- gsub("black", "#FFA500", cols)

cols <- gsub("red", "#009F62", cols)

#lik.anc gives us marginal ancestral states, or "empirical Bayesian posterior
#probabilities" this is using an all-rates-different substitution matrix, as we
#don’t know the mechanism we can’t assume anything about the transition rates.

fitARD<-ace(ptd,dt8,type="discrete", CI = TRUE, model = "ARD")

fitARD$loglik
```

```
#we can try this with an equal rates model (e.g Jukes-Cantor if talking about
#nucleotide substitution models)
fitER<-ace(ptd,dt8,type="discrete", CI = TRUE, model = "ER")
fitER$loglik
#and a symmetrical model
fitSYM<-ace(ptd,dt8,type="discrete", CI = TRUE, model = "SYM")
fitSYM$loglik
#p value for LRT with ML trees
LRT_ML_p_val <- (1 - pchisq(2*abs((fitER$loglik) - (fitARD$loglik)), 1))
LRT_ML_p_val
#plot this with the nodes on the tree
plotTree(dt8, fsize = 0.1, type = "fan")
nodelabels(node=1:dt8$Nnode+Ntip(dt8),
           pie=fitARD$lik.anc,piecol=cols,cex=0.1)
add.simmap.legend(colors=cols,prompt=FALSE,x=0.9*par()$usr[1],
                  y=-max(nodeHeights(dt8)),fsize=0.6)
# make pdf of ancestral estimation probability distribution pie charts on nodes
pdf("anc_state_dt8_piechart_bl_gold.pdf")
plotTree(dt8, ftype="i", fsize = 1, type = "fan")
plotTree(dt8, fsize = 0.1, type = "fan")
nodelabels(node=1:dt8$Nnode+Ntip(dt8),
           pie=fitARD$lik.anc,piecol=cols,cex=0.1)
add.simmap.legend(colors=cols,prompt=FALSE,x=0.9*par()$usr[1],
                  y=-max(nodeHeights(dt8)),fsize=0.6)
dev.off()
#or plot tree displaying one possible version of the ancestral histories across
#the branches
mtree<-make.simmap(dt8,ptd,model="ER")
mtree
plot(mtree,cols,fsize=0.4,ftype="i", type = "fan")
add.simmap.legend(colors=cols,prompt=FALSE,x=0.9*par()$usr[1],
                  y=-max(nodeHeights(dt8)),fsize=0.5)
#make pdf of MCMC tree
pdf("dt8_MCMC_anc_state_bl_gold.pdf")
plot(mtree,cols,fsize=0.4,type = "fan")
add.simmap.legend(colors=cols,prompt=FALSE,x=0.9*par()$usr[1],
                  y=-max(nodeHeights(dt8)),fsize=0.5)
```

```
dev.off()

#Generating a distribution of simulated trees

#making 1000 - afternote, this takes a long time. A hundred might be easier for
#speedy reproducibility but 1000 for effectiveness,
#as the Markov chain will stabilise into an equilibrium distribution as the
#number of simulations approaches infinity

#making 100 for now

equaltrees <-make.simmap(dt8,ptd,nsim=100, Q = "ER")

#So trying the same with 100 simulations but also substituted in a Markov chain
#Monte-Carlo approach to decipher the rate transition matrix (Q)

mtrees100mc<-make.simmap(dt8,ptd,nsim=100, Q = "mcmc")

mtrees100ard<-make.simmap(dt8,ptd,nsim=100, Q = "ARD")

#make a summary of these trees

pd<-summary(equaltrees,plot=FALSE)

pd

#extract the log likelihoods to use in LRT test

equal_loglik <- data.frame()

equal_loglik[1,] <- equaltrees[[1]]$logL

for(i in 1:100){equal_loglik[,i] <- equaltrees[[i]]$logL}

equal_loglik <- as.numeric(equal_loglik)

eqav <- mean(equal_loglik)

#and a summary of the MCMC map

pd2 <- summary(mtrees100mc,plot = FALSE)

pd2

#extract the log likelihoods to use in LRT test

mcmc_loglik <- data.frame()

mcmc_loglik[1,] <- mtrees100mc[[1]]$logL

for(i in 1:100){mcmc_loglik[,i] <- mtrees100mc[[i]]$logL}

mcmc_loglik <- as.numeric(mcmc_loglik)

mcav <- mean(mcmc_loglik)

LRT_stochastic_maps_p_value <- (1-pchisq(2*abs(mcav - eqav), 1))

LRT_stochastic_maps_p_value

#1.88x10^-15

#put the data into more friendly format for equal rates

fiter_DT30 <- fitER$lik.anc

pd_DT30 <- pd$ace

#and for all rates different

fitard_DT30 <- fitARD$lik.anc
```

```
pd2_DT30 <- pd2$ace

Dt30 <- as.data.frame(fitard_DT30, pd2_DT30)

Dt30ER <- as.data.frame(fiter_DT30, pd_DT30)

#compare the marginal states from ACE with our posterior output from 1000 trees

plot(fiter_DT30$DT30,pd_DT30$DT30, title = "Probability of each node DT30",
xlab="Scaled ancestral probability distribution",

     ylab="Posterior probabilities from stochastic mapping", col = "dark blue",
pch = 16)

abline(fit <- lm(pdw_DT30$DT30 ~ fitard_DT30$DT30), lty="dashed",col="red",lwd=2)

legend("topleft", bty="n", legend=paste("Rsquared = ",
format(summary(fit)$adj.r.squared, digits=4)))

#and check the ARDs together

pdf("mcmc_vs_ace_Dt30_only_with_r_2.pdf")

plot(fitard_DT30$DT30,pdw_DT30$DT30, title = "Probability of each node DT30",
xlab="Scaled ancestral probability distribution",

     ylab="Posterior probabilities from stochastic mapping", col = "dark blue",
pch = 16)

abline(fit <- lm (pdw_DT30$DT30 ~ fitard_DT30$DT30), lty="dashed",col="red",lwd=2)

legend("topleft", bty="n", legend=paste("Rsquared = ",
format(summary(fit)$adj.r.squared, digits=4)))

dev.off()

#make the ARD version into a ggplot as reasoning that more likely a switch between
#DT8 and DT30 would have different forward and backward rates... although a strong
#assumption. Must verify experimentally

library(ggplot2)

p <- ggplot(Dt30, aes(Dt30$fitard_DT30.DT30, Dt30$pd2_DT30.DT30)) + geom_point() +
geom_point(data = Dt30, aes(Dt30$fitard_DT30.DT30, Dt30$pd2_DT30.DT30), size = 1,
colour = "dark blue")

p

#change axis labels

p2 <- p + xlab("Scaled maximum likelihood ancestral state DT30 probabilities") +
ylab("Posterior DT30 probabilities from stochastic mapping")

p2

require(stats)

fit

coef = coefficients(fit)

eq <- paste0("y = ", round(coef[2],digits = 2), "*x + ", round(coef[1],digits =
2))

p2 + geom_abline(intercept = -0.0179, slope= 0.8656, colour = "red", linetype =
"dashed", size = 1) + ggtitle(eq)

#add R squared value
```

```
c<- cor(df$fitard_DT30.DT30, df$pd2_DT30.DT30)

rsq <- function (x, y) cor(x, y) ^ 2

rsq(df$fitard_DT30.DT30, df$pd2_DT30.DT30)

#pdf this with a few additions, mainly just the regression line

pdf("analysis_comparison_data_modelled.pdf", height = 5, width = 8)

eq <- paste0("y = ", round(coef[2],digits = 2), "*x + ", round(coef[1],digits =
2),    "      R-squared = 0.901        Probability of node being DT30")

p2 + geom_abline(intercept = -0.0179, slope= 0.8656, colour = "red", linetype =
"dashed", size = 1) + ggtitle(eq)

dev.off()

#also making a density map of the 100 simulated stochastic trees for the equal
#rates transition model, showing how many times each point on the tree was DT8 or
#DT30

obj2<-densityMap(equaltrees,lwd=4, outline = FALSE, fsize = 0.4, type = "fan")

obj2

#and the same for the MCMC generated Q model

obj3<-densityMap(mtrees100mc,lwd=4, outline = FALSE, fsize = 0.4, type = "fan")

#pdf of this

pdf("density_map_dt8_30_lrt_tests_ER.pdf")

densityMap(equaltrees,lwd=5, outline = FALSE, fsize = 0.4, type = "fan")

dev.off()

#pdf of this

pdf("density_map_dt8_30_LRT_tests_MCMC.pdf")

densityMap(mtrees100mc,lwd=5, outline = FALSE, fsize = 0.4, type = "fan")

dev.off()

# Constructing data using our improved model settings
# parsimony assumes that we always have small Q,
#so using joint probability distributions & then sampling from these is a good way
#to optimise the output

#In this case we are sampling from our distribution every 1000 sims to use an
#updated posterior probability of Q
# & then subsequently re-simulating the joint posterior probabilities

mtrees4<-make.simmap(dt8,ptd,nsim=1000, Q = "mcmc")

pd4<-summary(mtrees4,plot=FALSE)

pd4

obj4 <-densityMap(mtrees4,lwd=4, outline = TRUE, fsize = 0.3)

#and this is the same but assessing the difference when Q is unspecified (should
#be the same as we are not using the equal rates model)

mtrees5<-make.simmap(dt8,ptd,nsim=1000)

pd5<-summary(mtrees5,plot=FALSE)
```

```
obj5 <-densityMap(mtrees5,lwd=4, outline = TRUE, fsize = 0.3)
```

Appendix V: Programming Δ Δ Ct for qPCR Output Data

An example of the code generated for programming Δ Δ Ct for qPCR data in R is shown:

```
#Set the correct working directory

setwd("\\\\nbi-cfs2\\ifrprojects\\Rob-
Kingsley\\Oliver\\Genome_variation_S.typhimurium_phage_heavy_metals\\DT8_DT30_swit
ching_analysis\\qPCR_genotyping_wzy\\phage_10_3_techrep")

#read in libraries required, these are for data manipulation & plotting

library(dplyr)
library(tibble)
library(ggplot2)
library(tidyr)
library(gridExtra)

#read in table with qPCR data

t <- read.table("phage_10_replicate_2.txt", header = TRUE, sep = "\t")

#split based on the individual types of reaction occurring, just growth, phage
#challenege & wzy- control growth

t <- split.data.frame(t, t$reaction)

#read dataframes into separate items

g <- t$growth

p <- t$phage

d <- t$delta

#considered the data structure splitting again seems like a useful thing to do

g <- split.data.frame(g, g$Target.Name)

p <- split.data.frame(p, p$Target.Name)

d <- split.data.frame(d, d$Target.Name)

#correct formula for double delta Ct

gddp1 <- (-((g$`wzy +`$rep1 - g$`wzy +`[1,3]) - (g$rpoD$rep1 - g$rpoD[1,3])))

#make into dataframe

gddp1 <- as.data.frame(gddp1)

colnames(gddp1) <- c("rep1p")

gddp1$rep2p <- (-((g$`wzy +`$rep2 - g$`wzy +`[1,4]) - (g$rpoD$rep2 -
g$rpoD[1,4])))

gddp1$rep3p <- (-((g$`wzy +`$rep3 - g$`wzy +`[1,5]) - (g$rpoD$rep3 -
g$rpoD[1,5])))

#repeat for wzynegative
```

```
gddn1 <- (-((g$`wzy -`$rep1 - g$`wzy -`[1,3]) - (g$rpoD$rep1 - g$rpoD[1,3])))

#make into dataframe

gddn1 <- as.data.frame(gddn1)

colnames(gddn1) <- c("rep1n")

gddn1$rep2n <- (-((g$`wzy -`$rep2 - g$`wzy -`[1,4]) - (g$rpoD$rep2 -
g$rpoD[1,4])))

gddn1$rep3n <- (-((g$`wzy -`$rep3 - g$`wzy -`[1,5]) - (g$rpoD$rep3 -
g$rpoD[1,5])))

#repeat for phage wzy+

pddp1 <- (-((p$`wzy +`$rep1 - p$`wzy +`[1,3]) - (p$rpoD$rep1 - p$rpoD[1,3])))

pddp1 <- as.data.frame(pddp1)

colnames(pddp1) <- c("rep1p")

pddp1$rep2p <- (-((p$`wzy +`$rep2 - p$`wzy +`[1,4]) - (p$rpoD$rep2 -
p$rpoD[1,4])))

pddp1$rep3p <- (-((p$`wzy +`$rep3 - p$`wzy +`[1,5]) - (p$rpoD$rep3 -
p$rpoD[1,5])))

#repeat for phage wzy -

pddn1 <- (-((p$`wzy -`$rep1 - p$`wzy -`[1,3]) - (p$rpoD$rep1 - p$rpoD[1,3])))

pddn1 <- as.data.frame(pddn1)

colnames(pddn1) <- c("rep1n")

pddn1$rep2n <- (-((p$`wzy -`$rep2 - p$`wzy -`[1,4]) - (p$rpoD$rep2 -
p$rpoD[1,4])))

pddn1$rep3n <- (-((p$`wzy -`$rep3 - p$`wzy -`[1,5]) - (p$rpoD$rep3 -
p$rpoD[1,5])))

#repeat for negative control wzy -

dddn1 <- (-((d$`wzy -`$rep1 - d$`wzy -`[1,3]) - (d$rpoD$rep1 - d$rpoD[1,3])))

dddn1 <- as.data.frame(dddn1)

colnames(dddn1) <- c("rep1n")

dddn1$rep2n <- (-((d$`wzy -`$rep2 - d$`wzy -`[1,4]) - (d$rpoD$rep2 -
d$rpoD[1,4])))

dddn1$rep3n <- (-((d$`wzy -`$rep3 - d$`wzy -`[1,5]) - (d$rpoD$rep3 -
d$rpoD[1,5])))

ddd <- dddn1

gddp1[,4] <- gddn1$rep1n

gddp1[,5] <- gddn1$rep2n

gddp1[,6] <- gddn1$rep3n

gdd <- gddp1

pddp1[,4] <- pddn1$rep2n

pddp1[,5] <- pddn1$rep2n
```

```
pddp1[,6] <- pddn1$rep3n
pdd <- pddp1
colnames(gdd) <- c("rep1p", "rep2p", "rep3p", "rep1n", "rep2n","rep3n")
colnames(pdd) <- c("rep1p", "rep2p", "rep3p", "rep1n", "rep2n","rep3n")
colnames(ddd) <- c("rep1n", "rep2n", "rep3n")
#add mean and standard error to table of ddd
ddd$stderdn <- sd(ddd[1,1:3])/sqrt(3)
for (i in 1:6){ddd$stderdn[i] <- sd(ddd[i,1:3]/sqrt(3))}
ddd$meann <- (ddd$rep1n + ddd$rep2n + ddd$rep3n)/3
#add mean and standard error to table of gddp
gdd$stderdp <- sd(gdd[1,1:3])/sqrt(3)
for (i in 1:6){gdd$stderdp[i] <- sd(gdd[i,1:3]/sqrt(3))}
gdd$stderdn <- sd(gdd[1,4:6])/sqrt(3)
for (i in 1:6){gdd$stderdn[i] <- sd(gdd[i,4:6]/sqrt(3))}
gdd$meanp <- (gdd$rep1p + gdd$rep2p + gdd$rep3p)/3
gdd$meann <- (gdd$rep1n + gdd$rep2n + gdd$rep3n)/3
#add mean and standard error for pddp
pdd$stderdp <- sd(pdd[1,1:3])/sqrt(3)
for (i in 1:6){pdd$stderdp[i] <- sd(pdd[i,1:3]/sqrt(3))}
pdd$stderdn <- sd(pdd[1,4:6])/sqrt(3)
for (i in 1:6){pdd$stderdn[i] <- sd(pdd[i,4:6]/sqrt(3))}
pdd$meanp <- (pdd$rep1p + pdd$rep2p + pdd$rep3p)/3
pdd$meann <- (pdd$rep1n + pdd$rep2n + pdd$rep3n)/3
ddd$group <- c("n")
#put each output into a tidy format for ggplot
time <- c(0,2,4,6,8,24)
d2 <- data.frame(time,ddd$meann,ddd$stderdn,ddd$group)
colnames(d2) <- c("time","mean","stder","group")
d3 <- d2
d3$stder[1] <- 0
#and do for growth and phage
#making intermediate table then merging based on group
#for growthwzy+
gmidp <- data.frame(time)
gmidp$mean <- gdd$meanp
gmidp$stder <- gdd$stderdp
gmidp$group<- "p"
```

```
#for growth wzy-
gmidn <- data.frame(time)
gmidn$mean <- gdd$meann
gmidn$stder <- gdd$stderdn
gmidn$group<- "n"
#when merging here apparently you need to specific all = TRUE to get the values
g2 <- merge(gmidp, gmidn, all = T)
g3 <- g2
g3$stder[1] <- 0
g3$stder[2] <- 0
#for phagewzy+
pmidp <- data.frame(time)
pmidp$mean <-pdd$meanp
pmidp$stder <- pdd$stderdp
pmidp$group<- "p"
#for growth wzy-
pmidn <- data.frame(time)
pmidn$mean <- pdd$meann
pmidn$stder <- pdd$stderdn
pmidn$group<- "n"
p2 <- merge(pmidp, pmidn, all = T)
p3 <- p2
p3$mean[1] <- 0
p3$stder[1] <- 0
p3$stder[2] <- 0
d3$time <- as.factor(d3$time)
g3$time <- as.factor(g3$time)
p3$time <- as.factor(p3$time)
#change data to log format so that it makes sense
#d$mean <- log2(d$mean)
#d$stder <- log2(d$stder)
#g$mean <- log2(g$mean)
#g$stder <- log2(g$stder)
#p$mean <- log2(p$mean)
#p$stder <- log2(p$stder)
#plot each as line graph with standard error
dplot <- ggplot(d3, aes(time,mean)) +
```

```
geom_point(size = 3, color = "black") +
  xlab("Time after phage innoculum (h)") +
  ylab("- delta delta Ct (rpoD relative)") +
  ggtitle("L01157-10wzy- Growth") +
  scale_color_manual(values=c("#009F62","#FFA500")) +
  geom_errorbar(aes(ymin = mean - stder,
                    ymax = mean + stder), size = 1.2, width = 0.2, color =
"#FFA500")  +
  coord_cartesian(ylim = c(-10, 10)) + geom_line(color = "#FFA500", size = 2,
group = d3$group)
dplot
gplot <- ggplot(g3, aes(time,mean, col = group, group = group)) +
  geom_line(size = 1.5) +
  geom_point(size = 3, color = "black") +
  xlab("Time after phage innoculum (h)") +
  ylab("- delta delta Ct (rpoD relative)") +
  ggtitle("L01157-10 WT Growth") +
  scale_color_manual(values=c("#FFA500", "#009F62")) +
  coord_cartesian(ylim = c(-10,10)) +
                    geom_errorbar(aes(ymin = mean - stder, ymax = mean + stder),
size = 1.5, width = 0.2)
gplot
pplot <- ggplot(p3, aes(time,mean, col = group, group = group)) +
  geom_line(size = 1.5) +
  geom_point(size = 3, color = "black") +
  xlab("Time after phage innoculum (h)") +
  ylab("- delta delta Ct (rpoD relative)") +
  ggtitle("L01157-10 3.4e4 PFU Phage-10") +
  scale_color_manual(values=c("#FFA500", "#009F62")) +
  coord_cartesian(ylim = c(-10,10)) +
  geom_errorbar(aes(ymin = mean - stder, ymax = mean + stder), size = 1.5, width =
0.2)
pplot
all <- grid.arrange(gplot, dplot, pplot, nrow=3)
pdf("phage_10_triplicate.pdf")
all <- grid.arrange(gplot, dplot, pplot, nrow=3)
dev.off()
```

Appendix VI: Sorting ARIBA Data Output

```
#This script is for sorting the output "all reports" file from ARIBA,
#using a binary presence absence approach

#load packages

library(tidyverse)
library(reshape)
library(dplyr)
library(wrapr)
library(ggtree)
library(phangorn)
library(ggplot2)
library(ape)

#set correct working directory

setwd("\\\\IFR-Group-Data\\ifr-research-groups\\Rob-
Kingsley\\PHE\\fastq2\\fastq\\symlink_1.1.1_only_analyses\\ariba_SJ46_gene_by_gene
\\ariba_SJ46_1\\ALL_REPORTS")

#load up all the reports

temp <- list.files(pattern="*mTmII")

#making a dataframe list of lists

f <- lapply(temp, read.delim)

#make copy of frame list without changes

f3 <- f

#removing all except the first column

f2 <- lapply(f, function(x){x[,-c(2:29)]})

#removing all except the first column after checking numbers are correct

f2 <- lapply(f2, function(x){x[,-c(2:3)]})

#restore current version

f3 <- f2

#remove duplicates of any gene name or strain name

f4 <- lapply(f3, function(x){x[!duplicated(x),]})

#restore current version

f5 <- f4

#change column name of gene names to strain name

f6 <- lapply(f5, function(x){x[,3] = x[,1]})

#write each as a csv file

lapply(f6, function(x){for (i in 1:1697){write.csv(x, paste("mTmII", i,
".csv"))}})

#make new data into dataframes

f6 <- lapply(f6, function(x){as.data.frame(x)})
```

```
f7 <- f6
#add single cell with strain name
for(i in 1:1697){f6[[i]][1,2] <- f5[[i]][1,2]}
#restore version
f7 <- f6
#change column names
for (i in 1:1697){colnames(f7[[i]]) <- c(paste(f7[[i]][1,2]))}
#restore version
f8 <- f7
#and remove column 2 from all
for (i in 1:1697){f7[[i]][,2] <- NULL}
#transpose the dataframes
f9 <- lapply(f7, function(x){t(x)})
#bind all into single dataframe
f10 <- data.frame(matrix(unlist(f9), nrow=1697, byrow=T),stringsAsFactors=FALSE)
#write each as a csv file
for (i in 1:1697){write.csv(f9[[i]], paste0("mTmII",sep = "_",i,sep =
"_",".csv"))}
#load up new file after concatenation with linux (cat mTmII* >
#all_ariba_presence_absence.csv)
mTmII <- read.csv("all_ariba_presence_absence.csv", header = FALSE)
#remove excess from names
mTmII$V1 <- gsub("_._mTmII", "", mTmII$V1)
mTmII$V1 <- gsub("_.._mTmII", "", mTmII$V1)
mTmII$V1 <- gsub("_..._mTmII", "", mTmII$V1)
mTmII$V1 <- gsub("_...._mTmII", "", mTmII$V1)
#make column 1 the row names
mTmII <- `row.names<-`(mTmII, mTmII$V1)
mTmII$V1 <- NULL
#restore current version
mTmII2 <- mTmII
#read presence absence file
mTmII <- read.csv("mTmII_pres_abs.csv", header = TRUE, stringsAsFactors = F)
mTmII <- `row.names<-`(mTmII, mTmII$X)
mTmII$X <- NULL
#align all data using which and checking for the number existing in the row with
if(x %in % y)
l <- list()
```

```
u <-
c("O4698_09_04338","O4698_09_04339","O4698_09_04340","O4698_09_04341","O4698_09_04342","O4698_09_04343","O4698_09_04344","O4698_09_04345","O4698_09_04346","O4698_09_04347","O4698_09_04348","O4698_09_04349","O4698_09_04350","O4698_09_04351","O4698_09_04352","O4698_09_04353","O4698_09_04354","O4698_09_04355","O4698_09_04356","O4698_09_04357","O4698_09_04358","O4698_09_04359","O4698_09_04360","O4698_09_04361","O4698_09_04362","O4698_09_04363","O4698_09_04364","O4698_09_04365","O4698_09_04366","O4698_09_04367","O4698_09_04368","O4698_09_04369","O4698_09_04370","O4698_09_04371","O4698_09_04372","O4698_09_04373","O4698_09_04374","O4698_09_04375","O4698_09_04376","O4698_09_04377","O4698_09_04378","O4698_09_04379","O4698_09_04380","O4698_09_04381","O4698_09_04382","O4698_09_04383","O4698_09_04384","O4698_09_04385","O4698_09_04386","O4698_09_04387","O4698_09_04388","O4698_09_04389","O4698_09_04390","O4698_09_04391","O4698_09_04392","O4698_09_04393","O4698_09_04394","O4698_09_04395","O4698_09_04396","O4698_09_04397","O4698_09_04398","O4698_09_04399")

h <- NULL

df <- mTmII

for(k in 1:nrow(df)){

  # create a list for each row of the df

  l[[k]] <- df[k, ]

  for(i in 1:length(l[[k]])){

    #check if number exists in the row

    if(u[i] %in % l[[k]]){

      # find the index of the number given it exists

      a <- which(l[[k]] == u[i])

      #assign to "help" vector in order to not overwrite values

      h[i] <- l[[k]][a]

    }

    else{

      #numbers that do not exist in the vector are assigned NA

      h[i] <- NA

    }

  }

  #replace row by sorted vector with NA place holders ("help" vector)

  l[[k]] <- h

}

#transfer from df1 to df2 and make backups of data.frames

df2 <- df1; df3 <- df; for(i in 1:nrow(df)){df[i,] <- df1[i,]}

#change data type to vector

for (i in ncol(df)){df[,i] <- vapply(df[,i], paste, collapse = ", ",
character(1L))}

#change words to 1
```

```
for(i in 1:62){df[,i] <- gsub("......1..","1", df[,i])}
#change NA to 0
for(i in 1:62){df[,i] <- gsub(".....NA.","0", df[,i])}
#write csv of dataframe
write.csv(df, "aligned.binary.mTmII.csv")
#make heatmap with data
p <- read.tree("193_120_tree_no_outgroup.newick")

#plot tree to view and identify any gaps
p <- ggtree(p, size=0.01, color="black")
print(p)
p <- p +theme(
  panel.background = element_rect(fill = "transparent",colour = NA),
  plot.background = element_rect(fill = "transparent",colour = NA)
)
p1 <- p + geom_tiplab(size = 0.01, aes(label=label),
                      hjust=0, align= F) +geom_treescale(color="black", width =
0.0005, x = 0, y = 300, linesize=0.09, fontsize = 3, offset = 0)
print(p1)
#plot heat map
heat_pt <- gheatmap(p1, df, offset = 0.001, width = 3, colnames = TRUE, font.size
= 0.5)
plot(heat_pt)
#pdf this
pdf("mTmII_on_tree_heatmap.pdf")
heat_pt
dev.off()
```

Appendix VII: Linux Blast-Based Script for Finding and Sorting Alleles in Many Sequences at Once

```
#!/bin/bash
# polymorphism_finder.sh
echo "This is polymorphism finder 1.0"
echo "Current directory is"
pwd
```

```
#source blast on the HPC cluster
source blast-2.6.0;

#or if using CLIMB the previous command will spit an error, and you can just
continue with this:

#make a database for each .fasta file in the current folder

for f in *.fasta ;do makeblastdb -in $f  -parse_seqids -dbtype nucl; done;

name={insert_sequence_name}.fasta

rm $name.fasta.nhr
rm $name.fasta.nin
rm $name.fasta.nog
rm $name.fasta.nsd
rm $name.fasta.nsi
rm $name.fasta.nsq

mkdir $name

for f in *.fasta; do blastn -db $f -query $name.fasta -out $f.$name.region -outfmt
5; done;

for f in *.$name.region; do

#Different outputs can require different lines specified by sed '{line_number}p'
#This can easily be found by just looking in the output file and seeing which line
#the result is on; it always starts with <Hsp_hseq>

sed -n '111p' $f | sed -r 's/<Hsp_hseq>//' | awk '$1=$1' | sed -r
's/<\/Hsp_hseq>//' > $f.$name.fasta

done; for file in *$name.fasta; do echo ">$file"$'\n'"$(cat -- "$file")" >
"$file"; done; rm *.region

cat *.$name.fasta > concatenated_$name.fasta

cp *$name.fasta $name

cd $name

for f in *.fasta; do sed -e "s/.\{60\}/&\n/g" $f > ${f}.split.fasta; done

rm *.fasta.${name}.region.${name}.fasta

cd ..

rm *.fasta.${name}.region.${name}.fasta

#if finished with the databases of each sequence these can be removed with:

cd ..

rm *.fasta.nhr
rm *.fasta.nin
rm *.fasta.nog
rm *.fasta.nsd
rm *.fasta.nsi
rm *.fasta.nsq
```

Appendix VIII: Strain Information for Figure IV.4

strain	ST	PT[a]	accession no.	clade	chrom size[b]	pSLT[c]	P1[c]	P2[c]	ECC[d]	CDS[e]	SNPs[f]	Prophage[g]	total HAC[h]	spec. HAC	year isolation	source
SO9207-07	19	DT170B	PRJEB34598	α11	4.92	-	-	-	0	4592	973	7 (5:1:1)	48	7	2007	Pig
SO1960-05	19	U288	PRJEB34597	α12	4.90	154.4	19.4	18.2	3	4832	1272	8 (5:2:1)	57	16	2005	Pig
NCTC13348	19	DT104	HF937208.1	α15	4.93	94.0	-	-	1	4751	1083	9 (5:3:1)	63	25	1988	Human
SO4698-09	34	DT193	LN999997.1	α17	5.04	-	-	-	0	4754	1010	9 (6:2:1)	43	7	2009	Cattle
SO7676-03	19	DT56	PRJEB34599	β5	4.88	-	-	-	0	4570	797	6 (4:1:1)	75	26	2003	Bird
SO9304-02	19	DT41	PRJEB34596	β4	5.05	117.4	32.2		3	4685	943	8 (5:2:1)	78	15	2002	Cattle
SL1344	19	DT44	FQ312003.1	β1	4.88	93.8	86.9	8.7	3	4771	0	9 (4:3:2)	42	2	~1966	Cattle
LO1157-10	19	DT8	PRJEB34595	β2	4.86	93.8	23.9	22.2	3	4706	1101	7 (4:2:1)	62	19	2010	Duck
D23580	313	-	FN424405.1	β1	4.88	117.0	84.6	-	2	4790	901	7 (5:2:0)	57	9	2004	Human
A130	313	DT56v	PRJEB34594	β1	4.93	166.9	-	-	1	4812	901	8 (5:2:1)	62	13	2001	Human
94–213	98	DT2	HG326213.1	β3	4.82	93.8	-	-	1	4598	903	5 (3:1:1)	66	20	1994	Pigeon

[a] Phage type / Sequence type
[b] Chromosome size (Mbp)
[c] contig size (Kbp)
[d] extra chromosomal contigs (ECC), putative plasmid sequences
[e] Number of predicted coding sequences
[f] number of SNPs with reference to SL1344
[g] Number of prophage (intact, incomplete, candidate)
[h] total number of pseudogenes with reference to SL1344 allele, [i] number of clade-specific pseudogenes (S1 Table), [i] mean delta bitscore (DBS) with reference to SL1344 protein orthologue.

https://doi.org/10.1371/journal.pgen.1008850.t001

From Bawn et al., (2020) (39).

Appendix IX: Strains used in this study

Strain	Phage type	Host	Isolation date	Region	Reference/Institute
10354_1997	DT99	Duck	1997	UK	APHA
6398_05	DT30	Duck	2005	UK	APHA
S10_2762	DT8	Duck	2010	Ireland	APHA
2029_1997	DT99	Duck	1997	UK	APHA
568_1998	DT8	Duck	1998	UK	APHA
7392_05	DT8	Duck	2005	UK	APHA
S10_3104	DT8	Duck	2010	Ireland	APHA
S02130_05	?	Chicken	2005	UK	APHA
7942_2000	DT8	Duck	2000	UK	APHA
L01000_05	DT8	sheep	2005	UK	APHA
S10_1175	DT8	Duck	2010	Ireland	APHA
1590_2000	DT8	Duck	2000	UK	APHA
S05314_06	DT8	turkey	2006	UK	APHA
S04695_09	DT1	Duck	2009	UK	APHA
S01631_10	DT30	Duck	2010	UK	APHA

S04551_10	DT8	Duck	2010	UK	APHA
L00244_10	DT8	Horse	2010	UK	APHA
S05989_06	DT8	cattle	2006	UK	APHA
S05479_09	DT8	chicken	2009	UK	APHA
S01682_10	DT8	Duck	2010	UK	APHA
S04572_10	DT8	Duck	2010	UK	APHA
L00745_10	DT8	pheasant	2010	UK	APHA
3041_05	DT8	Cattle	2005	UK	APHA
S02889_11	DT193	Duck	2011	UK	APHA
S01045_06	DT193	Duck	2006	UK	APHA
S06614_06	DT8	Duck	2006	UK	APHA
S00869_09	DT8	cattle	2009	UK	APHA
S04584_10	DT8	Duck	2010	UK	APHA
S11_967	DT8	Duck	2011	Ireland	APHA
S02194_06	DT8	Duck	2006	UK	APHA
S08362_06	DT30	Duck	2006	UK	APHA
L00658_09	DT8	Duck	2009	UK	APHA
S04830_10	U302	Duck	2010	UK	APHA
S10_2446	DT8	Duck	2010	Ireland	APHA
S00687_11	DT8	Duck	2011	UK	APHA
S11_1001	DT8	Duck	2011	Ireland	APHA
S02784_06	DT30	Duck	2006	UK	APHA
L00778_06	DT41	Duck	2006	UK	APHA
L00625_09	DT8	pheasant	2009	UK	APHA
S01972_10	DT8	guinea fowl	2010	UK	APHA
S04838_10	DT8	Duck	2010	UK	APHA
S01710_11	DT8	Duck	2011	UK	APHA
4203_2013	DT8	Duck	2013	Italy	APHA
S11_2195	DT8	chicken	2011	Ireland	APHA
7792_1996	DT30	feed (home)	1996	UK	APHA
S02934_06	DT8	Duck	2006	UK	APHA
L00863_06	U288	Duck	2006	UK	APHA
S05191_07	DT8	Duck	2007	UK	APHA
S02493_10	DT30	chicken	2010	UK	APHA
S01838_11	DT8	cattle	2011	UK	APHA
4204_2013	DT8	Duck	2013	Italy	APHA
S11_2473	DT8	Duck	2011	Ireland	APHA
S03122_06	DT8	goose	2006	UK	APHA
S03899_08	DT8	goose	2008	UK	APHA
S05825_07	DT30	avian	2007	UK	APHA
3990_2000	DT104	Duck	2000	UK	APHA
S02532_10	DT8	chicken	2010	UK	APHA
S05376_10	U302	Duck	2010	UK	APHA

S01872_11	DT30	Duck	2011	UK	APHA
4361_2013	DT8	Duck	2013	France	APHA
S11_3247	DT8	Duck	2011	Ireland	APHA
S03232_06	DT195	Duck	2006	UK	APHA
S05611_08	DT8	chicken	2008	UK	APHA
S10101_07	DT8	chicken	2007	UK	APHA
S02996_10	DT8	chicken	2010	UK	APHA
S05940_10	U302	Duck	2010	UK	APHA
5672_03	DT8	Duck	2003	UK	APHA
S01887_11	DT8	Duck	2011	UK	APHA
4362_2013	DT8	Duck	2013	France	APHA
S11_3366	DT8	Duck	2011	Ireland	APHA
S03254_06	DT2	Duck	2006	UK	APHA
S05712_08	DT8	Duck	2008	UK	APHA
S0119_10	DT8	turkey	2010	UK	APHA
S04091_10	DT8	Duck	2010	UK	APHA
S05946_10	U302	Duck	2010	UK	APHA
S01944_11	DT8	Duck	2011	UK	APHA
4363_2013	DT8	Duck	2013	France	APHA
5486_05	DT30	Duck	2005	UK	APHA
S11_3652	DT8	Duck	2011	Ireland	APHA
S04001_06	DT30	goose	2006	UK	APHA
L00555_08	n/a	Duck	2008	UK	APHA
S00473_10	DT8	donkey	2010	UK	APHA
S04500_10	untypable	Duck	2010	UK	APHA
S05982_10	DT8	turkey	2010	UK	APHA
S02028_11	DT8	elephant	2011	UK	APHA
4264_2013	DT8	Duck	2013	France	APHA
S12_7	DT8	Duck	2012	Ireland	APHA
S01299_09	DT8	Duck	2009	UK	APHA
531_1998	DT8	Duck	1998	UK	APHA
S01345_10	DT8	Duck	2010	UK	APHA
S04501_10	DT8	Duck	2010	UK	APHA
L01157_10	DT8	chicken	2010	UK	APHA
S02230_11	DT8	horse	2011	UK	APHA
S12_71	DT8	Duck	2012	Ireland	APHA
S05303_06	DT8	cattle	2006	UK	APHA
S04527_10	DT8	feed	2010	UK	APHA
H100920317	DT8	Human	2010	UK	APHA
S02264_13	DT30	Duck	2013	UK	APHA
S10_002590	DT8	Duck	2010	Ireland	APHA
H102860398	DT8	Human	2010	UK	APHA
H122460450	DT30	Human	2012	UK	APHA
S01085_13	DT30	Duck	2013	UK	APHA

S00829_13	DT193	Duck	2013	UK	APHA
H100580475	DT8	Human	2010	UK	APHA
S10_002570	DT8	Duck	2010	UK	APHA
S02433_12	DT8	partridge	2010	UK	APHA
S04039_12	DT8	Duck	2010	UK	APHA
6440_1994	DT8	Cattle	1994	UK	APHA
H102500306	DT8	Human	2010	UK	APHA
H102640297	DT8	Human	2010	UK	APHA
H101540365	DT8	Human	2010	UK	APHA
S10_002446	DT8	Duck	2010	UK	APHA
4798_1994	DT8	Dog	1994	UK	APHA
S06292_12	DT30	Duck	2012	UK	APHA
H102420629	DT8	Human	2010	UK	APHA
H102520502	DT8	Human	2010	UK	APHA
S10_002696	DT8		2010	UK	APHA
Salmoporc	DT9			UK	APHA
S02693_12	DT8	horse	2012	UK	APHA
S10_002693	DT8	Duck	2010	UK	APHA
H102340424	DT8	Human	2010	UK	APHA
H102360334	DT8	Human	2010	UK	APHA
H122360467	DT30	Human	2012	UK	APHA
S03079_13	DT41	Duck	2013	UK	APHA
S12_000794	DT8	Duck	2012	UK	APHA
H102060263	DT8	Human	2010	UK	APHA
H102260537	DT8	Human	2010	UK	APHA
H101700306	DT8	Human	2010	UK	APHA
H114040550	DT30	Human	2011	UK	APHA
S02370_12	DT8	pheasant	2012	UK	APHA
S10_00037073	DT8	ND	2010	UK	APHA
S11_002572	DT8	Duck	2011	UK	APHA
S03086_11	DT8	parrot	2011	UK	APHA
H102060263	DT8	Human	2010	UK	APHA
H101000361	DT8	Human	2010	UK	APHA
H095080368	DT8	Human	2009	UK	APHA
H113860446	DT30	Human	2011	UK	APHA
H093700684	DT8	Human	2010	UK	APHA
S06231_12	DT41	Duck		UK	APHA
S02693_13	DT8	pheasant	2013	UK	APHA
H101660105	DT8	Human	2010	UK	APHA
H095080368	DT8	Human	2009	UK	APHA
H093060663	DT8	Human	2009	UK	APHA
H102860373	DT8	Human	2010	UK	APHA
H101320108	DT8	Human	2010	UK	APHA
H102980421	DT8	Human	2010	UK	APHA

S02482_12	DT8	pheasant	2012	UK	APHA
H101540365	DT8	Human	2010	UK	APHA
H094200556	DT8	Human	2009	UK	APHA
H092460168	DT8	Human	2009	UK	APHA
H122360547	DT30	Human	2012	UK	APHA
H102860371	DT8	Human	2010	UK	APHA
ST571_TM				UK	APHA
S10_000865	DT8			UK	APHA
H122380533	DT30	Human	2012	UK	APHA
H093980731	DT8	Human	2009	UK	APHA
H092040408	DT8	Human	2009	UK	APHA
H122420319	DT30	Human	2012	UK	APHA
H122360553	DT30	Human	2012	UK	APHA
H101660105	DT8	Human	2010	UK	APHA
S03645_11	DT30	Duck	2011	UK	APHA
S10_002348	DT8		2010	UK	APHA
H093880802	DT8	Human	2009	UK	APHA
H091040136	DT8	Human	2009	UK	APHA
H093400237	DT8	Human	9	UK	APHA
S10_003104	DT8	Duck	2010	UK	APHA
3232_07				UK	APHA
S06298_12	DT8	Duck	2012	UK	APHA
S06742_12	DT8	Duck	2012	UK	APHA
S02930_13	DT41	Duck	2013	UK	APHA
S03679_11	DT30	Duck	3011	UK	APHA
S2924_11	DT8	Turkey	2011	UK	APHA
6441_1994	DT8	Cattle	1994	UK	APHA
S02791_12	DT8	Duck	2012	UK	APHA
8424_1997	DT8	Duck	1997	UK	APHA
284_1998	DT104	Duck	1998	UK	APHA
6371_1993	DT8	Cattle	1993	UK	APHA
3825_1995	DT104	Duck	1995	UK	APHA
10028_1996	DT208	Duck	1996	UK	APHA
2270_1997			1997	UK	APHA
758_2001	DT125	Duck	2001	UK	APHA
3131_04	DT8	Duck	2004	UK	APHA
S10_698	DT8	Duck	2010	UK	APHA
5127_1995	DT40	Duck	1995	UK	APHA
S10_1059	DT8	Duck	2010	UK	APHA
5175_1995	DT104	Duck	1995	UK	APHA
11087_1996	DT193	Duck	1996	UK	APHA
2883_2001	DT30	Duck	2001	UK	APHA
3397_1993	DT8	chicken	1993	UK	APHA
4875_1997	DT195	Duck	1997	UK	APHA

7403_02	DT30	Duck	2002	UK	APHA
8049_1993	DT8	Dog	1993	UK	APHA
6923_1996	n/a	Duck	1996	UK	APHA
6157_1997	DT8	Duck	1997	UK	APHA
6425_1998	DT8	Duck	1998	UK	APHA
7399_1996	DT120	Duck	1996	UK	APHA
6886_1997	DT30	Duck	1997	UK	APHA
9606_02	DT30	Duck	2002	UK	APHA
4459_1994	DT193	Duck	1994	UK	APHA
7499_1996	DT8	Cattle	1996	UK	APHA
2291_05	DT8	Duck	2005	UK	APHA
S2733_1997	DT195	Duck	1997	UK	APHA
5983_1993	DT40	Duck	1993	UK	APHA
S10524_1996	DT9	Duck	1996	UK	APHA
3517_1997	DT195	Duck	1997	UK	APHA
2510_1998	DT30	feed	1998	UK	APHA
2437_2001	DT30	turkey	2001	UK	APHA
6281_04	DT30	Duck	2004	UK	APHA
103_1993	DT30	Duck	1993	UK	APHA
3519_1997	DT8	Duck	1997	UK	APHA
4433_1998	DT104	Duck	1998	UK	APHA
L01857_04	DT41	Duck	2004	UK	APHA
12232_1995	DT104	Duck	1995	UK	APHA
11088_1996	DT9	Duck	1996	UK	APHA
4974_1998	DT1	Duck	1998	UK	APHA
1280_05	DT66	Duck	2005	UK	APHA
S10_1916	DT8	avian	2010	UK	APHA
11553_1996	DT9	Duck	1996	UK	APHA
7410_02	DT8	chicken	2002	UK	APHA
1822_05	DT8	Duck	2005	UK	APHA
S10_1908	DT8	cattle	2010	UK	APHA
10359_1993	DT8	Duck	1993	UK	APHA
10017_1996	DT8	feed	1996	UK	APHA
9358_1998	DT99	Duck	1998	UK	APHA
2031_05	DT41	Duck	2005	UK	APHA
S10_2137	DT8	cat	2010	UK	APHA
137_1997	DT30	Duck	1997	UK	APHA
6888_1997	DT8	Duck	1997	UK	APHA
1426_2000	DT8	pheasant	2000	UK	APHA
60_03	DT8	Duck	2003	UK	APHA
2684_1994	DT8	Duck	1994	UK	APHA
7668_1996	DT8	Duck	1996	UK	APHA
759_1997	DT104	Duck	1997	UK	APHA
183_03	DT8	Duck	2003	UK	APHA

2780_1994	DT8	Duck	1994	UK	APHA
925_1997	DT99	Duck	1997	UK	APHA
156_1998	DT8	chicken	1998	UK	APHA
S10_2590	DT8	Duck	2010	UK	APHA
8674_1994	DT104	Duck	1994	UK	APHA
8270_1996	DT8	turkey	1996	UK	APHA
927_1997	DT8	Duck	1997	UK	APHA
5269_2000	DT40	Duck	2000	UK	APHA
7423_03	DT30	Duck	2003	UK	APHA
S01564_10	DT8	alpaca	2010	UK	APHA
8378_1996	DT8	Duck	1996	UK	APHA
5935_2000	DT8	turkey	2000	UK	APHA
L01344_03	DT8	cattle	2003	UK	APHA
9000_1996	DT40	Duck	1996	UK	APHA
2381_1997	DT195	Duck	1997	UK	APHA
870_1998	DT8	horse	1998	UK	APHA
65_04	DT8	Duck	2004	UK	APHA
S11_525	DT8	cattle	2011	UK	APHA
S12_118	DT8	Duck	2012	UK	APHA
S02888_11	DT8	Duck	2011	UK	APHA
S10_1308	DT8	Duck	2010	UK	APHA
S12_429	DT8	Duck	2012	UK	APHA
S11_966	DT8	cattle	2011	UK	APHA
S01690_10	DT30	Duck	2010	UK	APHA
S02918_11	DT8	Duck	2011	UK	APHA
S01836_10	DT8	cattle	2010	UK	APHA
S02919_11	DT8	Duck	2011	UK	APHA
S04839_10	DT30	Duck	2010	UK	APHA
S04003_06	DT8	goose	2006	UK	APHA
S02243_09	DT8	pheasant	2009	UK	APHA
S01509_10	DT30	goose	2010	UK	APHA
L1291_10	DT8	Duck	2010	UK	APHA
S02756_11	DT30	turkey	2011	UK	APHA
S05634_12	DT41	Duck	2012	UK	APHA
S3202_07	?	Duck	2007	UK	APHA
4904_1996	DT8	Duck	1996	UK	APHA
6758_1996	DT8	Duck	1996	UK	APHA
6882_1997	DT8	Duck	1997	UK	APHA
2358_1998	DT104	Duck	1998	UK	APHA
S10_002137	DT8	cat	2010	UK	APHA
S6656_12	DT9		2012	UK	APHA
MS57	DT30			Ireland	Mohammed et al., 2015
PB225	DT8			Ireland	Mohammed et al., 2015

PB469	DT8			ireland	Mohammed et al., 2015
PB880	DT8			Ireland	Mohammed et al., 2015
H124521012	DT 120	CHICKEN	2012	UK	Public Health England
H124521014	DT 120	CHICKEN	2012	UK	Public Health England
H124521018	DT 120	POULTRY	2012	UK	Public Health England
H124521082	DT 135	Human	2012	UK	Public Health England
H124521083	DT 8	Human	2012	UK	Public Health England
H124521119	DT 135	Human	2012	UK	Public Health England
H124521124	DT 9	Human	2012	UK	Public Health England
H124521125	DT 208	Human	2012	UK	Public Health England
H124521137	DT 64	Human	2012	UK	Public Health England
H124521169	DT 135	Human	2012	UK	Public Health England
H124521176	RDNC	Human	2012	UK	Public Health England
H124521199	DT 64	Human	2012	UK	Public Health England
H124521212	DT 208	Human	2012	UK	Public Health England
H124521302	DT 35	CHICKEN	2012	UK	Public Health England
H124521308	DT 135	CHICKEN	2012	UK	Public Health England
H124521316	DT 30	CHICKEN	2012	UK	Public Health England
H124521322	DT 135	POULTRY	2012	UK	Public Health England
H124521330	DT 8	ANIMAL	2012	UK	Public Health England
H124521333	DT 35	POULTRY	2012	UK	Public Health England
H124521334	DT 64	ANIMAL	2012	UK	Public Health England

H132260914	DT 120	CHICKEN	2013	UK	Public Health England
H132260959	DT 120	Human	2013	UK	Public Health England
H132260971	DT 64	Human	2013	UK	Public Health England
H133020731	DT 8	Human	2013	UK	Public Health England
H141420383	None	Human	2014	UK	Public Health England
H141420385	DT 193	Human	2014	UK	Public Health England
H141420391	DT 193	Human	2014	UK	Public Health England
H141420393	RDNC	Human	2014	UK	Public Health England
H141420398	DT 104	Human	2014	UK	Public Health England
H141420400	DT 193	Human	2014	UK	Public Health England
H141420402	DT 193	Human	2014	UK	Public Health England
H141460669	DT 193	Human	2014	UK	Public Health England
H141480357	DT 15	Human	2014	UK	Public Health England
H141480359	DT 177	Human	2014	UK	Public Health England
H141480439	DT 56	Human	2014	UK	Public Health England
H141500474	DT 193	Human	2014	UK	Public Health England
H141500475	None	Human	2014	UK	Public Health England
H141500476	DT 193	Human	2014	UK	Public Health England
H141500477	DT 193	Human	2014	UK	Public Health England
H141500478	RDNC	Human	2014	UK	Public Health England
H141500484	DT 193	Human	2014	UK	Public Health England
H141500485	DT 193	Human	2014	UK	Public Health England

H141520436	RDNC	Human	2014	UK	Public Health England
H141520468	None	Human	2014	UK	Public Health England
H141520470	DT 56	Human	2014	UK	Public Health England
H141540514	DT 104	Human	2014	UK	Public Health England
H141540529	DT 120	Human	2014	UK	Public Health England
H141560271	DT 193	Human	2014	UK	Public Health England
H141580495	DT 193	RAW MEAT	2014	UK	Public Health England
H141580499	DT 104	Human	2014	UK	Public Health England
H141580500	DT 195	Human	2014	UK	Public Health England
H141580501	DT 143	Human	2014	UK	Public Health England
H141580502	RDNC	Human	2014	UK	Public Health England
H141580503	DT 193	Human	2014	UK	Public Health England
H141580513	DT 56	Human	2014	UK	Public Health England
H141600085	DT 193	Human	2014	UK	Public Health England
H141600086	DT 193	Human	2014	UK	Public Health England
H141600088	DT 193	Human	2014	UK	Public Health England
H141600092	None	Human	2014	UK	Public Health England
H141600104	None	Human	2014	UK	Public Health England
H141600104	None	Human	2014	UK	Public Health England
H141600104	None	Human	2014	UK	Public Health England
H141600105	None	Human	2014	UK	Public Health England
H141620518	DT 56	Human	2014	UK	Public Health England

H141620518	DT 56	Human	2014	UK	Public Health England
H141620520	Untypable	Human	2014	UK	Public Health England
H141620520	Untypable	Human	2014	UK	Public Health England
H141620522	DT 8	Human	2014	UK	Public Health England
H141620522	DT 8	Human	2014	UK	Public Health England
H141620526	DT 193	Human	2014	UK	Public Health England
H141620528	DT 193	Human	2014	UK	Public Health England
H141620528	DT 193	Human	2014	UK	Public Health England
H141620529	DT 56	Human	2014	UK	Public Health England
H141620529	DT 56	Human	2014	UK	Public Health England
H141620533	DT 120	Human	2014	UK	Public Health England
H141640734	DT 104	Human	2014	UK	Public Health England
H141640739	None	Human	2014	UK	Public Health England
H141640739	None	Human	2014	UK	Public Health England
H141640740	None	Human	2014	UK	Public Health England
H141640740	None	Human	2014	UK	Public Health England
H141660473	DT 193	Human	2014	UK	Public Health England
H141660474	DT 104	Human	2014	UK	Public Health England
H141660475	PT U311	PORK - RAW	2014	UK	Public Health England
H141660476	DT 104	Human	2014	UK	Public Health England
H141660477	DT 120	Human	2014	UK	Public Health England
H141660478	None	Human	2014	UK	Public Health England

H141660479	DT 104	Human	2014	UK	Public Health England
H141660481	DT 193	PORK	2014	UK	Public Health England
H141660485	DT 193	RAW MEAT	2014	UK	Public Health England
H141741067	DT 104	Human	2014	UK	Public Health England
H141741091	DT 120	Human	2014	UK	Public Health England
H141741092	PT U323	Human	2014	UK	Public Health England
H141741094	None	Human	2014	UK	Public Health England
H141741095	None	Human	2014	UK	Public Health England
H141741114	DT 193	SOURCE UNKNOWN	2014	UK	Public Health England
H141741117	DT 193	SOURCE UNKNOWN	2014	UK	Public Health England
H141741118	DT 104	SOURCE UNKNOWN	2014	UK	Public Health England
H141741121	DT 86	NCTC 5713B3	2014	UK	Public Health England
H141760554	DT 193	Human	2014	UK	Public Health England
H141760555	PT U313	Human	2014	UK	Public Health England
H141760556	DT 193	Human	2014	UK	Public Health England
H141760557	None	Human	2014	UK	Public Health England
H141760561	None	Human	2014	UK	Public Health England
H141780732	DT 21	RAW MEAT	2014	UK	Public Health England
H141780739	None	Human	2014	UK	Public Health England
H141780740	DT 195	Human	2014	UK	Public Health England
H141780741	DT 193	Human	2014	UK	Public Health England

H141800491	DT 193	Human	2014	UK	Public Health England
H141800493	None	Human	2014	UK	Public Health England
H141800494	None	Human	2014	UK	Public Health England
H141820386	DT 193	Human	2014	UK	Public Health England
H141820387	DT 193	Human	2014	UK	Public Health England
H141820388	DT 193	Human	2014	UK	Public Health England
H141820389	DT 193	Human	2014	UK	Public Health England
H141820390	DT 104	Human	2014	UK	Public Health England
H141820391	PT U314	Human	2014	UK	Public Health England
H141820392	DT 193	Human	2014	UK	Public Health England
H141820393	PT U288	Human	2014	UK	Public Health England
H141820410	DT 193	Human	2014	UK	Public Health England
H141840455	PT U312	Human	2014	UK	Public Health England
H141840456	None	Human	2014	UK	Public Health England
H141840457	PT U312	Human	2014	UK	Public Health England
H141840458	DT 104	Human	2014	UK	Public Health England
H141840462	DT 193	Human	2014	UK	Public Health England
H141840463	DT 1	EQA	2014	UK	Public Health England
H141840464	DT 21	RAW MEAT	2014	UK	Public Health England
H141840475	None	Human	2014	UK	Public Health England
H141840480	DT 120	Human	2014	UK	Public Health England
H141840482	DT 120	Human	2014	UK	Public Health England

H141860444	RDNC	Human	2014	UK	Public Health England
H141860462	DT 120	Human	2014	UK	Public Health England
H141880251	PT U312	Human	2014	UK	Public Health England
H141880252	DT 193	Human	2014	UK	Public Health England
H141880253	PT U317	Human	2014	UK	Public Health England
H141880254	DT 120	UNKNOWN SOURCE (NON HUMAN)	2014	UK	Public Health England
H141880268	DT 193	Human	2014	UK	Public Health England
H141920581	DT 135	Human	2014	UK	Public Health England
H141920582	DT 193	Human	2014	UK	Public Health England
H141920583	DT 104	Human	2014	UK	Public Health England
H141920584	DT 22	Human	2014	UK	Public Health England
H141920596	None	Human	2014	UK	Public Health England
H141940684	None	Human	2014	UK	Public Health England
H141940696	None	Human	2014	UK	Public Health England
H141940697	DT 104	Human	2014	UK	Public Health England
H141940698	RDNC	Human	2014	UK	Public Health England
H141940699	DT 193	Human	2014	UK	Public Health England
H141940700	DT 104	Human	2014	UK	Public Health England
H141940702	DT 132	Human	2014	UK	Public Health England
H141940704	DT 170	Human	2014	UK	Public Health England
H141940705	DT 193	Human	2014	UK	Public Health England

H141960658	None	Human	2014	UK	Public Health England
H141960659	DT 120	Human	2014	UK	Public Health England
H141960666	PT U323	Human	2014	UK	Public Health England
H141960674	DT 7	Human	2014	UK	Public Health England
H141960675	DT 193	Human	2014	UK	Public Health England
H141960676	DT 193	Human	2014	UK	Public Health England
H141980453	DT 193	SOURCE UNKNOWN	2014	UK	Public Health England
H141980458	DT 104	NCTC AMPOULE	2014	UK	Public Health England
H141980465	DT 74	Human	2014	UK	Public Health England
H141980466	DT 1	Human	2014	UK	Public Health England
H141980469	DT 193	Human	2014	UK	Public Health England
H142000598	None	Human	2014	UK	Public Health England
H142000599	DT 132	Human	2014	UK	Public Health England
H142000600	DT 1	Human	2014	UK	Public Health England
H142020343	DT 193	Human	2014	UK	Public Health England
H142020344	DT 41	Human	2014	UK	Public Health England
H142020349	None	Human	2014	UK	Public Health England
H142020351	None	Human	2014	UK	Public Health England
H142020418	DT 193	Human	2014	UK	Public Health England
H142020419	DT 193	Human	2014	UK	Public Health England
H142020420	RDNC	Human	2014	UK	Public Health England

H142020441	DT 193	Human	2014	UK	Public Health England
H142040424	DT 104	Human	2014	UK	Public Health England
H142040426	DT 101	Human	2014	UK	Public Health England
H142040436	DT 104	Human	2014	UK	Public Health England
H142040442	DT 120	Human	2014	UK	Public Health England
H142040445	RDNC	Human	2014	UK	Public Health England
H142040455	None	Human	2014	UK	Public Health England
H142060410	DT 193	Human	2014	UK	Public Health England
H142060412	PT U277	Human	2014	UK	Public Health England
H142060415	DT 193	Human	2014	UK	Public Health England
H142080647	DT 193	Human	2014	UK	Public Health England
H142080648	None	Human	2014	UK	Public Health England
H142100254	DT 193	Human	2014	UK	Public Health England
H142100255	DT 170	Human	2014	UK	Public Health England
H142100256	PT U302	Human	2014	UK	Public Health England
H142100257	DT 193	Human	2014	UK	Public Health England
H142100258	DT 41	Human	2014	UK	Public Health England
H142100259	DT 29	Human	2014	UK	Public Health England
H142120282	DT 64	Human	2014	UK	Public Health England
H142120283	DT 193	Human	2014	UK	Public Health England
H142140689	DT 103	Human	2014	UK	Public Health England
H142140691	DT 1	Human	2014	UK	Public Health England

H142140696	None	Human	2014	UK	Public Health England
H142140702	DT 193	Human	2014	UK	Public Health England
H142160283	DT 120	Human	2014	UK	Public Health England
H142160284	DT 104	Human	2014	UK	Public Health England
H142160286	None	Human	2014	UK	Public Health England
H142160292	DT 193	RAW MEAT	2014	UK	Public Health England
H142160293	DT 193	Human	2014	UK	Public Health England
H142160295	DT 120	Human	2014	UK	Public Health England
H142180411	DT 193	Human	2014	UK	Public Health England
H142180412	DT 193	Human	2014	UK	Public Health England
H142180413	DT 193	Human	2014	UK	Public Health England
H142240388	PT U310	Human	2014	UK	Public Health England
H142240391	DT 193	Human	2014	UK	Public Health England
H142240393	DT 101	Human	2014	UK	Public Health England
H142240394	DT 161	Human	2014	UK	Public Health England
H142240396	DT 193	Human	2014	UK	Public Health England
H142240397	DT 22	Human	2014	UK	Public Health England
H142240472	DT 161	Human	2014	UK	Public Health England
H142260380	DT 15	Human	2014	UK	Public Health England
H142260381	DT 193	Human	2014	UK	Public Health England
H142260381	DT 193	Human	2014	UK	Public Health England
H142260451	DT 193	Human	2014	UK	Public Health England

H142280446	DT 193	Human	2014	UK	Public Health England
H142300580	DT 104	Human	2014	UK	Public Health England
H142300582	DT 193	Human	2014	UK	Public Health England
H142320339	DT 120	Human	2014	UK	Public Health England
H142320344	RDNC	Human	2014	UK	Public Health England
H142320354	DT 193	Human	2014	UK	Public Health England
H142320355	DT 193	Human	2014	UK	Public Health England
H142320356	DT 120	Human	2014	UK	Public Health England
H142320357	DT 193	Human	2014	UK	Public Health England
H142320358	DT 104	Human	2014	UK	Public Health England
H142320411	PT U310	Human	2014	UK	Public Health England
H142340624	DT 2	Human	2014	UK	Public Health England
H142340625	DT 193	Human	2014	UK	Public Health England
H142340626	DT 193	Human	2014	UK	Public Health England
H142340627	DT 193	Human	2014	UK	Public Health England
H142340629	DT 73	Human	2014	UK	Public Health England
H142360379	DT 8	Human	2014	UK	Public Health England
H142360387	DT 120	Human	2014	UK	Public Health England
H142360389	DT 120	Human	2014	UK	Public Health England
H142380312	PT U311	Human	2014	UK	Public Health England
H142380313	PT U323	Human	2014	UK	Public Health England
H142380314	DT 2	Human	2014	UK	Public Health England

H142380315	DT 193	Human	2014	UK	Public Health England
H142380316	DT 193	Human	2014	UK	Public Health England
H142380317	DT 104	Human	2014	UK	Public Health England
H142380318	DT 2	Human	2014	UK	Public Health England
H142380320	DT 193	Human	2014	UK	Public Health England
H142400424	DT 191	Human	2014	UK	Public Health England
H142420764	DT 193	Human	2014	UK	Public Health England
H142420765	DT 191	Human	2014	UK	Public Health England
H142420766	RDNC	Human	2014	UK	Public Health England
H142420768	DT 193	Human	2014	UK	Public Health England
H142420769	DT 193	Human	2014	UK	Public Health England
H142420807	RDNC	Human	2014	UK	Public Health England
H142420808	None	Human	2014	UK	Public Health England
H142420821	None	Human	2014	UK	Public Health England
H142420822	DT 193	Human	2014	UK	Public Health England
H142440682	PT U302	Human	2014	UK	Public Health England
H142440683	DT 193	Human	2014	UK	Public Health England
H142440684	DT 177	Human	2014	UK	Public Health England
H142440685	DT 193	Human	2014	UK	Public Health England
H142440703	DT 8	Human	2014	UK	Public Health England
H142440705	DT 193	Human	2014	UK	Public Health England
H142440707	None	Human	2014	UK	Public Health England

H142460293	DT 94	Human	2014	UK	Public Health England
H142500094	DT 104	Human	2014	UK	Public Health England
H142500101	DT 193	Human	2014	UK	Public Health England
H142520441	DT 203	Human	2014	UK	Public Health England
H142520442	DT 104	Human	2014	UK	Public Health England
H142520443	PT U302	Human	2014	UK	Public Health England
H142520444	PT U323	Human	2014	UK	Public Health England
H142520827	DT 193	Human	2014	UK	Public Health England
H142540301	DT 193	Human	2014	UK	Public Health England
H142540304	DT 49	Human	2014	UK	Public Health England
H142540308	DT 2	Human	2014	UK	Public Health England
H142540309	DT 193	Human	2014	UK	Public Health England
H142540310	DT 104	Human	2014	UK	Public Health England
H142540390	DT 193	Human	2014	UK	Public Health England
H142540395	DT 8	Human	2014	UK	Public Health England
H142540400	DT 178	Human	2014	UK	Public Health England
H142560489	DT 193	Human	2014	UK	Public Health England
H142560491	DT 193	Human	2014	UK	Public Health England
H142560538	DT 1	Human	2014	UK	Public Health England
H142580387	DT 120	Human	2014	UK	Public Health England
H142600609	DT 120	Human	2014	UK	Public Health England
H142600655	DT 104	Human	2014	UK	Public Health England

H142600656	DT 46	Human	2014	UK	Public Health England
H142600659	DT 193	Human	2014	UK	Public Health England
H142600660	DT 193	Human	2014	UK	Public Health England
H142600661	DT 193	Human	2014	UK	Public Health England
H142600665	DT 195	Human	2014	UK	Public Health England
H142620391	DT 2	Human	2014	UK	Public Health England
H142640381	DT 193	Human	2014	UK	Public Health England
H142640382	DT 193	Human	2014	UK	Public Health England
H142640383	DT 191	Human	2014	UK	Public Health England
H142640384	DT 193	Human	2014	UK	Public Health England
H142640385	PT U302	Human	2014	UK	Public Health England
H142640386	None	Human	2014	UK	Public Health England
H142640387	None	Human	2014	UK	Public Health England
H142640394	DT 193	Human	2014	UK	Public Health England
H142660316	DT 193	Human	2014	UK	Public Health England
H142680476	DT 193	Human	2014	UK	Public Health England
H142680477	PT U311	Human	2014	UK	Public Health England
H142680479	RDNC	NOT STATED	2014	UK	Public Health England
H142680481	DT 193	Human	2014	UK	Public Health England
H142700048	DT 104	Human	2014	UK	Public Health England
H142720293	PT U311	Human	2014	UK	Public Health England
H142720294	RDNC	Human	2014	UK	Public Health England

H142720295	PT U314	Human	2014	UK	Public Health England
H142720297	DT 193	Human	2014	UK	Public Health England
H142720298	DT 193	Human	2014	UK	Public Health England
H142720299	DT 193	Human	2014	UK	Public Health England
H142740617	DT 193	Human	2014	UK	Public Health England
H142740622	DT 193	Human	2014	UK	Public Health England
H142740624	DT 193	Human	2014	UK	Public Health England
H142740625	DT 193	Human	2014	UK	Public Health England
H142740626	PT U302	Human	2014	UK	Public Health England
H142740627	DT 104	Human	2014	UK	Public Health England
H142740628	DT 99	Human	2014	UK	Public Health England
H142740641	DT 56	Human	2014	UK	Public Health England
H142740643	DT 177	Human	2014	UK	Public Health England
H142760493	DT 132	Human	2014	UK	Public Health England
H142760495	PT U323	Human	2014	UK	Public Health England
H142760496	None	Human	2014	UK	Public Health England
H142760497	DT 193	Human	2014	UK	Public Health England
H142760499	DT 193	Human	2014	UK	Public Health England
H142760550	DT 193	Human	2014	UK	Public Health England
H142780347	None	Human	2014	UK	Public Health England
H142780372	DT 40	Human	2014	UK	Public Health England
H142780373	DT 35	Human	2014	UK	Public Health England

H142780374	DT 104	Human	2014	UK	Public Health England
H142780376	DT 193	Human	2014	UK	Public Health England
H142800188	DT 193	Human	2014	UK	Public Health England
H142800189	DT 193	Human	2014	UK	Public Health England
H142800190	DT 17	Human	2014	UK	Public Health England
H142800191	None	Human	2014	UK	Public Health England
H142800209	Untypable	Human	2014	UK	Public Health England
H142820396	None	Human	2014	UK	Public Health England
H142820412	DT 8	Human	2014	UK	Public Health England
H142820418	DT 193	Human	2014	UK	Public Health England
H142820420	DT 135	Human	2014	UK	Public Health England
H142820422	DT 193	Human	2014	UK	Public Health England
H142820424	DT 135	Human	2014	UK	Public Health England
H142820426	DT 193	Human	2014	UK	Public Health England
H142820427	DT 120	Human	2014	UK	Public Health England
H142820428	DT 101	Human	2014	UK	Public Health England
H142820429	DT 101	Human	2014	UK	Public Health England
H142820431	None	Human	2014	UK	Public Health England
H142820468	DT 193	Human	2014	UK	Public Health England
H142840476	None	Human	2014	UK	Public Health England
H142840477	DT 69	Human	2014	UK	Public Health England
H142840479	DT 101	Human	2014	UK	Public Health England

H142840480	DT 193	Human	2014	UK	Public Health England
H142840488	DT 193	Human	2014	UK	Public Health England
H142860385	DT 193	Human	2014	UK	Public Health England
H142860396	DT 135	Human	2014	UK	Public Health England
H142860397	DT 193	Human	2014	UK	Public Health England
H142860403	DT 193	Human	2014	UK	Public Health England
H142880242	DT 193	Human	2014	UK	Public Health England
H142880244	PT U310	Human	2014	UK	Public Health England
H142880245	None	Human	2014	UK	Public Health England
H142880246	None	Human	2014	UK	Public Health England
H142880253	DT 104	Human	2014	UK	Public Health England
H142900379	DT 104	Human	2014	UK	Public Health England
H142900475	PT U311	Human	2014	UK	Public Health England
H142900476	DT 193	Human	2014	UK	Public Health England
H142920569	DT 135	Human	2014	UK	Public Health England
H142920572	PT U310	Human	2014	UK	Public Health England
H142920573	DT 49	Human	2014	UK	Public Health England
H142920575	DT 1	Human	2014	UK	Public Health England
H142920576	DT 193	Human	2014	UK	Public Health England
H142920583	PT U288	Human	2014	UK	Public Health England
H142920585	DT 193	Human	2014	UK	Public Health England
H142920602	PT U310	Human	2014	UK	Public Health England

H142920606	DT 104	Human	2014	UK	Public Health England
H142920607	DT 8	Human	2014	UK	Public Health England
H142920608	DT 1	Human	2014	UK	Public Health England
H142920612	DT 195	Human	2014	UK	Public Health England
H142940532	DT 193	Human	2014	UK	Public Health England
H142940534	DT 193	Human	2014	UK	Public Health England
H142940536	None	Human	2014	UK	Public Health England
H142940545	DT 1	Human	2014	UK	Public Health England
H142940549	DT 1	Human	2014	UK	Public Health England
H142940551	DT 193	Human	2014	UK	Public Health England
H142940552	PT U288	Human	2014	UK	Public Health England
H142940554	PT U311	Human	2014	UK	Public Health England
H142940555	PT U311	Human	2014	UK	Public Health England
H142940556	PT U311	Human	2014	UK	Public Health England
H142940557	PT U311	Human	2014	UK	Public Health England
H142940559	PT U311	Human	2014	UK	Public Health England
H142940577	DT 104	Human	2014	UK	Public Health England
H142940579	DT 193	Human	2014	UK	Public Health England
H142960507	DT 74	Human	2014	UK	Public Health England
H142960508	DT 193	Human	2014	UK	Public Health England
H142960510	PT U312	Human	2014	UK	Public Health England
H142960521	DT 193	Human	2014	UK	Public Health England

H142960527	DT 2	Human	2014	UK	Public Health England
H142980299	None	Human	2014	UK	Public Health England
H142980300	DT 193	Human	2014	UK	Public Health England
H142980300	DT 193	Human	2014	UK	Public Health England
H142980302	DT 193	Human	2014	UK	Public Health England
H143000283	PT U288	Human	2014	UK	Public Health England
H143000285	DT 193	Human	2014	UK	Public Health England
H143000286	None	Human	2014	UK	Public Health England
H143000286	None	Human	2014	UK	Public Health England
H143000295	DT 193	Human	2014	UK	Public Health England
H143000297	DT 193	Human	2014	UK	Public Health England
H143020367	PT U302	Human	2014	UK	Public Health England
H143020368	DT 193	Human	2014	UK	Public Health England
H143020369	DT 22	Human	2014	UK	Public Health England
H143020372	DT 193	Human	2014	UK	Public Health England
H143020373	DT 193	Human	2014	UK	Public Health England
H143020374	PT U288	Human	2014	UK	Public Health England
H143020408	DT 193	Human	2014	UK	Public Health England
H143040502	PT U288	Human	2014	UK	Public Health England
H143040504	DT 193	Human	2014	UK	Public Health England
H143040505	DT 193	Human	2014	UK	Public Health England
H143040507	PT U311	Human	2014	UK	Public Health England

H143040509	DT 195	Human	2014	UK	Public Health England
H143040511	DT 193	Human	2014	UK	Public Health England
H143040550	DT 101	Human	2014	UK	Public Health England
H143060181	PT U311	Human	2014	UK	Public Health England
H143060187	DT 205	Human	2014	UK	Public Health England
H143060188	DT 193	Human	2014	UK	Public Health England
H143060189	DT 193	Human	2014	UK	Public Health England
H143060191	DT 193	Human	2014	UK	Public Health England
H143060192	PT U302	Human	2014	UK	Public Health England
H143060271	PT U311	Human	2014	UK	Public Health England
H143080507	DT 193	Human	2014	UK	Public Health England
H143080514	DT 193	Human	2014	UK	Public Health England
H143080520	DT 193	Human	2014	UK	Public Health England
H143080521	DT 104	Human	2014	UK	Public Health England
H143080522	RDNC	Human	2014	UK	Public Health England
H143100440	RDNC	Human	2014	UK	Public Health England
H143100441	RDNC	Human	2014	UK	Public Health England
H143100443	DT 193	Human	2014	UK	Public Health England
H143100447	DT 193	Human	2014	UK	Public Health England
H143100449	PT U276	Human	2014	UK	Public Health England
H143100450	DT 193	Human	2014	UK	Public Health England
H143120411	PT U311	Human	2014	UK	Public Health England

H143120415	DT 193	Human	2014	UK	Public Health England
H143120421	DT 2	Human	2014	UK	Public Health England
H143120429	DT 40	Human	2014	UK	Public Health England
H143120432	DT 193	Human	2014	UK	Public Health England
H143120433	DT 193	Human	2014	UK	Public Health England
H143120493	RDNC	Human	2014	UK	Public Health England
H143120493	RDNC	Human	2014	UK	Public Health England
H143140475	DT 193	Human	2014	UK	Public Health England
H143140481	DT 193	Human	2014	UK	Public Health England
H143140483	DT 120	Human	2014	UK	Public Health England
H143140486	DT 193	Human	2014	UK	Public Health England
H143140487	DT 120	Human	2014	UK	Public Health England
H143140488	DT 9	Human	2014	UK	Public Health England
H143140489	DT 49	Human	2014	UK	Public Health England
H143140492	DT 193	Human	2014	UK	Public Health England
H143140492	DT 193	Human	2014	UK	Public Health England
H143140493	DT 193	Human	2014	UK	Public Health England
H143140578	None	Human	2014	UK	Public Health England
H143140582	DT 42	Human	2014	UK	Public Health England
H143140589	DT 193	Human	2014	UK	Public Health England
H143140591	DT 43	Human	2014	UK	Public Health England
H143160363	DT 191	Human	2014	UK	Public Health England

H143160439	DT 193	Human	2014	UK	Public Health England
H143160444	DT 104	Human	2014	UK	Public Health England
H143160444	DT 104	Human	2014	UK	Public Health England
H143180347	PT U311	Human	2014	UK	Public Health England
H143180352	PT U288	Human	2014	UK	Public Health England
H143180353	DT 41	Human	2014	UK	Public Health England
H143180355	DT 193	Human	2014	UK	Public Health England
H143180357	DT 193	Human	2014	UK	Public Health England
H143180357	DT 193	Human	2014	UK	Public Health England
H143200210	DT 104	Human	2014	UK	Public Health England
H143200215	DT 193	Human	2014	UK	Public Health England
H143200218	DT 193	Human	2014	UK	Public Health England
H143200580	PT U311	Human	2014	UK	Public Health England
H143200611	DT 104	Human	2014	UK	Public Health England
H143200618	DT 8	Human	2014	UK	Public Health England
H143220671	DT 193	Human	2014	UK	Public Health England
H143220672	DT 8	Human	2014	UK	Public Health England
H143220674	None	Human	2014	UK	Public Health England
H143220678	DT 193	Human	2014	UK	Public Health England
H143220686	RDNC	Human	2014	UK	Public Health England
H143220688	DT 120	Human	2014	UK	Public Health England
H143220689	DT 193	Human	2014	UK	Public Health England

H143220691	DT 193	Human	2014	UK	Public Health England
H143220692	DT 104	Human	2014	UK	Public Health England
H143220695	DT 1	Human	2014	UK	Public Health England
H143220698	DT 193	Human	2014	UK	Public Health England
H143220699	PT U311	Human	2014	UK	Public Health England
H143220700	PT U311	Human	2014	UK	Public Health England
H143220701	DT 7	CATTLE	2014	UK	Public Health England
H143220702	RDNC	CATTLE	2014	UK	Public Health England
H143220707	DT 193	Human	2014	UK	Public Health England
H143220792	DT 193	Human	2014	UK	Public Health England
H143220793	DT 104	Human	2014	UK	Public Health England
H143240626	DT 193	Human	2014	UK	Public Health England
H143240631	DT 8	Human	2014	UK	Public Health England
H143240635	PT U312	Human	2014	UK	Public Health England
H143240637	PT U312	Human	2014	UK	Public Health England
H143240640	DT 193	FROZEN RAW CHICKEN	2014	UK	Public Health England
H143240642	PT U302	Human	2014	UK	Public Health England
H143240643	DT 193	Human	2014	UK	Public Health England
H143240644	DT 193	Human	2014	UK	Public Health England
H143240645	None	Human	2014	UK	Public Health England
H143240647	DT 120	Human	2014	UK	Public Health England
H143240650	RDNC	Human	2014	UK	Public Health England

H143240651	DT 101	Human	2014	UK	Public Health England
H143240683	DT 193	Human	2014	UK	Public Health England
H143240685	DT 2	Human	2014	UK	Public Health England
H143260616	DT 193	Human	2014	UK	Public Health England
H143260617	None	Human	2014	UK	Public Health England
H143260618	None	Human	2014	UK	Public Health England
H143260626	DT 193	Human	2014	UK	Public Health England
H143260627	DT 104	Human	2014	UK	Public Health England
H143260629	DT 193	Human	2014	UK	Public Health England
H143260631	DT 10	Human	2014	UK	Public Health England
H143260632	PT U276	Human	2014	UK	Public Health England
H143260633	DT 120	Human	2014	UK	Public Health England
H143280585	DT 2	Human	2014	UK	Public Health England
H143280585	DT 2	Human	2014	UK	Public Health England
H143280587	DT 101	Human	2014	UK	Public Health England
H143280603	DT 193	Human	2014	UK	Public Health England
H143280604	PT U302	Human	2014	UK	Public Health England
H143280605	DT 193	Human	2014	UK	Public Health England
H143280606	DT 104	Human	2014	UK	Public Health England
H143280607	DT 193	Human	2014	UK	Public Health England
H143280610	DT 193	Human	2014	UK	Public Health England
H143300400	DT 193	Human	2014	UK	Public Health England

H143300401	DT 193	Human	2014	UK	Public Health England
H143300402	PT U323	Human	2014	UK	Public Health England
H143300403	DT 12	Human	2014	UK	Public Health England
H143300404	DT 193	Human	2014	UK	Public Health England
H143300405	DT 120	Human	2014	UK	Public Health England
H143320328	None	Human	2014	UK	Public Health England
H143320447	DT 56	Human	2014	UK	Public Health England
H143320448	DT 193	Human	2014	UK	Public Health England
H143320465	DT 193	Human	2014	UK	Public Health England
H143320471	DT 104	Human	2014	UK	Public Health England
H143320472	RDNC	Human	2014	UK	Public Health England
H143320490	DT 4	Human	2014	UK	Public Health England
H143340483	None	Human	2014	UK	Public Health England
H143340495	DT 193	Human	2014	UK	Public Health England
H143340496	PT U331	Human	2014	UK	Public Health England
H143340500	None	Human	2014	UK	Public Health England
H143340502	DT 104	Human	2014	UK	Public Health England
H143340503	DT 193	Human	2014	UK	Public Health England
H143340505	DT 193	Human	2014	UK	Public Health England
H143340506	PT U329	Human	2014	UK	Public Health England
H143340515	None	Human	2014	UK	Public Health England
H143340517	DT 193	Human	2014	UK	Public Health England

H143340518	DT 104	Human	2014	UK	Public Health England
H143340519	DT 193	Human	2014	UK	Public Health England
H143340521	DT 8	Human	2014	UK	Public Health England
H143340522	DT 8	Human	2014	UK	Public Health England
H143360556	DT 193	Human	2014	UK	Public Health England
H143380360	DT 193	Human	2014	UK	Public Health England
H143380438	PT U323	Human	2014	UK	Public Health England
H143380440	DT 193	Human	2014	UK	Public Health England
H143380441	DT 7	Human	2014	UK	Public Health England
H143380442	None	Human	2014	UK	Public Health England
H143400432	PT U323	Human	2014	UK	Public Health England
H143400433	DT 193	Human	2014	UK	Public Health England
H143400434	PT U323	Human	2014	UK	Public Health England
H143400436	DT 195	Human	2014	UK	Public Health England
H143400440	DT 8	Human	2014	UK	Public Health England
H143400441	RDNC	Human	2014	UK	Public Health England
H143400442	None	Human	2014	UK	Public Health England
H143400481	DT 193	Human	2014	UK	Public Health England
H143420455	DT 135	Human	2014	UK	Public Health England
H143420586	None	Human	2014	UK	Public Health England
H143420588	None	Human	2014	UK	Public Health England
H143420589	None	Human	2014	UK	Public Health England

H143420590	DT 193	Human	2014	UK	Public Health England
H143420594	DT 193	Human	2014	UK	Public Health England
H143420595	DT 193	Human	2014	UK	Public Health England
H143420596	DT 101	Human	2014	UK	Public Health England
H143420602	DT 101	Human	2014	UK	Public Health England
H143420603	DT 193	Human	2014	UK	Public Health England
H143420605	DT 193	Human	2014	UK	Public Health England
H143440341	DT 193	Human	2014	UK	Public Health England
H143440347	None	Human	2014	UK	Public Health England
H143440350	DT 1	Human	2014	UK	Public Health England
H143440352	DT 208	Human	2014	UK	Public Health England
H143480449	PT U311	Human	2014	UK	Public Health England
H143520409	DT 193	Human	2014	UK	Public Health England
H143520413	DT 193	Human	2014	UK	Public Health England
H143520415	DT 193	Human	2014	UK	Public Health England
H143520418	DT 193	Human	2014	UK	Public Health England
H143520444	DT 49	Human	2014	UK	Public Health England
H143520446	RDNC	Human	2014	UK	Public Health England
H143520449	DT 120	Human	2014	UK	Public Health England
H143540763	DT 8	Human	2014	UK	Public Health England
H143540847	DT 193	Human	2014	UK	Public Health England
H143540860	DT 8	Human	2014	UK	Public Health England

H143540866	DT 120	Human	2014	UK	Public Health England
H143540867	DT 135	Human	2014	UK	Public Health England
H143540869	DT 193	Human	2014	UK	Public Health England
H143540869	DT 193	Human	2014	UK	Public Health England
H143540870	DT 193	Human	2014	UK	Public Health England
H143540872	DT 116	Human	2014	UK	Public Health England
H143540873	DT 135	Human	2014	UK	Public Health England
H143540874	DT 193	Human	2014	UK	Public Health England
H143540875	DT 177	Human	2014	UK	Public Health England
H143540876	DT 56	CAT	2014	UK	Public Health England
H143540877	PT U308	CANARY	2014	UK	Public Health England
H143540878	PT U308	PIG	2014	UK	Public Health England
H143540879	PT U323	Human	2014	UK	Public Health England
H143540881	DT 193	Human	2014	UK	Public Health England
H143540882	DT 193	Human	2014	UK	Public Health England
H143540883	DT 193	Human	2014	UK	Public Health England
H143540884	PT U302	Human	2014	UK	Public Health England
H143540885	DT 135	Human	2014	UK	Public Health England
H143540899	PT U312	Human	2014	UK	Public Health England
H143540899	PT U312	Human	2014	UK	Public Health England
H143560431	DT 193	Human	2014	UK	Public Health England
H143560436	DT 193	Human	2014	UK	Public Health England

H143560439	DT 193	Human	2014	UK	Public Health England
H143560440	None	Human	2014	UK	Public Health England
H143560441	DT 193	Human	2014	UK	Public Health England
H143560443	PT U311	Human	2014	UK	Public Health England
H143560444	PT U323	Human	2014	UK	Public Health England
H143580507	None	Human	2014	UK	Public Health England
H143580513	DT 1	Human	2014	UK	Public Health England
H143580522	DT 8	Human	2014	UK	Public Health England
H143580525	None	Human	2014	UK	Public Health England
H143580526	RDNC	Human	2014	UK	Public Health England
H143580530	None	Human	2014	UK	Public Health England
H143580533	DT 193	Human	2014	UK	Public Health England
H143580535	DT 193	Human	2014	UK	Public Health England
H143580536	DT 99	Human	2014	UK	Public Health England
H143580538	PT U311	Human	2014	UK	Public Health England
H143580621	PT U314	Human	2014	UK	Public Health England
H143600533	DT 193	Human	2014	UK	Public Health England
H143600534	DT 120	Human	2014	UK	Public Health England
H143600536	DT 193	Human	2014	UK	Public Health England
H143600539	DT 104	Human	2014	UK	Public Health England
H143600540	PT U311	Human	2014	UK	Public Health England
H143600541	PT U311	Human	2014	UK	Public Health England

H143600542	DT 193	Human	2014	UK	Public Health England
H143600543	PT U311	Human	2014	UK	Public Health England
H143600544	RDNC	Human	2014	UK	Public Health England
H143600562	DT 193	Human	2014	UK	Public Health England
H143620791	PT U311	Human	2014	UK	Public Health England
H143620791	PT U311	Human	2014	UK	Public Health England
H143620793	DT 193	Human	2014	UK	Public Health England
H143620794	DT 193	Human	2014	UK	Public Health England
H143620798	DT 120	Human	2014	UK	Public Health England
H143620798	DT 120	Human	2014	UK	Public Health England
H143640556	DT 193	Human	2014	UK	Public Health England
H143640578	DT 104	Human	2014	UK	Public Health England
H143640578	DT 104	Human	2014	UK	Public Health England
H143640580	DT 193	Human	2014	UK	Public Health England
H143640580	DT 193	Human	2014	UK	Public Health England
H143640580	DT 193	Human	2014	UK	Public Health England
H143640583	DT 193	Human	2014	UK	Public Health England
H143640584	None	Human	2014	UK	Public Health England
H143640585	None	Human	2014	UK	Public Health England
H143640585	None	Human	2014	UK	Public Health England
H143660447	DT 193	Human	2014	UK	Public Health England
H143660449	DT 193	Human	2014	UK	Public Health England

H143660450	DT 120	Human	2014	UK	Public Health England
H143660462	DT 193	Human	2014	UK	Public Health England
H143660470	DT 193	Human	2014	UK	Public Health England
H143660473	DT 193	Human	2014	UK	Public Health England
H143660478	DT 135	Human	2014	UK	Public Health England
H143660478	DT 135	Human	2014	UK	Public Health England
H143660479	PT U323	Human	2014	UK	Public Health England
H143660479	PT U323	Human	2014	UK	Public Health England
H143660507	DT 193	Human	2014	UK	Public Health England
H143660508	DT 120	Human	2014	UK	Public Health England
H143680693	DT 120	Human	2014	UK	Public Health England
H143680693	DT 120	Human	2014	UK	Public Health England
H143680695	DT 193	Human	2014	UK	Public Health England
H143680695	DT 193	Human	2014	UK	Public Health England
H143680696	DT 193	Human	2014	UK	Public Health England
H143680696	DT 193	Human	2014	UK	Public Health England
H143700616	DT 101	Human	2014	UK	Public Health England
H143700617	DT 193	Human	2014	UK	Public Health England
H143700618	DT 193	Human	2014	UK	Public Health England
H143700633	PT U311	Human	2014	UK	Public Health England
H143720723	PT U312	Human	2014	UK	Public Health England
H143720725	DT 104	Human	2014	UK	Public Health England

H143720758	DT 193	Human	2014	UK	Public Health England
H143720760	DT 2	Human	2014	UK	Public Health England
H143720761	DT 193	Human	2014	UK	Public Health England
H143720762	DT 193	Human	2014	UK	Public Health England
H143720763	None	Human	2014	UK	Public Health England
H143720765	DT 193	Human	2014	UK	Public Health England
H143720766	DT 193	Human	2014	UK	Public Health England
H143720768	DT 193	Human	2014	UK	Public Health England
H143720769	DT 193	Human	2014	UK	Public Health England
H143740686	DT 81	Human	2014	UK	Public Health England
H143740703	None	Human	2014	UK	Public Health England
H143740704	DT 193	PIG	2014	UK	Public Health England
H143740705	DT 193	PIG	2014	UK	Public Health England
H143740708	DT 120	Human	2014	UK	Public Health England
H143740710	DT 193	Human	2014	UK	Public Health England
H143740713	PT U311	Human	2014	UK	Public Health England
H143740716	DT 44	Human	2014	UK	Public Health England
H143740718	DT 193	Human	2014	UK	Public Health England
H143740719	PT U311	Human	2014	UK	Public Health England
H143740720	PT U311	Human	2014	UK	Public Health England
H143740721	DT 80	Human	2014	UK	Public Health England
H143740722	DT 193	Human	2014	UK	Public Health England

H143740724	DT 193	Human	2014	UK	Public Health England
H143740725	DT 99	Human	2014	UK	Public Health England
H143740729	DT 193	SOURCE UNKNOWN	2014	UK	Public Health England
H143760527	DT 193	Human	2014	UK	Public Health England
H143760573	Untypable	Human	2014	UK	Public Health England
H143760577	DT 41	Human	2014	UK	Public Health England
H143760578	None	Human	2014	UK	Public Health England
H143760579	DT 193	Human	2014	UK	Public Health England
H143760580	DT 193	Human	2014	UK	Public Health England
H143760581	DT 193	Human	2014	UK	Public Health England
H143760583	None	Human	2014	UK	Public Health England
H143760585	DT 8	Human	2014	UK	Public Health England
H143760590	DT 193	Human	2014	UK	Public Health England
H143760594	DT 8	Human	2014	UK	Public Health England
H143780536	DT 193	Human	2014	UK	Public Health England
H143780537	PT U302	Human	2014	UK	Public Health England
H143780538	DT 80	Human	2014	UK	Public Health England
H143780539	DT 120	Human	2014	UK	Public Health England
H143780540	DT 193	Human	2014	UK	Public Health England
H143780543	DT 1	Human	2014	UK	Public Health England
H143780564	DT 110	Human	2014	UK	Public Health England
H143800497	DT 193	Human	2014	UK	Public Health England

H143800498	PT U291	Human	2014	UK	Public Health England
H143800499	DT 46	Human	2014	UK	Public Health England
H143800500	DT 193	Human	2014	UK	Public Health England
H143800502	DT 193	Human	2014	UK	Public Health England
H143800504	DT 46	Human	2014	UK	Public Health England
H143820790	DT 8	Human	2014	UK	Public Health England
H143820791	DT 193	Human	2014	UK	Public Health England
H143820792	PT U323	Human	2014	UK	Public Health England
H143820794	PT U312	Human	2014	UK	Public Health England
H143820795	DT 104	Human	2014	UK	Public Health England
H143820797	DT 193	Human	2014	UK	Public Health England
H143820842	Untypable	Human	2014	UK	Public Health England
H143840498	PT U277	Human	2014	UK	Public Health England
H143840570	DT 46	Human	2014	UK	Public Health England
H143840571	DT 80	Human	2014	UK	Public Health England
H143840575	None	Human	2014	UK	Public Health England
H143840576	PT U302	Human	2014	UK	Public Health England
H143860639	DT 195	Human	2014	UK	Public Health England
H143860655	PT U312	Human	2014	UK	Public Health England
H143860656	PT U311	Human	2014	UK	Public Health England
H143860657	DT 120	Human	2014	UK	Public Health England
H143860658	DT 120	Human	2014	UK	Public Health England

H143880696	PT U311	Human	2014	UK	Public Health England
H143880701	PT U277	Human	2014	UK	Public Health England
H143880703	DT 193	Human	2014	UK	Public Health England
H143880705	DT 120	Human	2014	UK	Public Health England
H143880724	DT 193	PIG	2014	UK	Public Health England
H143880725	DT 193	PIG	2014	UK	Public Health England
H143880726	DT 193	PIG	2014	UK	Public Health England
H143900512	DT 193	Human	2014	UK	Public Health England
H143900570	DT 193	Human	2014	UK	Public Health England
H143900571	DT 193	Human	2014	UK	Public Health England
H143900574	PT U308	Human	2014	UK	Public Health England
H143920789	DT 193	Human	2014	UK	Public Health England
H143920792	RDNC	Human	2014	UK	Public Health England
H143920793	DT 7	Human	2014	UK	Public Health England
H143920796	DT 104	Human	2014	UK	Public Health England
H143920797	PT U302	Human	2014	UK	Public Health England
H143920800	PT U311	Human	2014	UK	Public Health England
H143920801	PT U311	Human	2014	UK	Public Health England
H143920802	DT 193	Human	2014	UK	Public Health England
H143920803	DT 101	Human	2014	UK	Public Health England
H143940533	DT 193	Human	2014	UK	Public Health England
H143940609	DT 193	Human	2014	UK	Public Health England

H143940629	DT 193	Human	2014	UK	Public Health England
H143940630	DT 8	Human	2014	UK	Public Health England
H143940635	DT 193	Human	2014	UK	Public Health England
H143940637	DT 193	Human	2014	UK	Public Health England
H143940638	DT 120	Human	2014	UK	Public Health England
H143940639	DT 193	Human	2014	UK	Public Health England
H143960298	PT U323	Human	2014	UK	Public Health England
H143960356	DT 104	Human	2014	UK	Public Health England
H143960357	DT 193	Human	2014	UK	Public Health England
H143960359	PT U323	Human	2014	UK	Public Health England
H143960360	PT U302	Human	2014	UK	Public Health England
H143960363	DT 193	Human	2014	UK	Public Health England
H143960367	PT U312	Human	2014	UK	Public Health England
H143980708	DT 193	Human	2014	UK	Public Health England
H143980710	DT 193	Human	2014	UK	Public Health England
H143980721	DT 193	NECK FLAP	2014	UK	Public Health England
H143980724	DT 193	NECK FLAP	2014	UK	Public Health England
H143980725	DT 193	NECK FLAP	2014	UK	Public Health England
H143980726	DT 193	NECK FLAP	2014	UK	Public Health England
H143980747	DT 120	Human	2014	UK	Public Health England
H143980748	PT U302	Human	2014	UK	Public Health England
H144000584	DT 193	Human	2014	UK	Public Health England

H144000589	DT 101	Human	2014	UK	Public Health England
H144000591	DT 2	Human	2014	UK	Public Health England
H144000593	DT 1	Human	2014	UK	Public Health England
H144000594	DT 120	Human	2014	UK	Public Health England
H144020656	DT 195	Human	2014	UK	Public Health England
H144020667	DT 193	Human	2014	UK	Public Health England
H144020673	RDNC	Human	2014	UK	Public Health England
H144020708	PT U302	Human	2014	UK	Public Health England
H144020710	DT 193	Human	2014	UK	Public Health England
H144020711	DT 193	Human	2014	UK	Public Health England
H144020714	None	Human	2014	UK	Public Health England
H144020715	DT 104	Human	2014	UK	Public Health England
H144020716	DT 193	Human	2014	UK	Public Health England
H144020717	DT 104	Human	2014	UK	Public Health England
H144020720	PT U323	Human	2014	UK	Public Health England
H144020722	DT 101	Human	2014	UK	Public Health England
H144040605	DT 8	60.23504	2014	UK	Public Health England
H144040606	DT 8	60.23503	2014	UK	Public Health England
H144040612	DT 104	Human	2014	UK	Public Health England
H144040613	DT 193	Human	2014	UK	Public Health England
H144040615	DT 193	Human	2014	UK	Public Health England
H144040617	DT 193	Human	2014	UK	Public Health England

H144040620	DT 104	Human	2014	UK	Public Health England
H144040621	DT 104	Human	2014	UK	Public Health England
H144040622	DT 193	Human	2014	UK	Public Health England
H144060558	PT U302	Human	2014	UK	Public Health England
H144060572	PT U302	Human	2014	UK	Public Health England
H144060575	DT 104	Human	2014	UK	Public Health England
H144060576	PT U302	Human	2014	UK	Public Health England
H144060577	PT U323	Human	2014	UK	Public Health England
H144080387	DT 104	Human	2014	UK	Public Health England
H144080426	None	Human	2014	UK	Public Health England
H144080428	None	Human	2014	UK	Public Health England
H144080429	None	Human	2014	UK	Public Health England
H144080430	DT 193	Human	2014	UK	Public Health England
H144080431	DT 193	Human	2014	UK	Public Health England
H144100484	DT 104	Human	2014	UK	Public Health England
H144100490	DT 161	Human	2014	UK	Public Health England
H144120507	DT 193	Human	2014	UK	Public Health England
H144120530	DT 191	Human	2014	UK	Public Health England
H144120537	DT 193	Human	2014	UK	Public Health England
H144120538	None	Human	2014	UK	Public Health England
H144120539	PT U323	Human	2014	UK	Public Health England
H144120542	DT 41	Human	2014	UK	Public Health England

H144120543	PT U302	Human	2014	UK	Public Health England
H144120556	DT 1	Human	2014	UK	Public Health England
H144120557	DT 104	Human	2014	UK	Public Health England
H144120558	PT U302	Human	2014	UK	Public Health England
H144120559	PT U323	Human	2014	UK	Public Health England
H144140610	DT 104	Human	2014	UK	Public Health England
H144140619	DT 193	Human	2014	UK	Public Health England
H144140627	PT U311	Human	2014	UK	Public Health England
H144140628	None	Human	2014	UK	Public Health England
H144140631	DT 195	Human	2014	UK	Public Health England
H144140632	DT 193	Human	2014	UK	Public Health England
H144140634	DT 104	Human	2014	UK	Public Health England
H144140636	PT U302	Human	2014	UK	Public Health England
H144140639	PT U323	Human	2014	UK	Public Health England
H144160425	DT 15	Human	2014	UK	Public Health England
H144160430	DT 193	Human	2014	UK	Public Health England
H144160459	DT 120	Human	2014	UK	Public Health England
H144160463	PT U323	Human	2014	UK	Public Health England
H144160466	PT U302	Human	2014	UK	Public Health England
H144160467	None	Human	2014	UK	Public Health England
H144160468	DT 99	Human	2014	UK	Public Health England
H144160469	DT 9	Human	2014	UK	Public Health England

H144180582	DT 104	Human	2014	UK	Public Health England
H144180586	DT 2	Human	2014	UK	Public Health England
H144180588	DT 41	Human	2014	UK	Public Health England
H144180590	PT U302	Human	2014	UK	Public Health England
H144180592	DT 193	Human	2014	UK	Public Health England
H144200657	DT 208	Human	2014	UK	Public Health England
H144200658	DT 1	Human	2014	UK	Public Health England
H144200659	DT 7	Human	2014	UK	Public Health England
H144200661	DT 193	Human	2014	UK	Public Health England
H144220658	PT U302	Human	2014	UK	Public Health England
H144220659	Untypable	Human	2014	UK	Public Health England
H144220661	None	Human	2014	UK	Public Health England
H144220662	PT U323	Human	2014	UK	Public Health England
H144220667	DT 193	Human	2014	UK	Public Health England
H144220669	DT 193	Human	2014	UK	Public Health England
H144220671	DT 120	Human	2014	UK	Public Health England
H144220672	DT 8	Human	2014	UK	Public Health England
H144220673	DT 193	Human	2014	UK	Public Health England
H144220674	PT U311	Human	2014	UK	Public Health England
H144220675	DT 208	Human	2014	UK	Public Health England
H144220712	PT U277	GREEN FINCH	2014	UK	Public Health England
H144220720	DT 208	Human	2014	UK	Public Health England

H144220720	DT 208	Human	2014	UK	Public Health England
H144240607	DT 104	Human	2014	UK	Public Health England
H144240609	DT 193	Human	2014	UK	Public Health England
H144240612	DT 104	Human	2014	UK	Public Health England
H144240613	DT 193	Human	2014	UK	Public Health England
H144240614	DT 193	Human	2014	UK	Public Health England
H144240634	DT 135	Human	2014	UK	Public Health England
H144240637	DT 208	Human	2014	UK	Public Health England
H144240638	DT 208	Human	2014	UK	Public Health England
H144260658	DT 208	Human	2014	UK	Public Health England
H144260669	PT U288	Human	2014	UK	Public Health England
H144260673	DT 104	Human	2014	UK	Public Health England
H144260674	DT 8	Human	2014	UK	Public Health England
H144260675	DT 193	Human	2014	UK	Public Health England
H144260676	RDNC	Human	2014	UK	Public Health England
H144280521	DT 193	Human	2014	UK	Public Health England
H144280569	DT 101	Human	2014	UK	Public Health England
H144280570	None	Human	2014	UK	Public Health England
H144280571	DT 104	Human	2014	UK	Public Health England
H144300569	DT 193	Human	2014	UK	Public Health England
H144300576	PT U311	Human	2014	UK	Public Health England
H144300578	PT U311	Human	2014	UK	Public Health England

H144300579	PT U311	Human	2014	UK	Public Health England
H144320727	DT 120	Human	2014	UK	Public Health England
H144320757	DT 193	Human	2014	UK	Public Health England
H144320758	DT 193	Human	2014	UK	Public Health England
H144320759	PT U302	Human	2014	UK	Public Health England
H144320760	DT 101	Human	2014	UK	Public Health England
H144320771	DT 195	Human	2014	UK	Public Health England
H144320772	PT U323	Human	2014	UK	Public Health England
H144320773	PT U323	Human	2014	UK	Public Health England
H144340671	DT 193	Human	2014	UK	Public Health England
H144340706	PT U311	Human	2014	UK	Public Health England
H144340707	DT 96	Human	2014	UK	Public Health England
H144340709	DT 104	Human	2014	UK	Public Health England
H144340709	DT 104	Human	2014	UK	Public Health England
H144340710	PT U323	Human	2014	UK	Public Health England
H144340711	DT 193	Human	2014	UK	Public Health England
H144340712	DT 1	Human	2014	UK	Public Health England
H144340714	DT 193	Human	2014	UK	Public Health England
H144340715	DT 93	Human	2014	UK	Public Health England
H144340716	DT 193	Human	2014	UK	Public Health England
H144340718	DT 193	Human	2014	UK	Public Health England
H144340719	PT U311	Human	2014	UK	Public Health England

H144360255	DT 193	Human	2014	UK	Public Health England
H144360256	PT U302	Human	2014	UK	Public Health England
H144360257	DT 193	Human	2014	UK	Public Health England
H144360258	PT U302	Human	2014	UK	Public Health England
H144360259	RDNC	Human	2014	UK	Public Health England
H144380544	DT 2	Human	2014	UK	Public Health England
H144380574	PT U311	Human	2014	UK	Public Health England
H144380583	None	NCTC1188 1B2D	2014	UK	Public Health England
H144380585	PT U323	Human	2014	UK	Public Health England
H144380586	PT U323	Human	2014	UK	Public Health England
H144400569	DT 193	Human	2014	UK	Public Health England
H144400598	DT 193	Human	2014	UK	Public Health England
H144400600	PT U302	Human	2014	UK	Public Health England
H144400601	DT 193	Human	2014	UK	Public Health England
H144420807	DT 208	Human	2014	UK	Public Health England
H144420810	DT 2	Human	2014	UK	Public Health England
H144420811	DT 193	Human	2014	UK	Public Health England
H144420896	DT 2	Human	2014	UK	Public Health England
H144420905	DT 104	Human	2014	UK	Public Health England
H144420906	DT 2	Human	2014	UK	Public Health England
H144420908	DT 193	Human	2014	UK	Public Health England
H144420910	DT 120	Human	2014	UK	Public Health England

H144420913	DT 193	Human	2014	UK	Public Health England
H144420914	DT 193	Human	2014	UK	Public Health England
H144420915	DT 193	Human	2014	UK	Public Health England
H144420916	DT 193	Human	2014	UK	Public Health England
H144420918	PT U329	Human	2014	UK	Public Health England
H144440743	DT 193	Human	2014	UK	Public Health England
H144440744	DT 193	Human	2014	UK	Public Health England
H144440750	DT 193	Human	2014	UK	Public Health England
H144440753	PT U288	Human	2014	UK	Public Health England
H144440765	DT 193	Human	2014	UK	Public Health England
H144460488	Untypable	Human	2014	UK	Public Health England
H144480567	DT 193	Human	2014	UK	Public Health England
H144480568	DT 193	Human	2014	UK	Public Health England
H144480569	DT 120	Human	2014	UK	Public Health England
H144500515	Untypable	Human	2014	UK	Public Health England
H144500519	PT U302	Human	2014	UK	Public Health England
H144500533	PT U302	Human	2014	UK	Public Health England
H144500534	DT 193	Human	2014	UK	Public Health England
H144500535	PT U302	Human	2014	UK	Public Health England
H144520686	DT 193	Human	2014	UK	Public Health England
H144520687	DT 193	Human	2014	UK	Public Health England
H144520703	DT 1	Human	2014	UK	Public Health England

H144520714	PT U323	Human	2014	UK	Public Health England
H144520715	DT 2	Human	2014	UK	Public Health England
H144520716	DT 104	Human	2014	UK	Public Health England
H144520719	DT 193	Human	2014	UK	Public Health England
H144520720	PT U312	Human	2014	UK	Public Health England
H144520723	DT 193	Human	2014	UK	Public Health England
H144520724	DT 193	Human	2014	UK	Public Health England
H144520725	DT 193	Human	2014	UK	Public Health England
H144540567	None	Human	2014	UK	Public Health England
H144540642	DT 56	Human	2014	UK	Public Health England
H144540644	DT 193	Human	2014	UK	Public Health England
H144540652	DT 193	Human	2014	UK	Public Health England
H144540654	DT 193	Human	2014	UK	Public Health England
H144540655	PT U302	Human	2014	UK	Public Health England
H144540656	DT 132	Human	2014	UK	Public Health England
H144540658	DT 193	Human	2014	UK	Public Health England
H144540660	DT 2	Human	2014	UK	Public Health England
H144560569	DT 104	Human	2014	UK	Public Health England
H144560589	DT 132	UNKNOWN SOURCE	2014	UK	Public Health England
H144560590	DT 193	Human	2014	UK	Public Health England
H144560591	PT U329	Human	2014	UK	Public Health England
H144580647	DT 193	Human	2014	UK	Public Health England

H144580651	PT U311	Human	2014	UK	Public Health England
H144580657	DT 21	Human	2014	UK	Public Health England
H144600638	DT 193	Human	2014	UK	Public Health England
H144600640	RDNC	Human	2014	UK	Public Health England
H144600656	DT 193	Human	2014	UK	Public Health England
H144600660	DT 193	Human	2014	UK	Public Health England
H144620685	DT 193	Human	2014	UK	Public Health England
H144620695	DT 193	Human	2014	UK	Public Health England
H144620698	RDNC	Human	2014	UK	Public Health England
H144620699	DT 193	Human	2014	UK	Public Health England
H144620700	DT 2	Human	2014	UK	Public Health England
H144620701	DT 1	Human	2014	UK	Public Health England
H144620718	None	Human	2014	UK	Public Health England
H144620719	DT 193	Human	2014	UK	Public Health England
H144620721	DT 11	Human	2014	UK	Public Health England
H144620722	DT 193	Human	2014	UK	Public Health England
H144640572	DT 132	SOURCE UNKNOWN	2014	UK	Public Health England
H144640573	DT 132	SOURCE UNKNOWN	2014	UK	Public Health England
H144640576	DT 193	Human	2014	UK	Public Health England
H144640577	PT U312	Human	2014	UK	Public Health England
H144640578	DT 193	Human	2014	UK	Public Health England

H144640579	DT 193	Human	2014	UK	Public Health England
H144640580	DT 120	Human	2014	UK	Public Health England
H144640583	DT 104	Human	2014	UK	Public Health England
H144640584	PT U288	Human	2014	UK	Public Health England
H144640585	DT 120	Human	2014	UK	Public Health England
H144660684	RDNC	Human	2014	UK	Public Health England
H144660687	DT 132	SOURCE UNKNOWN	2014	UK	Public Health England
H144660698	DT 193	Human	2014	UK	Public Health England
H144660699	PT U281	Human	2014	UK	Public Health England
H144660700	PT U302	Human	2014	UK	Public Health England
H144660701	PT U323	Human	2014	UK	Public Health England
H144680508	DT 193	Human	2014	UK	Public Health England
H144680543	None	Human	2014	UK	Public Health England
H144680544	None	Human	2014	UK	Public Health England
H144680548	DT 193	Human	2014	UK	Public Health England
H144680549	DT 120	Human	2014	UK	Public Health England
H144680550	DT 193	Human	2014	UK	Public Health England
H144680551	DT 101	Human	2014	UK	Public Health England
H144700423	DT 193	Human	2014	UK	Public Health England
H144700424	RDNC	Human	2014	UK	Public Health England
H144700426	DT 132	Human	2014	UK	Public Health England
H144720736	DT 193	Human	2014	UK	Public Health England

H144720785	DT 193	Human	2014	UK	Public Health England
H144720834	None	Human	2014	UK	Public Health England
H144720840	DT 193	Human	2014	UK	Public Health England
H144720845	DT 193	Human	2014	UK	Public Health England
H144720847	DT 193	Human	2014	UK	Public Health England
H144740651	DT 2	Human	2014	UK	Public Health England
H144740658	DT 15	Human	2014	UK	Public Health England
H144740702	None	Human	2014	UK	Public Health England
H144740709	DT 193	Human	2014	UK	Public Health England
H144740710	DT 8	Human	2014	UK	Public Health England
H144740711	None	Human	2014	UK	Public Health England
H144740712	RDNC	Human	2014	UK	Public Health England
H144740718	DT 1	Human	2014	UK	Public Health England
H144740719	DT 193	Human	2014	UK	Public Health England
H144740720	DT 101	Human	2014	UK	Public Health England
H144740723	DT 179	Human	2014	UK	Public Health England
H144760703	DT 132	FOOD SOURCE UNKNOWN	2014	UK	Public Health England
H144760704	DT 132	FOOD SOURCE UNKNOWN	2014	UK	Public Health England
H144760705	DT 132	FOOD SOURCE UNKNOWN	2014	UK	Public Health England

H144760707	DT 132	FOOD SOURCE UNKNOWN	2014	UK	Public Health England
H144760708	DT 191	Human	2014	UK	Public Health England
H144760719	DT 132	FOOD SOURCE UNKNOWN	2014	UK	Public Health England
H144780621	DT 195	Human	2014	UK	Public Health England
H144780661	DT 132	SOURCE UNKNOWN	2014	UK	Public Health England
H144780666	DT 132	Human	2014	UK	Public Health England
H144780666	DT 132	Human	2014	UK	Public Health England
H144780673	DT 120	Human	2014	UK	Public Health England
H144800748	DT 193	Human	2014	UK	Public Health England
H144800782	PT U311	Human	2014	UK	Public Health England
H144800784	DT 120	Human	2014	UK	Public Health England
H144820727	DT 21	Human	2014	UK	Public Health England
H144820733	PT U311	Human	2014	UK	Public Health England
H144840757	DT 120	Human	2014	UK	Public Health England
H144840760	PT U311	Human	2014	UK	Public Health England
H144840761	DT 120	Human	2014	UK	Public Health England
H144840762	DT 193	Human	2014	UK	Public Health England
H144840763	DT 193	Human	2014	UK	Public Health England
H144840764	PT U311	Human	2014	UK	Public Health England
H144880865	DT 12	Human	2014	UK	Public Health England

H144880866	None	Human	2014	UK	Public Health England
H144880867	DT 193	Human	2014	UK	Public Health England
H144880869	DT 41	Human	2014	UK	Public Health England
H144880871	DT 193	Human	2014	UK	Public Health England
H144880873	DT 120	Human	2014	UK	Public Health England
H144900572	DT 104	Human	2014	UK	Public Health England
H144900575	DT 161	Human	2014	UK	Public Health England
H144900578	PT U323	Human	2014	UK	Public Health England
H144920410	PT U311	Human	2014	UK	Public Health England
H144920411	PT U311	Human	2014	UK	Public Health England
H144920412	PT U311	Human	2014	UK	Public Health England
H144920413	PT U311	Human	2014	UK	Public Health England
H144920618	None	Human	2014	UK	Public Health England
H144920624	DT 120	Human	2014	UK	Public Health England
H144920627	RDNC	Human	2014	UK	Public Health England
H144920628	Untypable	Human	2014	UK	Public Health England
H144920630	DT 120	Human	2014	UK	Public Health England
H144920634	DT 193	Human	2014	UK	Public Health England
H144920635	DT 120	Human	2014	UK	Public Health England
H144920636	PT U311	Human	2014	UK	Public Health England
H144920637	DT 56	Human	2014	UK	Public Health England
H144940487	PT U311	Human	2014	UK	Public Health England

H144940490	DT 41	Human	2014	UK	Public Health England
H144940553	DT 120	Human	2014	UK	Public Health England
H144940556	DT 120	Human	2014	UK	Public Health England
H144940560	DT 193	Human	2014	UK	Public Health England
H144960545	DT 120	Human	2014	UK	Public Health England
H144960547	DT 120	Human	2014	UK	Public Health England
H144980592	DT 208	Human	2014	UK	Public Health England
H144980717	DT 193	Human	2014	UK	Public Health England
H144980718	None	Human	2014	UK	Public Health England
H144980731	DT 120	Human	2014	UK	Public Health England
H144980732	DT 120	Human	2014	UK	Public Health England
H144980733	DT 120	Human	2014	UK	Public Health England
H144980734	DT 120	Human	2014	UK	Public Health England
H144980735	DT 120	Human	2014	UK	Public Health England
H145000544	DT 20	Human	2014	UK	Public Health England
H145000545	PT U317	Human	2014	UK	Public Health England
H145020733	DT 191	Human	2014	UK	Public Health England
H145020735	DT 193	Human	2014	UK	Public Health England
H145040692	DT 193	Human	2014	UK	Public Health England
H145040715	PT U311	Human	2014	UK	Public Health England
H145040717	DT 120	Human	2014	UK	Public Health England
H145040718	DT 120	Human	2014	UK	Public Health England

H145040721	DT 120	Human	2014	UK	Public Health England
H145040722	DT 120	Human	2014	UK	Public Health England
H145040728	DT 193	Human	2014	UK	Public Health England
H145060536	DT 193	Human	2014	UK	Public Health England
H145060537	DT 193	Human	2014	UK	Public Health England
H145060538	None	Human	2014	UK	Public Health England
H145060720	DT 104	Human	2014	UK	Public Health England
H145060728	Untypable	Human	2014	UK	Public Health England
H145080408	DT 193	Human	2014	UK	Public Health England
H145080409	DT 193	Human	2014	UK	Public Health England
H145080411	DT 120	Human	2014	UK	Public Health England
H145080413	DT 193	Human	2014	UK	Public Health England
H145080422	DT 12	Human	2014	UK	Public Health England
H145100624	DT 120	Human	2014	UK	Public Health England
H145100673	DT 40	Human	2014	UK	Public Health England
H145100674	DT 193	Human	2014	UK	Public Health England
H145100676	DT 104	Human	2014	UK	Public Health England
H145100680	DT 193	Human	2014	UK	Public Health England
H145100686	DT 195	Human	2014	UK	Public Health England
H145100804	None	Human	2014	UK	Public Health England
H145120591	DT 104	Human	2014	UK	Public Health England
H145120596	DT 2	Human	2014	UK	Public Health England

H145120597	DT 2	Human	2014	UK	Public Health England
H145120612	DT 193	Human	2014	UK	Public Health England
H145120615	DT 193	Human	2014	UK	Public Health England
H145120616	DT 193	Human	2014	UK	Public Health England
H145120619	PT U329	Human	2014	UK	Public Health England
H145120620	DT 18	Human	2014	UK	Public Health England
H145120625	None	Human	2014	UK	Public Health England
H145140590	DT 193	Human	2014	UK	Public Health England
H145140591	PT U302	Human	2014	UK	Public Health England
H145140592	DT 193	Human	2014	UK	Public Health England
H145140594	DT 1	Human	2014	UK	Public Health England
H145140595	PT U311	Human	2014	UK	Public Health England
H145140599	PT U311	Human	2014	UK	Public Health England
H145140600	DT 120	Human	2014	UK	Public Health England
H145140601	DT 120	Human	2014	UK	Public Health England
H145160571	DT 1	Human	2014	UK	Public Health England
H145160574	DT 104	Human	2014	UK	Public Health England
H145160676	DT 10	Human	2014	UK	Public Health England
H145180570	PT U302	Human	2014	UK	Public Health England
H145180710	DT 7	Human	2014	UK	Public Health England
H145180715	PT U311	Human	2014	UK	Public Health England
H145180717	DT 193	Human	2014	UK	Public Health England

H145180718	None	Human	2014	UK	Public Health England
H145180719	DT 193	Human	2014	UK	Public Health England
H145180720	PT U323	Human	2014	UK	Public Health England
H145200602	DT 41	Human	2014	UK	Public Health England
H145200606	PT U311	Human	2014	UK	Public Health England
H145200609	DT 2	Human	2014	UK	Public Health England
H145200611	None	Human	2014	UK	Public Health England
H145220846	DT 193	Human	2014	UK	Public Health England
H145220852	DT 193	Human	2014	UK	Public Health England
H145220853	DT 1	Human	2014	UK	Public Health England
H145240540	PT U311	Human	2014	UK	Public Health England
H145240594	RDNC	Human	2014	UK	Public Health England
H150120636	DT 193	Human	2015	UK	Public Health England
H150120639	DT 135	Human	2015	UK	Public Health England
H150120641	DT 193	Human	2015	UK	Public Health England
H150120645	DT 193	Human	2015	UK	Public Health England
H150120668	DT 1	Human	2015	UK	Public Health England
H150140575	DT 193	Human	2015	UK	Public Health England
H150140577	DT 193	Human	2015	UK	Public Health England
H150200808	PT U288	Human	2015	UK	Public Health England
H150200809	DT 193	Human	2015	UK	Public Health England
H150200814	None	Human	2015	UK	Public Health England

H150220500	RDNC	Human	2015	UK	Public Health England
H150220501	DT 120	Human	2015	UK	Public Health England
H150220502	DT 120	Human	2015	UK	Public Health England
H150220503	DT 120	Human	2015	UK	Public Health England
H150220505	DT 193	Human	2015	UK	Public Health England
H150220507	DT 193	Human	2015	UK	Public Health England
H150220512	None	Human	2015	UK	Public Health England
H150220513	DT 193	Human	2015	UK	Public Health England
H150220521	DT 41	Human	2015	UK	Public Health England
H150240436	DT 193	Human	2015	UK	Public Health England
H150240437	PT U311	Human	2015	UK	Public Health England
H150240438	DT 41	Human	2015	UK	Public Health England
H150240440	DT 120	Human	2015	UK	Public Health England
H150240441	DT 120	Human	2015	UK	Public Health England
H150260321	DT 193	Human	2015	UK	Public Health England
H150260323	DT 193	Human	2015	UK	Public Health England
H150260324	PT U323	Human	2015	UK	Public Health England
H150260326	DT 104	Human	2015	UK	Public Health England
H150260329	DT 193	RAW POULTRY	2015	UK	Public Health England
H150260336	DT 2	Human	2015	UK	Public Health England
H150280543	DT 104	Human	2015	UK	Public Health England
H150280548	DT 193	Human	2015	UK	Public Health England

H150300415	PT U331	Human	2015	UK	Public Health England
H150300545	DT 193	Human	2015	UK	Public Health England
H150300551	DT 193	Human	2015	UK	Public Health England
H150300552	PT U311	Human	2015	UK	Public Health England
H150320538	PT U311	Human	2015	UK	Public Health England
H150320614	DT 41	Human	2015	UK	Public Health England
H150320692	DT 104	Human	2015	UK	Public Health England
H150320697	RDNC	Human	2015	UK	Public Health England
H150320702	DT 193	Human	2015	UK	Public Health England
H150320703	DT 104	Human	2015	UK	Public Health England
H150320704	DT 193	Human	2015	UK	Public Health England
H150320705	DT 120	Human	2015	UK	Public Health England
H150320711	DT 193	Human	2015	UK	Public Health England
H150320714	PT U288	Human	2015	UK	Public Health England
H150320715	DT 193	Human	2015	UK	Public Health England
H150340498	RDNC	Human	2015	UK	Public Health England
H150340501	DT 137	Human	2015	UK	Public Health England
H150340503	DT 2	Human	2015	UK	Public Health England
H150380351	DT 104	Human	2015	UK	Public Health England
H150400536	RDNC	Human	2015	UK	Public Health England
H150400537	DT 193	Human	2015	UK	Public Health England
H150400539	Untypable	Human	2015	UK	Public Health England

H150400540	DT 193	Human	2015	UK	Public Health England
H150440669	DT 193	Human	2015	UK	Public Health England
H150440671	DT 193	Human	2015	UK	Public Health England
H150440673	DT 193	Human	2015	UK	Public Health England
H150440674	DT 193	Human	2015	UK	Public Health England
H150460715	PT U323	Human	2015	UK	Public Health England
H150460717	PT U288	Human	2015	UK	Public Health England
H150460720	DT 193	Human	2015	UK	Public Health England
H150460724	DT 193	Human	2015	UK	Public Health England
H150480665	DT 104	Human	2015	UK	Public Health England
H150480666	DT 193	Human	2015	UK	Public Health England
H150520684	PT U311	Human	2015	UK	Public Health England
H150520687	DT 193	Human	2015	UK	Public Health England
H150540596	DT 120	Human	2015	UK	Public Health England
H150540602	DT 193	Human	2015	UK	Public Health England
H150540604	PT U311	Human	2015	UK	Public Health England
H150560629	DT 193	Human	2015	UK	Public Health England
H150560633	DT 132	Human	2015	UK	Public Health England
H150560638	RDNC	Human	2015	UK	Public Health England
H150560642	DT 41	Human	2015	UK	Public Health England
H150580636	RDNC	RAW PRODUCE	2015	UK	Public Health England
H150580642	DT 104	Human	2015	UK	Public Health England

H150580643	RDNC	Human	2015	UK	Public Health England
H150580648	DT 56	Human	2015	UK	Public Health England
H150600548	DT 104	Human	2015	UK	Public Health England
H150620544	DT 193	Human	2015	UK	Public Health England
H150620545	DT 193	Human	2015	UK	Public Health England
H150620591	DT 193	Human	2015	UK	Public Health England
H150620593	DT 193	Human	2015	UK	Public Health England
H150640534	DT 193	Human	2015	UK	Public Health England
H150640559	DT 8	Human	2015	UK	Public Health England
H150640562	RDNC	Human	2015	UK	Public Health England
H150640564	DT 193	RAW SAUSAGE	2015	UK	Public Health England
H150640565	DT 193	Human	2015	UK	Public Health England
H150660556	DT 135	Human	2015	UK	Public Health England
H150660587	DT 10	Human	2015	UK	Public Health England
H150660588	DT 193	Human	2015	UK	Public Health England
H150660591	DT 193	Human	2015	UK	Public Health England
H150680611	DT 193	Human	2015	UK	Public Health England
H150700258	PT U288	Human	2015	UK	Public Health England
H150700259	PT U288	Human	2015	UK	Public Health England
H150700279	RDNC	Human	2015	UK	Public Health England
H150700280	DT 10	Human	2015	UK	Public Health England
H150700283	DT 135	Human	2015	UK	Public Health England

H150720593	RDNC	IQA	2015	UK	Public Health England
H150720596	DT 193	Human	2015	UK	Public Health England
H150720604	DT 193	Human	2015	UK	Public Health England
H150740512	DT 104	Human	2015	UK	Public Health England
H150740514	None	Human	2015	UK	Public Health England
H150740515	Untypable	Human	2015	UK	Public Health England
H150740519	PT U323	Human	2015	UK	Public Health England
H150740521	PT U313	Human	2015	UK	Public Health England
H150760395	DT 193	Human	2015	UK	Public Health England
H150780479	DT 41	Human	2015	UK	Public Health England
H150780487	PT U302	Human	2015	UK	Public Health England
H150780489	RDNC	Human	2015	UK	Public Health England
H150780494	PT U302	Human	2015	UK	Public Health England
H150800485	RDNC	Human	2015	UK	Public Health England
H150800523	DT 193	Human	2015	UK	Public Health England
H150820630	DT 193	Human	2015	UK	Public Health England
H150820634	Untypable	Human	2015	UK	Public Health England
H150820638	DT 104	Human	2015	UK	Public Health England
H150820681	DT 44	Human	2015	UK	Public Health England
H150820682	DT 1	Human	2015	UK	Public Health England
H150840619	RDNC	Human	2015	UK	Public Health England
H150840621	DT 41	Human	2015	UK	Public Health England

H150840622	DT 193	Human	2015	UK	Public Health England
H150880555	RDNC	Human	2015	UK	Public Health England
H150880559	PT U288	RAW SAUSAGE	2015	UK	Public Health England
H150880560	PT U288	RAW SAUSAGE	2015	UK	Public Health England
H150880561	PT U288	RAW SAUSAGE	2015	UK	Public Health England
H150880563	DT 193	RAW SAUSAGE	2015	UK	Public Health England
H150900461	DT 56	Human	2015	UK	Public Health England
H150900562	RDNC	Human	2015	UK	Public Health England
H150900563	DT 193	Human	2015	UK	Public Health England
H150900564	DT 120	Human	2015	UK	Public Health England
H150900566	PT U274	Human	2015	UK	Public Health England
H150900568	DT 193	Human	2015	UK	Public Health England
H150900571	DT 193	Human	2015	UK	Public Health England
H150920652	DT 193	Human	2015	UK	Public Health England
H150920673	Untypable	Human	2015	UK	Public Health England
H150920675	DT 193	Human	2015	UK	Public Health England
H150920678	DT 193	Human	2015	UK	Public Health England
H150920681	DT 193	Human	2015	UK	Public Health England
H150940605	RDNC	Human	2015	UK	Public Health England
H150940620	PT U274	Human	2015	UK	Public Health England
H150940621	PT U312	Human	2015	UK	Public Health England
H150940623	DT 193	Human	2015	UK	Public Health England

H150940624	Untypable	Human	2015	UK	Public Health England
H150940627	DT 193	Human	2015	UK	Public Health England
H150940630	DT 193	Human	2015	UK	Public Health England
H150960615	DT 2	Human	2015	UK	Public Health England
H150960623	DT 193	Human	2015	UK	Public Health England
H150980442	DT 193	Human	2015	UK	Public Health England
H151000369	DT 12	Human	2015	UK	Public Health England
H151000372	DT 1	Human	2015	UK	Public Health England
H151020588	DT 193	Human	2015	UK	Public Health England
H151020623	PT U323	Human	2015	UK	Public Health England
H151020625	DT 193	Human	2015	UK	Public Health England
H151020626	Untypable	Human	2015	UK	Public Health England
H151020627	PT U323	Human	2015	UK	Public Health England
H151020629	DT 12	UKNEQAS	2015	UK	Public Health England
H151020630	DT 193	Human	2015	UK	Public Health England
H151020631	DT 193	Human	2015	UK	Public Health England
H151020632	DT 193	Human	2015	UK	Public Health England
H151020634	DT 2	Human	2015	UK	Public Health England
H151020638	PT U311	Human	2015	UK	Public Health England
H151040666	PT U312	Human	2015	UK	Public Health England
H151040668	DT 137	Human	2015	UK	Public Health England
H151040669	DT 131	Human	2015	UK	Public Health England

H151040672	DT 193	Human	2015	UK	Public Health England
H151060483	DT 193	Human	2015	UK	Public Health England
H151060484	Untypable	Human	2015	UK	Public Health England
H151060485	Untypable	Human	2015	UK	Public Health England
H151060493	DT 193	Human	2015	UK	Public Health England
H151060497	DT 120	Human	2015	UK	Public Health England
H151080426	Untypable	Human	2015	UK	Public Health England
H151080428	DT 120	Human	2015	UK	Public Health England
H151080429	DT 208	Human	2015	UK	Public Health England
H151080433	DT 193	Human	2015	UK	Public Health England
H151080437	Untypable	Human	2015	UK	Public Health England
H151100483	DT 193	Human	2015	UK	Public Health England
H151100485	DT 190	Human	2015	UK	Public Health England
H151100486	Untypable	Human	2015	UK	Public Health England
H151100487	Untypable	Human	2015	UK	Public Health England
H151100488	Untypable	Human	2015	UK	Public Health England
H151100488	Untypable	Human	2015	UK	Public Health England
H151100490	Untypable	Human	2015	UK	Public Health England
H151100491	DT 193	Human	2015	UK	Public Health England
H151120549	DT 8	Human	2015	UK	Public Health England
H151120553	DT 193	Human	2015	UK	Public Health England
H151120554	RDNC	Human	2015	UK	Public Health England

H151120555	Untypable	Human	2015	UK	Public Health England
H151120556	Untypable	Human	2015	UK	Public Health England
H151120557	Untypable	Human	2015	UK	Public Health England
H151120558	Untypable	Human	2015	UK	Public Health England
H151120559	Untypable	Human	2015	UK	Public Health England
H151120560	Untypable	Human	2015	UK	Public Health England
H151120561	Untypable	Human	2015	UK	Public Health England
H151120562	DT 120	Human	2015	UK	Public Health England
H151120568	DT 193	Human	2015	UK	Public Health England
H151140648	DT 193	Human	2015	UK	Public Health England
H151140654	DT 193	Human	2015	UK	Public Health England
H151140674	DT 120	Human	2015	UK	Public Health England
H151160603	DT 193	Human	2015	UK	Public Health England
H151180241	None	Human	2015	UK	Public Health England
H151180245	DT 193	Human	2015	UK	Public Health England
H151180246	Untypable	Human	2015	UK	Public Health England
H151180253	Untypable	Human	2015	UK	Public Health England
H151200223	PT U288	Human	2015	UK	Public Health England
H151200224	RDNC	Human	2015	UK	Public Health England
H151220453	DT 193	Human	2015	UK	Public Health England
H151220457	DT 193	Human	2015	UK	Public Health England
H151220461	Untypable	Human	2015	UK	Public Health England

H151220462	Untypable	Human	2015	UK	Public Health England
H151220463	Untypable	Human	2015	UK	Public Health England
H151220464	Untypable	Human	2015	UK	Public Health England
H151220465	Untypable	Human	2015	UK	Public Health England
H151220466	Untypable	Human	2015	UK	Public Health England
H151220467	DT 193	Human	2015	UK	Public Health England
H151220468	DT 193	Human	2015	UK	Public Health England
H151220469	Untypable	Human	2015	UK	Public Health England
H151220470	DT 193	Human	2015	UK	Public Health England
H151220472	DT 56	HOUSE SPARROW	2015	UK	Public Health England
H151220474	DT 56	REED BUNTING	2015	UK	Public Health England
H151240457	DT 193	Human	2015	UK	Public Health England
H151280436	DT 120	Human	2015	UK	Public Health England
H151280438	DT 104	Human	2015	UK	Public Health England
H151280509	DT 193	Human	2015	UK	Public Health England
H151280512	DT 193	Human	2015	UK	Public Health England
H151280513	PT U323	Human	2015	UK	Public Health England
H151300474	DT 193	Human	2015	UK	Public Health England
H151300475	Untypable	Human	2015	UK	Public Health England
H151300476	Untypable	Human	2015	UK	Public Health England
H151300477	Untypable	Human	2015	UK	Public Health England
H151300479	DT 193	Human	2015	UK	Public Health England

H151300480	Untypable	Human	2015	UK	Public Health England
H151300482	RDNC	Human	2015	UK	Public Health England
H151320496	DT 193	Human	2015	UK	Public Health England
H151320498	PT U288	Human	2015	UK	Public Health England
H151320500	DT 12	Human	2015	UK	Public Health England
H151320504	DT 193	Human	2015	UK	Public Health England
H151340523	DT 193	Human	2015	UK	Public Health England
H151340531	PT U323	Human	2015	UK	Public Health England
H151340533	DT 195	Human	2015	UK	Public Health England
H151340538	DT 120	Human	2015	UK	Public Health England
H151360443	DT 195	Human	2015	UK	Public Health England
H151360458	DT 193	Human	2015	UK	Public Health England
H151380459	DT 193	Human	2015	UK	Public Health England
H151380461	DT 195	Human	2015	UK	Public Health England
H151380467	DT 193	Human	2015	UK	Public Health England
H151400464	DT 193	Human	2015	UK	Public Health England
H151420580	DT 104	Human	2015	UK	Public Health England
H151420583	Untypable	Human	2015	UK	Public Health England
H151420588	DT 193	Human	2015	UK	Public Health England
H151440507	DT 193	Human	2015	UK	Public Health England
H151440566	DT 120	Human	2015	UK	Public Health England
H151540860	DT 104	Human	2015	UK	Public Health England

H151540864	DT 120	Human	2015	UK	Public Health England
H151540868	Untypable	Human	2015	UK	Public Health England
H151540879	DT 193	Human	2015	UK	Public Health England
H151540883	DT 193	Human	2015	UK	Public Health England
H151540888	DT 193	Human	2015	UK	Public Health England
H151540890	DT 193	Human	2015	UK	Public Health England
H151540892	DT 120	Human	2015	UK	Public Health England
H151560674	DT 193	SOURCE UNKNOWN	2015	UK	Public Health England
H151560676	DT 193	SOURCE UNKNOWN	2015	UK	Public Health England
H151560680	PT U331	Human	2015	UK	Public Health England
H151560685	Untypable	Human	2015	UK	Public Health England
H151580353	DT 10	Human	2015	UK	Public Health England
H151620337	PT U309	Human	2015	UK	Public Health England
H151620362	DT 104	Human	2015	UK	Public Health England
H151640374	DT 193	Human	2015	UK	Public Health England
H151640376	PT U302	Human	2015	UK	Public Health England
H151640382	DT 191	Human	2015	UK	Public Health England
H151640383	Untypable	Human	2015	UK	Public Health England
H151640411	DT 120	Human	2015	UK	Public Health England
H151660628	Untypable	Human	2015	UK	Public Health England
H151660629	Untypable	Human	2015	UK	Public Health England

H151680561	DT 193	Human	2015	UK	Public Health England
H151700443	DT 120	Human	2015	UK	Public Health England
H151700448	DT 193	Human	2015	UK	Public Health England
H151700449	DT 93	Human	2015	UK	Public Health England
H151700453	DT 193	Human	2015	UK	Public Health England
H151720436	DT 193	Human	2015	UK	Public Health England
H151720439	DT 56	Human	2015	UK	Public Health England
H151720461	DT 56	Human	2015	UK	Public Health England
H151720643	DT 104	Human	2015	UK	Public Health England
H151720657	DT 193	Human	2015	UK	Public Health England
H151720671	DT 104	Human	2015	UK	Public Health England
H151740324	Untypable	Human	2015	UK	Public Health England
H151740325	Untypable	Human	2015	UK	Public Health England
H151740327	Untypable	Human	2015	UK	Public Health England
H151760296	Untypable	Human	2015	UK	Public Health England
H151760308	Untypable	Human	2015	UK	Public Health England
H151800443	Untypable	Human	2015	UK	Public Health England
H151820415	Untypable	Human	2015	UK	Public Health England
H151820416	Untypable	Human	2015	UK	Public Health England
H151820500	DT 193	Human	2015	UK	Public Health England
H151820505	DT 104	Human	2015	UK	Public Health England
H151820511	DT 104	Human	2015	UK	Public Health England

H151820523	DT 193	Human	2015	UK	Public Health England
H151820524	DT 120	Human	2015	UK	Public Health England
H151840449	RDNC	Human	2015	UK	Public Health England
H151840455	DT 120	Human	2015	UK	Public Health England
H151840468	PT U311	Human	2015	UK	Public Health England
H151840476	DT 104	Human	2015	UK	Public Health England
H151860632	DT 1	Human	2015	UK	Public Health England
H151920509	PT U311	Human	2015	UK	Public Health England
H151920511	PT U311	Human	2015	UK	Public Health England
H151940558	PT U288	Human	2015	UK	Public Health England
H151960469	RDNC	Human	2015	UK	Public Health England
H151980548	DT 1	Human	2015	UK	Public Health England
H152020819	Untypable	Human	2015	UK	Public Health England
H152020821	DT 193	Human	2015	UK	Public Health England
H152020822	DT 42	Human	2015	UK	Public Health England
H152020836	DT 132	Human	2015	UK	Public Health England
H152020838	DT 193	Human	2015	UK	Public Health England
H152020863	DT 87	Human	2015	UK	Public Health England
H152040583	DT 193	Human	2015	UK	Public Health England
H152060596	DT 193	Human	2015	UK	Public Health England
H152060607	DT 120	Human	2015	UK	Public Health England
H152060609	PT U298	Human	2015	UK	Public Health England

H152080598	DT 193	RAW MEAT	2015	UK	Public Health England
H152080602	DT 193	Human	2015	UK	Public Health England
H152080603	DT 193	RAW MEAT	2015	UK	Public Health England
H152080608	DT 135	Human	2015	UK	Public Health England
H152100636	DT 56	Human	2015	UK	Public Health England
H152120540	DT 193	Human	2015	UK	Public Health England
H152120542	DT 193	Human	2015	UK	Public Health England
H152120543	PT U288	Human	2015	UK	Public Health England
H152120554	DT 193	Human	2015	UK	Public Health England
H152180624	PT U298	Human	2015	UK	Public Health England
H152180625	PT U302	Human	2015	UK	Public Health England
H152180628	DT 193	Human	2015	UK	Public Health England
H152240454	RDNC	Human	2015	UK	Public Health England
H152240468	DT 193	Human	2015	UK	Public Health England
H152240488	DT 193	Human	2015	UK	Public Health England
H152260543	DT 193	Human	2015	UK	Public Health England
H152280417	DT 41	Human	2015	UK	Public Health England
H152280418	PT U302	Human	2015	UK	Public Health England
H152280421	RDNC	Human	2015	UK	Public Health England
H152300644	DT 1	Human	2015	UK	Public Health England
H152300647	PT U329	Human	2015	UK	Public Health England
H152300650	PT U329	Human	2015	UK	Public Health England

H152320501	DT 120	Human	2015	UK	Public Health England
H152320524	PT U302	Human	2015	UK	Public Health England
H152320527	DT 193	Human	2015	UK	Public Health England
H152320528	PT U329	Human	2015	UK	Public Health England
H152320546	DT 120	Human	2015	UK	Public Health England
H152320547	PT U329	Human	2015	UK	Public Health England
H152340669	DT 104	Human	2015	UK	Public Health England
H152340700	PT U302	Human	2015	UK	Public Health England
H152340701	DT 10	Human	2015	UK	Public Health England
H152340702	DT 10	Human	2015	UK	Public Health England
H152340712	DT 120	Human	2015	UK	Public Health England
H152340713	DT 193	Human	2015	UK	Public Health England
H152340799	DT 1	EMRU	2015	UK	Public Health England
H152360402	DT 193	Human	2015	UK	Public Health England
H152360403	PT U329	Human	2015	UK	Public Health England
H152360408	DT 116	Human	2015	UK	Public Health England
H152380439	DT 193	Human	2015	UK	Public Health England
H152380466	PT U330	CAT	2015	UK	Public Health England
H152380468	PT U302	Human	2015	UK	Public Health England
H152380479	PT U302	Human	2015	UK	Public Health England
H152400594	DT 193	Human	2015	UK	Public Health England
H152400631	PT U288	Human	2015	UK	Public Health England

H152400632	DT 193	Human	2015	UK	Public Health England
H152400637	PT U329	Human	2015	UK	Public Health England
H152400637	PT U329	Human	2015	UK	Public Health England
H152420810	Untypable	Human	2015	UK	Public Health England
H152420817	DT 1	Human	2015	UK	Public Health England
H152420820	PT U329	Human	2015	UK	Public Health England
H152440505	DT 193	Human	2015	UK	Public Health England
H152440529	DT 120	Human	2015	UK	Public Health England
H152440533	DT 104	Human	2015	UK	Public Health England
H152460484	DT 193	Human	2015	UK	Public Health England
H152460491	PT U329	Human	2015	UK	Public Health England
H152460492	PT U311	Human	2015	UK	Public Health England
H152480461	PT U307	Human	2015	UK	Public Health England
H152480473	RDNC	Human	2015	UK	Public Health England
H152520807	DT 7	Human	2015	UK	Public Health England
H152520822	Untypable	Human	2015	UK	Public Health England
H152520823	DT 193	Human	2015	UK	Public Health England
H152520830	DT 116	Human	2015	UK	Public Health England
H152540595	DT 193	Human	2015	UK	Public Health England
H152540598	PT U329	Human	2015	UK	Public Health England
H152540610	RDNC	Human	2015	UK	Public Health England
H152540619	DT 66	Human	2015	UK	Public Health England

H152580650	RDNC	Human	2015	UK	Public Health England
H152580658	DT 46	Human	2015	UK	Public Health England
H152600555	DT 193	Human	2015	UK	Public Health England
H152600561	Untypable	Human	2015	UK	Public Health England
H152620495	DT 193	Human	2015	UK	Public Health England
H152640594	RDNC	Human	2015	UK	Public Health England
H152660386	PT U288	Human	2015	UK	Public Health England
H152660397	DT 135	Human	2015	UK	Public Health England
H152700573	PT U302	Human	2015	UK	Public Health England
H152700576	PT U329	Human	2015	UK	Public Health England
H152700580	DT 56	Human	2015	UK	Public Health England
H152700586	DT 120	Human	2015	UK	Public Health England
H152700587	DT 12	Human	2015	UK	Public Health England
H152720601	DT 46	Human	2015	UK	Public Health England
H152720602	DT 46	Human	2015	UK	Public Health England
H152720605	DT 8	Human	2015	UK	Public Health England
H152720608	DT 104	Human	2015	UK	Public Health England
H152720613	DT 85	Human	2015	UK	Public Health England
H152720615	DT 104	Human	2015	UK	Public Health England
H152720617	PT U311	Human	2015	UK	Public Health England
H152720637	DT 116	Human	2015	UK	Public Health England
H152720640	PT U311	Human	2015	UK	Public Health England

H152720641	DT 193	RAW MATERIAL SPICE	2015	UK	Public Health England
H152720642	DT 120	Human	2015	UK	Public Health England
H152720643	RDNC	Human	2015	UK	Public Health England
H152720648	DT 120	Human	2015	UK	Public Health England
H152720655	DT 120	Human	2015	UK	Public Health England
H152720661	PT U312	Human	2015	UK	Public Health England
H152720664	DT 193	Human	2015	UK	Public Health England
H152740235	DT 193	Human	2015	UK	Public Health England
H152740236	PT U298	Human	2015	UK	Public Health England
H152740238	DT 193	Human	2015	UK	Public Health England
H152740240	Untypable	Human	2015	UK	Public Health England
H152740273	DT 8	Human	2015	UK	Public Health England
H152760467	DT 193	Human	2015	UK	Public Health England
H152760469	DT 120	Human	2015	UK	Public Health England
H152760473	RDNC	Human	2015	UK	Public Health England
H152820418	DT 9	Human	2015	UK	Public Health England
H152840417	DT 193	Human	2015	UK	Public Health England
H152860554	DT 101	Human	2015	UK	Public Health England
H152860555	DT 120	Human	2015	UK	Public Health England
H152880399	DT 120	Human	2015	UK	Public Health England
H152900244	DT 104	Human	2015	UK	Public Health England
H152900267	PT U278	Human	2015	UK	Public Health England

H152920486	DT 40	Human	2015	UK	Public Health England
H152920498	PT U278	Human	2015	UK	Public Health England
H152940468	DT 2	Human	2015	UK	Public Health England
H152940469	RDNC	Human	2015	UK	Public Health England
H152980464	RDNC	Human	2015	UK	Public Health England
H153000432	DT 104	Human	2015	UK	Public Health England
H153020667	RDNC	Human	2015	UK	Public Health England
H153040541	PT U329	Human	2015	UK	Public Health England
H153120511	DT 193	Human	2015	UK	Public Health England
H153120515	PT U331	Human	2015	UK	Public Health England
H153140404	DT 193	Human	2015	UK	Public Health England
H153160274	DT 193	Human	2015	UK	Public Health England
H153160280	PT U323	Human	2015	UK	Public Health England
H153220457	DT 193	Human	2015	UK	Public Health England
H153260506	PT U323	Human	2015	UK	Public Health England
H153260510	DT 12	Human	2015	UK	Public Health England
H153300540	RDNC	Human	2015	UK	Public Health England
H153400468	DT 104	Human	2015	UK	Public Health England
H153400470	DT 8	Human	2015	UK	Public Health England
H153440473	DT 12	Human	2015	UK	Public Health England
H153440473	DT 12	Human	2015	UK	Public Health England
H153440508	DT 120	Human	2015	UK	Public Health England

H153460592	DT 104	Human	2015	UK	Public Health England
H153460595	DT 193	Human	2015	UK	Public Health England
H153460616	DT 104	Human	2015	UK	Public Health England
H153480514	DT 193	Human	2015	UK	Public Health England
H153520446	DT 120	Human	2015	UK	Public Health England
H153520451	DT 193	Human	2015	UK	Public Health England
H153520457	DT 193	Human	2015	UK	Public Health England
H153520466	PT U312	Human	2015	UK	Public Health England
H153520481	Untypable	Human	2015	UK	Public Health England
H153540331	Untypable	Human	2015	UK	Public Health England
H153560137	Untypable	Human	2015	UK	Public Health England
H153560143	PT U302	Human	2015	UK	Public Health England
H153640677	DT 195	Human	2015	UK	Public Health England
H153640680	DT 193	Human	2015	UK	Public Health England
H153640688	PT U302	Human	2015	UK	Public Health England
H153680363	DT 193	Human	2015	UK	Public Health England
H153700324	DT 193	Human	2015	UK	Public Health England
H153700446	PT U329	Human	2015	UK	Public Health England
H153740470	DT 120	Human	2015	UK	Public Health England
H153820390	RDNC	Human	2015	UK	Public Health England
H153840505	DT 193	Human	2015	UK	Public Health England
H153840508	PT U311	Human	2015	UK	Public Health England

H153840513	DT 193	Human	2015	UK	Public Health England
H153840518	PT U311	Human	2015	UK	Public Health England
H153840520	Untypable	Human	2015	UK	Public Health England
H153940639	PT U302	Human	2015	UK	Public Health England
H153960420	DT 104	Human	2015	UK	Public Health England
H153960432	DT 101	Human	2015	UK	Public Health England
H153980204	DT 195	Human	2015	UK	Public Health England
H154040409	DT 1	Human	2015	UK	Public Health England
H154060438	RDNC	Human	2015	UK	Public Health England
H154080147	DT 15	Human	2015	UK	Public Health England
H154100393	Untypable	Human	2015	UK	Public Health England
H154120570	DT 120	Human	2015	UK	Public Health England
H154140372	DT 120	Human	2015	UK	Public Health England
H154160457	DT 193	Human	2015	UK	Public Health England
H154220780	DT 35	Human	2015	UK	Public Health England
H154220788	Untypable	Human	2015	UK	Public Health England
H154240483	DT 120	Human	2015	UK	Public Health England
H154280562	PT U311	Human	2015	UK	Public Health England
H154300382	PT U329	Human	2015	UK	Public Health England
H154320744	DT 41	Human	2015	UK	Public Health England
H154320745	DT 1	Human	2015	UK	Public Health England
H154360432	DT 1	Human	2015	UK	Public Health England

H154360435	DT 193	Human	2015	UK	Public Health England
H154460357	DT 193	Human	2015	UK	Public Health England
H154500452	None	Human	2015	UK	Public Health England
H154520625	None	Human	2015	UK	Public Health England
H154520647	None	Human	2015	UK	Public Health England
H154520679	None	Human	2015	UK	Public Health England
H154620518	None	Human	2015	UK	Public Health England
H154660433	None	Human	2015	UK	Public Health England
H154700345	None	Human	2015	UK	Public Health England
H154720547	None	Human	2015	UK	Public Health England
H154720549	None	Human	2015	UK	Public Health England
H154720556	None	Human	2015	UK	Public Health England
H154760207	None	Human	2015	UK	Public Health England
H154780417	None	Human	2015	UK	Public Health England
H154800359	None	Human	2015	UK	Public Health England
H154820551	None	Human	2015	UK	Public Health England
H154920516	None	Human	2015	UK	Public Health England
H154940333	None	Human	2015	UK	Public Health England
H154960546	None	Human	2015	UK	Public Health England
H154960760	None	SEWER SWAB	2015	UK	Public Health England
H155020271	None	OLD SPRAY BOTTLE (FORMAL SAMPLE)	2015	UK	Public Health England

H155060512	None	Human	2015	UK	Public Health England
H155060514	None	Human	2015	UK	Public Health England
H155120548	None	Human	2015	UK	Public Health England
H155140536	None	DIRTY USED OVEN CLOTH-FORMAL SAMPLE	2015	UK	Public Health England
H155140539	None	DIRTY USED OVEN CLOTH-FORMAL SAMPLE	2015	UK	Public Health England
H155140542	None	Human	2015	UK	Public Health England
H155160053	None	Human	2015	UK	Public Health England
H155220490	None	Human	2015	UK	Public Health England
H155360251	None	Human	2015	UK	Public Health England
H160120678	None	Human	2016	UK	Public Health England
H160140786	None	Human	2016	UK	Public Health England
H160160247	None	Human	2016	UK	Public Health England
H160220779	None	SWAB OF WHB	2016	UK	Public Health England
H160220781	None	Human	2016	UK	Public Health England
H160240656	None	Human	2016	UK	Public Health England
H160300259	None	Human	2016	UK	Public Health England
S0337107	DT193	PIG	2007	UK	Petrovska et al., (2016)
S0292307	DT193	CATTLE	2007	UK	Petrovska et al., (2016)

S0344705	DT193	HORSE	2005	UK	Petrovska et al., (2016)
L0173006	DT193	PIG	2006	UK	Petrovska et al., (2016)
S0565506	DT193	CATTLE	2006	UK	Petrovska et al., (2016)
S0657807	DT193	CAT	2007	UK	Petrovska et al., (2016)
S0806007	193	PIG	2007	UK	Petrovska et al., (2016)
L0064707	DT193	PIG	2007	UK	Petrovska et al., (2016)
S0509207	RDNC	DOG	2007	UK	Petrovska et al., (2016)
S0344408	DT193	SHEEP	2008	UK	Petrovska et al., (2016)
L0085709	DT120	DOG	2009	UK	Petrovska et al., (2016)
S0354909	DT193	CATTLE	2009	UK	Petrovska et al., (2016)
L0004109	DT193	DOG	2009	UK	Petrovska et al., (2016)
S0387409	DT193	CATTLE	2009	UK	Petrovska et al., (2016)
S0433209	DT193	PIG	2009	UK	Petrovska et al., (2016)
L01001-10	120	CHICKEN	2010	UK	Petrovska et al., (2016)
S01569-10	120	PIG	2010	UK	Petrovska et al., (2016)
S04797-08	191a	BUSTARD	2008	UK	Petrovska et al., (2016)
S00130-09	7	CAT	2009	UK	Petrovska et al., (2016)
1334-1997	ND	PIGS	1997	UK	Petrovska et al., (2016)
L00938-09	21	PIG	2009	UK	Petrovska et al., (2016)
L01176-08	ND	PIG	2008	UK	Petrovska et al., (2016)
L01189-08	ND	PIG	2008	UK	Petrovska et al., (2016)
S00814-10	193	CHICKEN	2010	UK	Petrovska et al., (2016)

L1101-10	ND	CHICKEN	2010	UK	Petrovska et al., (2016)
S4812-10	ND	PIG rearing	2010	UK	Petrovska et al., (2016)
S03445-08	193	SHEEP	2008	UK	Petrovska et al., (2016)
S07300-05	ND	CATTLE	2005	UK	Petrovska et al., (2016)
S00065-06	ND	CATTLE	2006	UK	Petrovska et al., (2016)
S01364-10	193	CATTLE	2010	UK	Petrovska et al., (2016)
4824-10	ND	PIG	2010	UK	Petrovska et al., (2016)
4797-10	ND	PRODMEAT	2010	UK	Petrovska et al., (2016)
S3659-10	ND	CHICKEN	2010	UK	Petrovska et al., (2016)
S03113-10	193	PIG	2010	UK	Petrovska et al., (2016)
S04698-09	193	CATTLE	2009	UK	Petrovska et al., (2016)
S00176-09	RDNC	PIG	2009	UK	Petrovska et al., (2016)
L01730-06	ND	PIG	2006	UK	Petrovska et al., (2016)
S02909-08	193	PIG	2008	UK	Petrovska et al., (2016)
S05893-09	193	PIG	2009	UK	Petrovska et al., (2016)
S05894-09	193	PIG	2009	UK	Petrovska et al., (2016)
H100120548	ND	Human	2010	UK	Petrovska et al., (2016)
H100800267	ND	Human	2010	UK	Petrovska et al., (2016)
H102120667	DT 193	Human	2010	UK	Petrovska et al., (2016)
H10234093302	ND	Human	2010	UK	Petrovska et al., (2016)
H103260370	ND	Human	2010	UK	Petrovska et al., (2016)
H105260826	DT 193	Human	2010	UK	Petrovska et al., (2016)

H105280433	DT 193	Human	2010	UK	Petrovska et al., (2016)
S680/10	ND	Human	ND	UK	Petrovska et al., (2016)
H103920583	ND	Human	2010	UK	Petrovska et al., (2016)
H1041406001	ND	Human	2010	UK	Petrovska et al., (2016)
H104240404	ND	Human	2010	UK	Petrovska et al., (2016)
H104680513	DT 193	Human	2010	UK	Petrovska et al., (2016)
H105000301	DT 193	Human	2010	UK	Petrovska et al., (2016)
45/16	ND	PIG	ND	UK	Petrovska et al., (2016)
2200/2	ND	PIG	ND	UK	Petrovska et al., (2016)
2448/2	ND	CATTLE	ND	UK	Petrovska et al., (2016)
496/10	ND	CATTLE	2010	UK	Petrovska et al., (2016)
1686/1	ND	Human	ND	UK	Petrovska et al., (2016)
1790/1	ND	Turkey	ND	UK	Petrovska et al., (2016)
242/2	ND	Shellfish	ND	UK	Petrovska et al., (2016)
2617/20	ND	PIG	ND	UK	Petrovska et al., (2016)
1948/2	ND	PIG	ND	UK	Petrovska et al., (2016)
1693/1	ND	Human	ND	UK	Petrovska et al., (2016)
3046/11	ND	CATTLE	ND	UK	Petrovska et al., (2016)
1365/1	ND	ND	ND	UK	Petrovska et al., (2016)
2841/2	ND	Human	ND	UK	Petrovska et al., (2016)
692/26	ND	PIG	ND	UK	Petrovska et al., (2016)
2117/2	ND	Shellfish	ND	UK	Petrovska et al., (2016)

629/2	ND	Swine	ND	UK	Petrovska et al., (2016)
2223/2	ND	PIG	ND	UK	Petrovska et al., (2016)
H07 246 0339	193	Human	2007	UK	Petrovska et al., (2016)
H07 362 0321	U302L	Human	2007	UK	Petrovska et al., (2016)
H07 166 0082	193	Human	2007	UK	Petrovska et al., (2016)
H07 182 0182	21 variant	Human	2007	UK	Petrovska et al., (2016)
H07 230 0280	193	Human	2007	UK	Petrovska et al., (2016)
H07 246 0338	120	Human	2007	UK	Petrovska et al., (2016)
H07 276 0382	193	Human	2007	UK	Petrovska et al., (2016)
H07 338 0264	120	Human	2007	UK	Petrovska et al., (2016)
H07 394 0379	120	Human	2007	UK	Petrovska et al., (2016)
SSI_AA527		Human	2014	Denmark	STATENS SERUM INSTITUT
17EP001816		Human	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
17EP001883		Human	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
17EP002422		Human	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
2017-22-798-13		Environment	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
2017-01-4449-1		Environment	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
2017-01-4449		Environment	2017	Norway	NORWEGIAN INSTITUTE OF PUBLIC HEALTH

Vulture-STm-CyP-1		wild vulture	2015	Spain	FISABIO - Public Health
Vulture-STm-CyP-3		wild vulture	2015	Spain	FISABIO - Public Health
Vulture-STm-CyP-4		wild vulture	2015	Spain	FISABIO - Public Health
Vulture-STm-CyP-5		wild vulture	2015	Spain	FISABIO - Public Health
SSI_AA811		Food	2013	Denmark	STATENS SERUM INSTITUT
Vulture-STm-CyP-7		wild vulture	2015	Spain	FISABIO - Public Health
Vulture-STm-CyP-10		wild vulture	2015	Spain	Fisabio
Vulture-STm-CyP-12		wild vulture	2015	Spain	FISABIO - Public Health
ERS3559641		Livestock	2016	China	YANGZHOU UNIVERSITY
ERR3581184		Environment	2019	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581195		Environment	2019	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581381		Environment	2017	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
SAL_ZA9125AA		Food	2013	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581443		Food	2018	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581506		Food	2019	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
STM110		Human	2013	Italy	IZSLER

ERR3581570		Food	2017	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581757		Food	2018	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
ERR3581911		Environment	2018	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
NA		NA	NA	NA	NA
FD01846628		missing	NA	Europe	EI
NA		NA	NA	NA	NA
FD01846635		missing	NA	Europe	EI
NA		NA	NA	NA	NA
NA		NA	NA	NA	NA
FD01846658		SAMEA6143436	2013	NONE	EI
Typhimurium		Poultry	2012	Poland	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU), Denmark
NA		NA	NA	NA	NA
FD01846674		SAMEA6143450	2013	NONE	EI
FD01846676		missing	NA	Europe	EI
FD01846682		SAMEA6143457	2013	NONE	EI
FD01846690		SAMEA6143465	2013	NONE	EI
FD01847436		SAMEA6143939	2013	South America	EI
FD01847440		SAMEA6143943	2013	South America	EI
FD01847949		missing	2014	South America	EI
FD01848038		missing	2015	South America	EI

FD01848055		missing	2016	South America	EI
monofasisk		Poultry	2012	Poland	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU), Denmark
FD01848263		SAMEA6144432	2005	Europe	EI
FD01849113		missing	2013	Africa	EI
FD01849278		missing	2011	Africa	EI
FD01849283		missing	2007	Africa	EI
FD01849290		missing	2001	Africa	EI
FD01874779		missing	2008	Africa	EI
FD01875900		missing	2011	Africa	EI
FD01875908		missing	2011	Africa	EI
NA		NA	NA	NA	NA
2011-60-538-1		Livestock	2011	Denmark	Technical University of Denmark
monofasisk		Poultry	2012	Poland	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU), Denmark
2011-60-176-1		Livestock	2011	Denmark	Technical University of Denmark
S368		missing	2014	Portugal	Animal and Plant Health Agency (UK)
SAL-18-VL-OH-ON-0043		Companion Animal	2018	Canada	FDA Center for Food Safety and Applied Nutrition
NC_WHO_S096		Human	2012	Lebanon	FDA Center for Food Safety and

					Applied Nutrition
VL-0618-5		Companion Animal	2018	Cambodia	FDA Center for Food Safety and Applied Nutrition
AG19-0312		Companion Animal	2018	United States	FDA Center for Food Safety and Applied Nutrition
Thai*Salmonella*_541013		Livestock	2011	Thailand	University of Bath
Thai*Salmonella*_543009		Companion Animal	2012	Thailand	NA
Thai*Salmonella*_193		Livestock	2013	Thailand	University of Bath
monofasisk		Livestock	2012	Poland	University of Bath
Thai*Salmonella*_60		Companion Animal	2013	Thailand	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU), Denmark
UNAM2018123_Sa_AN13		Livestock	2018	Mexico	University of Bath
PNCS005052		Human	2018	Canada	NA
815513		Human	2019	United Kingdom	SENASICA
SX17G598		Human	2017	China	National Microbiology Laboratory
SX17G597		Human	2017	China	Public Health England
SX17G596		Human	2017	China	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION

SX17G595		Human	2017	China	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
SX17G594		Human	2017	China	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
monofasisk		Livestock	2013	Germany	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
SX17G593		Human	2017	China	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
SX17G592		Human	2017	China	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU)
SX17G590		Human	2017	China	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
826239		Human	2019	United Kingdom	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
AZ_TG68559		Livestock	2010	United States	CHINESE PLA CENTER FOR DISEASE CONTROL AND PREVENTION
FNE0135		ND	2013	China	Public Health England

2010K-2457		Human	2010	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
2011K-0863		ND	2011	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
2012K-0073		Human	2011	United States	Centers for Disease Control and Prevention - PulseNet
H120260565		Human	2012	United Kingdom	Centers for Disease Control and Prevention - PulseNet
NA		NA	NA	NA	Centre for Diease Control and Prevention - Enteric Diseases Laboratory Branch
WAPHL_SAL-A00665		Environment	2014	United States	Public Health England - Gastrointestinal Bacteria Reference Unit
CFSAN024425		Swine	2008	Belgium	NA
PNUSAS000227		ND	2013	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)

71411		Food	2013	United Kingdom	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
PNUSAS000226		ND	2013	United States	Centers for Disease Control and Prevention - PulseNet
CFSAN031352		Livestock	2013	Thailand	Public Health England - Gastrointestinal Bacteria Reference Unit
CFSAN031337		Livestock	2013	Thailand	Centers for Disease Control and Prevention - PulseNet
PNUSAS000405		Human	2014	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
PNUSAS000440		ND	2015	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
CFSAN035087		Human	2015	United States	Centers for Disease Control and Prevention - PulseNet
STM206		Human	2014	Italy	Centers for Disease Control and Prevention

PNUSAS000613		Livestock	2015	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
1731		Human	2010	Australia	IZSLER
1732		Human	2010	Australia	Centers for Disease Control and Prevention - PulseNet
1733		Human	2010	Australia	University of New South Wales
1730		Human	2011	Australia	University of New South Wales
PNUSAS000920		Human	2015	United States	University of New South Wales
CVM N51306		Livestock	2013	United States	University of New South Wales
2015K-0418		Human	2014	United States	Centers for Disease Control and Prevention - PulseNet
FSIS1503463		Poultry	2015	United States	Food and Drug Administration (US)
ADRDL-15-5149		Livestock	2015	United States	CDC
LB-7		Poultry	2010	Italy	Department of Agriculture - Food Safety and Inspection Service (US)
S01413-16		Wild Animal	2016	United Kingdom	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)

Typhimurium 135497		Human	2010	Israel	National Institute of Infectious Diseases - Japan
FSIS1503558		Livestock	2015	United States	Animal and Plant Health Agency (APHA)
R9_3269_R2		Human	2008	United States	Sheba Medical Center
R9_3270_R2		Human	2008	United States	Department of Agriculture - Food Safety and Inspection Service (US)
CFSAN045348		Livestock	2014	Thailand	
FSIS1605973		Livestock	2016	United States	Cornell University
CVM N57359F		Porcine	2014	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
FSIS1606261		Livestock	2016	United States	Department of Agriculture - Food Safety and Inspection Service (US)
CFSAN045335		Livestock	2013	Thailand	FDA
CFSAN045319		Livestock	2013	Thailand	Department of Agriculture - Food Safety and Inspection Service (US)
S05355-15		Wild Animal	2015	United Kingdom	Food and Drug Administration - Center for Food Safety and Applied

					Nutrition (US)
FDA00003755		Environment	2012	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
FSIS1606397		Livestock	2016	United States	Animal and Plant Health Agency (APHA)
PNUSAS002132		ND	2016	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
FSIS1606634		Porcine	2016	United States	Department of Agriculture - Food Safety and Inspection Service (US)
CVM N40368		Livestock	2012	United States	Center for Diease Control and Prevention - Enteric Diseases Laboratory Branch (US)
FDA00010558		Environment	2016	Mexico	USDA
FSIS1607596		Livestock	2016	United States	Food and Drug Administration (US)
FSIS1607729		Livestock	2016	United States	Food and Drug Administration - Center for Food Safety and Applied

					Nutrition (US)
PNUSAS004211		ND	2011	United States	Department of Agriculture - Food Safety and Inspection Service (US)
PNUSAS004213		ND	2011	United States	Department of Agriculture - Food Safety and Inspection Service (US)
ERR1815499_22		Human	2016	Ireland	Center for Diease Control and Prevention - Enteric Diseases Laboratory Branch (US)
PNUSAS003936		ND	2009	United States	Center for Diease Control and Prevention - Enteric Diseases Laboratory Branch (US)
PNUSAS003935		Human	2008	United States	
SRR4897062		Foodproduct	2016	United States	Center for Diease Control and Prevention - Enteric Diseases Laboratory Branch (US)
SRR5045027		Foodproduct	2016	United States	1
967864		Human	NA	Canada	USDA-FSIS
L1874		Human	2008	Australia	USDA-FSIS
FSIS1710414		Livestock	2017	United States	National Microbiology Laboratory

OH-17-537		ND	2017	United States	University of New South Wales
FSIS1710449		Livestock	2017	United States	Department of Agriculture - Food Safety and Inspection Service (US)
FSIS1710787		Poultry	2017	United States	Food and Drug Administration - Center for Food Safety and Applied Nutrition (US)
ERR1857451		Human	2017	Ireland	Department of Agriculture - Food Safety and Inspection Service (US)
PNUSAS010064		ND	2012	United States	Department of Agriculture - Food Safety and Inspection Service (US)
FSIS1710756		Livestock	2017	United States	University Hospital Galway
FDA00000366		Food	2003	Portugal	Centers for Disease Control and Prevention - PulseNet
CCK-B-038		Poultry	2016	Barbados	Department of Agriculture - Food Safety and Inspection Service (US)
11-1166		Human	2011	Belgium	Food and Drug Administration - Center for Food Safety and Applied

					Nutrition (US)
37956		Human	2012	Belgium	
PNUSAS032218		missing	2012	United States	Scientific Institute of Public Health (WIV-ISP)
FDA00000385		Feed	2003	Portugal	Scientific Institute of Public Health (WIV-ISP)
486527		Human	2018	United Kingdom	Centers for Disease Control and Prevention Enteric Diseases Laboratory Branch
391748		Human	2017	United Kingdom	FDA Center for Food Safety and Applied Nutrition
SAL_UA1879AA		Livestock	2015	United Kingdom	PHE
442694		Human	2017	United Kingdom	Public Health England
431399		Human	2017	United Kingdom	Animal and Plant Health Agency (APHA)
N09-0923		Human	2009	Switzerland	
N09-0868		Human	2009	Switzerland	Public Health England
N08-2826		Human	2008	Switzerland	CFSAN
N10-2374		Human	2010	Switzerland	CFSAN
N10-1200		missing	2010	Switzerland	CFSAN
N10-0623		missing	2010	Switzerland	CFSAN
N07-563		Human	2007	Switzerland	CFSAN
N07-851		Human	2007	Switzerland	CFSAN
ISZ-13		Pig Meat	2010	Italy	CFSAN

N07-1209		Human	2007	Switzerland	CFSAN
N06-1830		Livestock	2006	Switzerland	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
399936		Human	2017	United Kingdom	CFSAN
N11-1381		Human	2011	Switzerland	CFSAN
N14-1016		Human	2014	Switzerland	Public Health England
N14-2154		Food	2014	Switzerland	CFSAN
N13-0144		Human	2013	Switzerland	CFSAN
N13-0004		Human	2013	Switzerland	CFSAN
470218		Human	2018	United Kingdom	CFSAN
N16-1247		Human	2016	Switzerland	CFSAN
DTU_1045		Fresh Cured Pork Meat	2015	Italy	Public Health England
N15-1290		Poultry	2015	Switzerland	CFSAN
N17-0794		Poultry	2017	Switzerland	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
N16-1477		missing	2016	Switzerland	CFSAN
N17-0346		Human	2017	Switzerland	CFSAN
N11-2679		Human	2011	Switzerland	CFSAN
N17-0904		Human	2017	Switzerland	CFSAN
542015		Human	2018	United Kingdom	CFSAN
475839		Human	2018	United Kingdom	CFSAN

493485		Human	2018	United Kingdom	Public Health England
NC_WHO_S213		Human	2014	Lebanon	PHE
3399-2012		Human	2012	Italy	Public Health England
NC_WHO_S047		Human	2011	Lebanon	FDA Center for Food Safety and Applied Nutrition
NC_WHO_S050		Human	2011	Lebanon	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
Sa_V42		Environment	2017	Mexico	FDA Center for Food Safety and Applied Nutrition
E2018009940		Livestock	2018	United States	FDA Center for Food Safety and Applied Nutrition
E2018009938		Livestock	2018	United States	SENASICA
11447/15		Dolphin	2015	Italy	FDA Center for Food Safety and Applied Nutrition
78938/17		Dolphin	2017	Italy	FDA Center for Food Safety and Applied Nutrition
SL3-12		missing	NA	Russia	Istituto Zooprofilattico Sperimentale del Piemonte Liguria e Valle d'Aosta
591934		Human	2018	United Kingdom	Istituto Zooprofilattico Sperimentale del Piemonte Liguria e Valle d'Aosta

SAL-18-VL-OH-ON-0013		Livestock	2018	Canada	Central Research Institute for Epidemiology
4007-2011		Meleagris	2011	Italy	Public Health England
R15.0289		Human	2015	Taiwan	FDA Center for Food Safety and Applied Nutrition
R15.0265		Human	2015	Taiwan	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
R15.0285		Human	2015	Taiwan	Centers for Disease Control, Taiwan
E028		Poultry	2013	Taiwan	Centers for Disease Control, Taiwan
E027		Poultry	2013	Taiwan	Centers for Disease Control, Taiwan
R15.0305		Human	2015	Taiwan	Centers for Disease Control, Taiwan
CS109		Poultry	2011	Taiwan	Centers for Disease Control, Taiwan
R15.0446		Human	2015	Taiwan	Centers for Disease Control, Taiwan
R15.0414		Human	2015	Taiwan	Centers for Disease Control, Taiwan
CB12.026		Human	2012	Taiwan	Centers for Disease Control, Taiwan
15Q003631		Livestock	2015	France	Centers for Disease Control, Taiwan
R16.3277		Human	2016	Taiwan	Centers for Disease

					Control, Taiwan
R14.0686		Human	2014	Taiwan	
R14.0884		Human	2014	Taiwan	Centers for Disease Control, Taiwan
PS17		Livestock	2011	Taiwan	Centers for Disease Control, Taiwan
NL09.115		Human	2009	Taiwan	Centers for Disease Control, Taiwan
PS32		Livestock	2012	Taiwan	Centers for Disease Control, Taiwan
R15.0059		Human	2015	Taiwan	Centers for Disease Control, Taiwan
R13.1996		Human	2013	Taiwan	Centers for Disease Control, Taiwan
SAL-18-VL-LA-KS-0017		Livestock	2018	United States	Centers for Disease Control, Taiwan
AHD24		Human	2011	United States	Centers for Disease Control, Taiwan
ERR2023105		Human	2017	Ireland	FDA Center for Food Safety and Applied Nutrition
AHD3		Human	2009	United States	FDA Center for Food Safety and Applied Nutrition
H2 - 038		Livestock	2012	Thailand	University Hospital Galway
H2 - 020		Livestock	2012	Thailand	FDA Center for Food Safety and Applied Nutrition

H2 - 014		Food	2012	Thailand	FDA Center for Food Safety and Applied Nutrition
H2 - 012		Livestock	2012	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 136		Companion Animal	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
H2 - 111		Livestock	2012	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 135		Aquatic	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 142		Aquatic	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 121		ND	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
DTU2016_1375_PRJproj1048_*Salmonella*_monofasisk_F14_1101_1		Livestock	2014	Germany	FDA Center for Food Safety and Applied Nutrition
H1 - 134		Environment	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 137		ND	2010	Thailand	Technical University of Denmark - Centre for Genomic Epidemiology, National Food Institute

H2 - 070		Livestock	2012	Thailand	FDA Center for Food Safety and Applied Nutrition
H1 - 143		Environment	2010	Thailand	FDA Center for Food Safety and Applied Nutrition
AMC 238		Aquatic	2015	Spain	FDA Center for Food Safety and Applied Nutrition
678048		Human	2019	United Kingdom	FDA Center for Food Safety and Applied Nutrition
NC_WHO_S073		Human	2012	Lebanon	FDA/CFSAN
FDA00013864		Environment	2018	China	Public Health England
697899		Human	2019	United Kingdom	FDA Center for Food Safety and Applied Nutrition
NC_WHO_S302		Human	2012	Lebanon	FDA Center for Food Safety and Applied Nutrition
DTU2016_872_PRJ1048_*Salmonella*_monofasisk_2012_60_1628_1_1		Livestock	2012	Germany	Public Health England
01/185/LJ		Human	2019	Slovenia	FDA Center for Food Safety and Applied Nutrition
45/531/CE		Human	2019	Slovenia	Centre for Genomic Epidemiology, National Food Institute, Technical University of

					Denmark (DTU)
3B/248/KP		Human	2019	Slovenia	NATIONAL LABORATORY FOR HEALTH, ENVIRONMENT AND FOOD
738652		Human	2019	United Kingdom	NATIONAL LABORATORY FOR HEALTH, ENVIRONMENT AND FOOD
769122		Human	2019	United Kingdom	NATIONAL LABORATORY FOR HEALTH, ENVIRONMENT AND FOOD
NC_WHO_S048		Human	2011	Lebanon	Public Health England
779879		Human	2019	United Kingdom	Public Health England
VNB1166		Human	2010	Vietnam	FDA Center for Food Safety and Applied Nutrition
VNS10314		Human	2010	Vietnam	Public Health England
VNB455		Human	2009	Vietnam	ND
DTU2016_874		Livestock	2012	France	ND
VNB712		Human	2009	Vietnam	ND
VNS20235		Human	2009	Vietnam	Centre for Genomic Epidemiology, National Food Institute, Technical University of Denmark (DTU)
VNB1264		Human	2011	Vietnam	ND
VNS30144		Human	2009	Vietnam	ND

VNB1701		Human	2012	Vietnam	ND
VNB176		Human	2008	Vietnam	ND
VNB617		Human	2009	Vietnam	ND
VNB1479		Human	2011	Vietnam	ND
VNB1140		Human	2010	Vietnam	ND
71_G_450		Chicken	2012	Vietnam	ND
ERR2124052		Human	2017	Ireland	ND
71_H_035		Pig	2012	Vietnam	ND
VNB1779		Human	2012	Vietnam	University Hospital Galway
VNB2315		Human	2013	Vietnam	ND
Hue_11		Human	2009	Vietnam	ND
CT69_2		Chicken	2013	Vietnam	ND
71_H_455		Pig	2012	Vietnam	ND
71_H_455		Pig	2012	Vietnam	ND
L-4233		Bos taurus	2014	Japan	ND
L-4259		Bos taurus	2013	Japan	ND
17-71865_S18		Human	2016	Italy	National Institute of Infectious Diseases - Japan
L-4257		Bos taurus	2013	Japan	National Institute of Infectious Diseases - Japan
L-3841		Sus scrofa domesticus	2009	Japan	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
L-3844		Sus scrofa domesticus	2002	Japan	National Institute of Infectious Diseases - Japan
L-3835		Homo sapiens	2007	Japan	National Institute of Infectious Diseases - Japan
L-3846		Sus scrofa domesticus	2008	Japan	National Institute of Infectious Diseases - Japan

L-4126		Bos taurus	1998	Japan	National Institute of Infectious Diseases - Japan
17-71882_S32		Swine	2016	Italy	National Institute of Infectious Diseases - Japan
L-4127		Bos taurus	1998	Japan	
17-71888_S37		Buffalo	2016	Italy	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
ERR2173689		Human	2017	Ireland	National Institute of Infectious Diseases - Japan
SSI-AC256		Human	2009	Denmark	ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLE VENEZIE
L-4234		Sus scrofa domesticus	2014	Japan	University Hospital Galway
L-4261		Bos taurus	2014	Japan	STATENS SERUM INSTITUT
ERR2193125		Human	2017	Ireland	ND
L-3838		Livestock	2008	Japan	ND
L-3837		Bos taurus	2008	Japan	University Hospital Galway
L-4071		Bos taurus	2013	Japan	National Institute of Infectious Diseases - Japan
ERR2193191		Human	2017	Ireland	National Institute of Infectious Diseases - Japan
16-SA00669		food	2016	Germany	National Institute of Infectious

					Diseases - Japan
16-SA00754		animal	2016	Germany	University Hospital Galway
13-SA02497		animal	2013	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
14-SA02723		animal	2014	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
15-SA00407		animal	2015	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
15-SA02945		food	2015	Germany	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
L00541-16		Livestock	2016	United Kingdom	ND
S00160-16		Livestock	2016	United Kingdom	FEDERAL INSTITUTE FOR RISK ASSESSMENT (BFR)
S05117-15		Poultry	2015	United Kingdom	Animal and Plant Health Agency (APHA)
DTU2016-418		Poultry	2014	Italy	Animal and Plant Health Agency (APHA)
136-16		Human	2016	Poland	Animal and Plant Health Agency (APHA)
201-15		Human	2015	Poland	IZSLT (Istituto Zooprofilattico Sperimentale del Lazio e la Toscana)

205-14		Human	2014	Poland	National Institute of Public Health - National Institute of Hygiene
272-15		Human	2015	Poland	National Institute of Public Health - National Institute of Hygiene
288-15		Human	2015	Poland	National Institute of Public Health - National Institute of Hygiene
303-14		Human	2014	Poland	National Institute of Public Health - National Institute of Hygiene
332-14		Human	2014	Poland	National Institute of Public Health - National Institute of Hygiene
51-15		Human	2015	Poland	National Institute of Public Health - National Institute of Hygiene
LB-22		Livestock	2011	Italy	National Institute of Public Health - National Institute of Hygiene
51-16		Human	2016	Poland	National Institute of Public Health - National Institute of Hygiene
56-16		Human	2016	Poland	National Institute of Infectious Diseases - Japan

69-16		Human	2016	Poland	National Institute of Public Health - National Institute of Hygiene
86-15		Human	2015	Poland	National Institute of Public Health - National Institute of Hygiene
86-16		Food	2016	Poland	National Institute of Public Health - National Institute of Hygiene
350-13		Human	2013	Poland	National Institute of Public Health - National Institute of Hygiene
297-13		Human	2013	Poland	ND
181-13		Human	2013	Poland	National Institute of Public Health - National Institute of Hygiene
74675_2014_salm		Livestock	2015	Italy	National Institute of Public Health - National Institute of Hygiene
ERR2215656		Human	2017	Ireland	National Institute of Public Health - National Institute of Hygiene
LB-21		Poultry	2012	Italy	IZSLT (Istituto Zooprofilattico Sperimentale del Lazio e la Toscana)
ERR2698814		Companion Animal	2016	Cambodia	University Hospital Galway

ERR2698816		Companion Animal	2016	Cambodia	National Institute of Infectious Diseases - Japan
53_Typhimurium2_p		Feed	2008	Netherlands	Institut Pasteur
58_monophasic Typhimurium1_h		Human	2012	Netherlands	Institut Pasteur
59_monophasic Typhimurium2_h		Human	2012	Netherlands	RIVM - National Institute for Public Health and the Environment
60_monophasic Typhimurium3_h		Human	2012	Netherlands	RIVM - National Institute for Public Health and the Environment
61_monophasic Typhimurium1_p		Livestock	2009	Netherlands	RIVM - National Institute for Public Health and the Environment
62_monophasic Typhimurium2_h		Livestock	2009	Netherlands	RIVM - National Institute for Public Health and the Environment
63_monophasic Typhimurium3_p		Livestock	2009	Netherlands	RIVM - National Institute for Public Health and the Environment
0608F12523		Human	2006	Denmark	RIVM - National Institute for Public Health and the Environment
LB-11		Livestock	2011	Italy	RIVM - National Institute for Public Health and the Environment

0511R6988		Human	2005	Denmark	STATENS SERUM INSTITUT
0501M38866		Human	2005	Denmark	National Institute of Infectious Diseases - Japan
1805S63638		Human	2018	Denmark	STATENS SERUM INSTITUT
1705F15563		Human	2017	Denmark	STATENS SERUM INSTITUT
1503F57888		Human	2015	Denmark	STATENS SERUM INSTITUT
11CEB2251SAL		sheep goat - dairy products - cheese	2011	France	STATENS SERUM INSTITUT
11CEB3015SAL		sheep goat - dairy products - cheese	2011	France	STATENS SERUM INSTITUT
11CEB6557SAL		pigs - pork back	2011	France	ND
2011_02155		human - faeces	2011	France	ND
2011_02228		human - faeces	2011	France	ANSES
SSI_AA378		Human	2013	Denmark	ND
2011_08647		human	2011	France	ND
2011_08648		human	2011	France	STATENS SERUM INSTITUT
2014LSAL04847		Environment	2014	France	ANSES
2014_02059		Human	2014	France	ANSES
1902T41853		Human	2019	Denmark	ANSES
06-01900		Human	2006	Germany	ND
16-04913		Human	2016	Germany	STATENS SERUM INSTITUT

16-04940		Human	2016	Germany	ROBERT KOCH-INSTITUTE
17EP001571		Human	2017	Norway	ROBERT KOCH-INSTITUTE
17EP001668		Human	2017	Norway	ROBERT KOCH-INSTITUTE
ERR1828951_108		Human	2015	Ireland	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
ERR1837626_122		Human	2015	Ireland	NORWEGIAN INSTITUTE OF PUBLIC HEALTH
ERR1806844_22		Human	2016	Ireland	ND
ERR1815681_40		Human	2016	Ireland	ND
ERR1816626_49		Human	2016	Ireland	ND
ERR1802428_7		Human	2016	Ireland	ND
ERR1817394_71		Human	2016	Ireland	ND
ERR1817513_76		Human	2016	Ireland	ND
ERR1817536_81		Human	2016	Ireland	ND
ERR1823827_94		Human	2014	Ireland	ND
ERR1823828_95		Human	2014	Ireland	ND
0153D		Pig	2012	Ireland	ND
0197B		Pig	2012	Ireland	ND
0286B		Pig	2012	Ireland	ND
0309C		Pig	2012	Ireland	ND
1792A		Pig	2012	Ireland	ND
3500B		Pig	2013	Ireland	ND
3524A		Pig	2013	Ireland	ND
0329B		Pig	2012	Ireland	ND
0591C		Pig	2012	Ireland	ND
1446A		Pig	2012	Ireland	ND
2002A		Pig	2012	Ireland	ND
2207A		Pig	2013	Ireland	ND
2213A		Pig	2013	Ireland	ND

2985A		Pig	2013	Ireland	ND
2994A		Pig	2013	Ireland	ND
2999A		Pig	2013	Ireland	ND
3083C		Pig	2013	Ireland	ND
3836A		Pig	2013	Ireland	ND
3845A		Pig	2013	Ireland	ND
STY1		Human	2013	Italy	ND
STY102		Human	2012	Italy	ND
STY107		Human	2012	Italy	University of Bologna
STY111		Human	2012	Italy	University of Bologna
STY122		Human	2012	Italy	University of Bologna
STY129		Swine	2012	Italy	University of Bologna
STY136		Swine	2012	Italy	University of Bologna
STY149		Swine	2012	Italy	University of Bologna
STY158		Swine	2013	Italy	University of Bologna
STY189		Swine	2014	Italy	University of Bologna
STY194		Swine	2014	Italy	University of Bologna
STY196		Swine	2014	Italy	University of Bologna
STY199		Swine	2014	Italy	University of Bologna
NA		NA	NA	NA	University of Bologna
STY27		Human	2013	Italy	University of Bologna
STY32		Human	2014	Italy	NA
STY33		Human	2014	Italy	University of Bologna
STY51		Human	2014	Italy	University of Bologna
STY6		Human	2013	Italy	University of Bologna
STY68		Human	2012	Italy	University of Bologna
S04698-09Δ*bar:cat*	DT193	NA	NA	UK	This study

SL1344Δ*copA:aphII*	DT44	NA	NA	NA	This study
4/74:*aphII*	ND	NA	NA	NA	This study (Luke Acton)
4/74 Δphage	ND	NA	NA	NA	Rodwell et al., (2021)
S04698-09:mTmII:*cat*	DT193	NA	NA	NA	This study (Luke Acton)
S04332-09:mTmII:*cat*	DT193	NA	NA	NA	This study (Luke Acton)
L00979-07:mTmII:*cat*	DT193	NA	NA	NA	This study (Luke Acton)
L00745-07:mTmII:*cat*	DT193	NA	NA	NA	This study (Luke Acton)
L01157-10:*aphII*	DT8	NA	NA	NA	This study (Gaetan Thilliez)
L01157-10:*aphII, wzy-*	DT30	NA	NA	NA	This study
L01157-10:*wzy-*	DT30	NA	NA	NA	This study
L01157-10Δ*wjx:cat*	DT8	NA	NA	NA	This study
L01157-10:*aphII*, *wzy-*, nalR	DT30	NA	NA	NA	This study
L01157-10Δ*wjx:cat,aphII,nalR*	DT31	NA	NA	NA	This study

Appendix X: First Author Publications Associated With This Study

frontiers in Microbiology

ORIGINAL RESEARCH
published: 24 May 2019
doi: 10.3389/fmicb.2019.01118

SGI-4 in Monophasic *Salmonella* Typhimurium ST34 Is a Novel ICE That Enhances Resistance to Copper

Priscilla Branchu[1†], Oliver J. Charity[1†], Matt Bawn[1‡], Gaetan Thilliez[1], Timothy J. Dallman[2], Liljana Petrovska[3] and Robert A. Kingsley[1,4]*

[1] Quadram Institute Bioscience, Norwich, United Kingdom, [2] Gastrointestinal Bacteria Reference Unit, National Infection Service, Public Health England, London, United Kingdom, [3] Animal and Plant Health Agency (APHA), Addlestone, United Kingdom, [4] School of Biological Sciences, University of East Anglia, Norwich, United Kingdom

OPEN ACCESS

Edited by:
Giovanna Suzzi,
University of Teramo, Italy

Reviewed by:
Qingping Wu,
Guangdong Institute of Microbiology (CAS), China
Joana Mourão,
University of Coimbra, Portugal
Soraya Chaturongakul,
Mahidol University, Thailand

***Correspondence:**
Robert A. Kingsley
rob.kingsley@quadram.ac.uk

[†]These authors have contributed equally to this work

[‡]**Present address:**
Matt Bawn,
Earlham Institute, Norwich, United Kingdom

Specialty section:
This article was submitted to Food Microbiology, a section of the journal Frontiers in Microbiology

Received: 06 February 2019
Accepted: 03 May 2019
Published: 24 May 2019

Citation:
Branchu P, Charity OJ, Bawn M, Thilliez G, Dallman TJ, Petrovska L and Kingsley RA (2019) SGI-4 in Monophasic Salmonella Typhimurium ST34 Is a Novel ICE That Enhances Resistance to Copper. Front. Microbiol. 10:1118. doi: 10.3389/fmicb.2019.01118

A multi drug resistant *Salmonella enterica* 4,[5],12:i- of sequence type 34 (monophasic *S.* Typhimurium ST34) is a current pandemic clone associated with livestock, particularly pigs, and numerous outbreaks in the human population. A large genomic island, termed SGI-4, is present in the monophasic Typhimurium ST34 clade and absent from other *S.* Typhimurium strains. SGI-4 consists of 87 open reading frames including *sil* and *pco* genes previously implicated in resistance to copper (Cu) and silver, and multiple genes predicted to be involved in mobilization and transfer by conjugation. SGI-4 was excised from the chromosome, circularized, and transferred to recipient strains of *S.* Typhimurium at a frequency influenced by stress induced by mitomycin C, and oxygen tension. The presence of SGI-4 was associated with increased resistance to Cu, particularly but not exclusively under anaerobic conditions. The presence of *silCBA* genes, predicted to encode an RND family efflux pump that transports Cu from the periplasm to the external milieu, was sufficient to impart the observed enhanced resistance to Cu, above that commonly associated with *S.* Typhimurium isolates. The presence of these genes resulted in the absence of Cu-dependent induction of *pco* genes encoding multiple proteins linked to Cu resistance, also present on SGI-4, suggesting that the system effectively limits the Cu availability in the periplasm, but did not affect SodCI-dependent macrophage survival.

Keywords: *Salmonella*, monophasic, integrative conjugative element, SGI-4, copper resistance

INTRODUCTION

Salmonella enterica serovar Typhimurium (*S.* Typhimurium), including monophasic variants, accounts for approximately 25% of all human cases of non-typhoidal *Salmonella* (NTS) infection in Europe, and is widespread in multiple animal reservoirs (Hugas and Beloeil, 2014; Animal and Plant Health Agency [APHA], 2017; Branchu et al., 2018). The epidemiological record of human multidrug-resistant (MDR) *S.* Typhimurium infections in Europe is characterized by successive waves of dominant MDR variants that persist for 10–15 years (Rabsch et al., 2001; Rabsch, 2007). *S.* Typhimurium DT104 emerged around 1990, becoming a globally pandemic clone that affected numerous domesticated and wild animal species (Threlfall, 2000). Subsequently, in 2007, a monophasic *S.* Typhimurium variant (*S.* 4,[5],12:i-) of sequence type 34 (ST34)

emerged in European pig populations and spread globally (hereafter referred to as monophasic *S.* Typhimurium ST34; Hauser et al., 2010; Antunes et al., 2011; Hopkins et al., 2012; Arguello et al., 2013; Mourao et al., 2015; Andres-Barranco et al., 2016; Bonardi, 2017). The mechanisms that drive succession of *S.* Typhimurium variants are not known, but selection by commonly used antibiotics is unlikely since successive variants share similar AMR profiles (Rabsch, 2007): ACSSuT (ampicillin, chloramphenicol, streptomycin, sulphonamide, tetracycline) for *S.* Typhimurium DT104 and ASSuT for monophasic *S.* Typhimurium ST34.

Cu is both an essential nutrient, due to its role as a cofactor in multiple enzymes in all aerobic organisms, and highly toxic due to its ability to displace iron from iron-sulfur clusters in dehydratases (Macomber and Imlay, 2009). To reduce toxicity, bacteria control the amount of free Cu in the cytoplasm and periplasm using transport systems, and by the oxidation of cuprous (Cu^{1+}) to less toxic cupric (Cu^{2+}) ions (Rensing and Grass, 2003). *Escherichia coli*, a close relative of *Salmonella*, encodes multiple transport systems on its chromosome to maintain Cu homeostasis. To overcome the toxicity of Cu, *E. coli* transports Cu from the cytoplasm into the periplasm via a P_{1B}-type ATPase, CopA, and from the cytoplasm and periplasm to the external milieu via a multicomponent Cu RND family efflux pump, CusCFBA, and oxidizes cuprous ions by the action of the multi-Cu oxidase, CueO (Grass and Rensing, 2001; Franke et al., 2003). CueO and CopA, which are co-regulated by the cytosolic CueR, are the primary Cu homeostasis systems active during aerobic growth (Stoyanov et al., 2001), while CusCFBA is important during anaerobic growth, under the transcriptional control of the periplasmic CusRS two component regulator (Outten et al., 2001). In addition, some *E. coli* isolated from Cu-rich environments have additional plasmid-encoded genes *pcoABCDRSE*, that encode several proteins including a multicopper oxidase system which is active in the periplasm, Cu transporters, and a Cu binding protein. The *pco* locus is regulated by a two-component system (PcoR/PcoS) with a sensor kinase that extends into the periplasm (Brown et al., 1995; Rensing and Grass, 2003).

Cu homeostasis in the genus *Salmonella* appears to be fundamentally different from that in *E. coli* due to the deletion of the *cusRSCFBA* genes from the chromosome noted in a number of commonly used lab strains (McClelland et al., 2001; Espariz et al., 2007; Pontel and Soncini, 2009; Fookes et al., 2011). Therefore, although *Salmonella* encode CopA and CueO, the lack of the CusCFBA RND family efflux pump means they lack the ability to transport Cu out of the cell entirely, which is likely to have a significant impact on the distribution of Cu within the cell. However, plasmid-borne *silRSECBAP* genes have been described in a *S.* Typhimurium strain associated with an outbreak in burn patients that had been treated topically with silver nitrate; these genes encode an RND family efflux pump that is closely related to the CusCFBA system and conferred resistance to silver and Cu (McHugh et al., 1975; Gupta et al., 1999). Although the *sil* and *cus* genes are generally absent from the whole genome sequence of reference strains of *Salmonella*, these genes have been reported in multiple distinct monophasic *S.* Typhimurium and *S.* Rissen clones associated primarily with pigs in the past two decades (Mourao et al., 2015; Campos et al., 2016; Mastrorilli et al., 2018).

When the whole genome sequence of monophasic *S.* Typhimurium ST34 was compared with other *S.* Typhimurium whole genome sequence, an 80kb genomic island was identified and designated as *Salmonella* genomic island 4 [SGI-4, note addendum for nomenclature change from SGI-3 (Petrovska et al., 2016)]; this genomic island was present in over 95% of monophasic *S.* Typhimurium ST34 strains from the United Kingdom and Italy, but absent from a diverse collection of *S.* Typhimurium including DT104 (Petrovska et al., 2016). Ancestral reconstruction analysis was consistent with acquisition of SGI-4 by horizontal transfer concomitant with clonal expansion of monophasic *S.* Typhimurium ST34, and rare sporadic loss of the genetic island (Petrovska et al., 2016). SGI-4 contained three clusters of genes predicted to be involved in resistance to Cu and silver or arsenic metal ions (Petrovska et al., 2016).

Here we addressed the hypothesis that SGI-4 is a mobile genetic element (MGE) that encodes multiple metal ion resistance determinants that alter the growth of *Salmonella* in concentrations of Cu relevant to the host and farm environments. Furthermore, we test the hypothesis that transferable Cu and silver resistance genes alter expression of endogenous Cu homeostasis genes that sense Cu levels in the periplasm and that this affects SocCI-mediated survival in macrophages.

MATERIALS AND METHODS

Bacterial Strains and Growth Conditions

Salmonella Typhimurium strain SL1344 was isolated from a calf in 1973 as described previously (Kroger et al., 2012). Monophasic *S.* Typhimurium ST34 strains (S04698-09, S04332-09, L00857-09, S05092-07, S06578-07, and L0938-09), *S.* Typhimurium U288 strains S01960-05, S05968-02 and 11020-1996, *S.* Typhimurium DT104 strains (NCTC13348, 4582-1995 and S00914-05), and two strains that were closely related to strain SL1344 (9115-1996 and 6164-1997) were isolated from various host species and their whole genome sequences were determined as described previously (Mather et al., 2016; Petrovska et al., 2016). For planktonic culture, bacteria were inoculated from single colonies subcultured on LB agar plates into 10ml of growth medium and incubated at 37°C in aerobic (normal atmospheric) conditions, or microaerobic (10% CO_2, 5% H_2, 5% O_2, and 80% N_2) or anaerobic (10% CO_2, 10% H_2, and 80% N_2) conditions using a Whitely A95 Anaerobic Workstation (don whitley scientific). Strains of *S.* Typhimurium or monophasic *S.* Typhimurium in which specific genes were replaced with either the *cat* gene conferring chloramphenicol resistance or the *aph* gene conferring resistance to kanamycin were constructed using allelic exchange methodology as described previously (Datsenko and Wanner, 2000). Briefly, pairs of oligonucleotide primers specific for amplification of the *cat* gene or the *aph* gene from plasmids pKD3 or pKD4 were synthesized with 50 nucleotide sequences on the 3′ end that were identical to sequence immediately proximal to the ATG start and distal to the stop codon of each

gene targeted for deletion (**Supplementary Table S1**). Bacterial strains were routinely cultured in Luria Bertani broth (Oxoid) supplemented with 0.03 mg/l chloramphenicol or 0.05 mg/l kanamycin as appropriate.

Sequence Analysis

A gene model for SGI-4 was constructed using Prokka (Seemann, 2014) with a minimum open reading frame (ORF) size of 100. Annotation of SGI-4 was achieved by aligning ORF sequences to those in the NCBI database using BLASTn to identify genes with the greatest sequence identity.

Phylogenetic trees from whole genome sequence data were constructed and single nucleotide polymorphisms (SNPs) identified in the whole genome sequences by aligning reads using BWA-MEM (Li, 2013), variant calling with Freebayes (Garrison and Marth, 2012) and SNP filtering using vcflib/vcftools (Danecek et al., 2011), combined as a pipeline using Snippy v3.0.[1] Maximum-likelihood trees were constructed using a general time-reversible substitution model with gamma correction for amongst-site rate variation with RAxML v8.0.20 with 1000 bootstraps (Stamatakis, 2006). Representative isolates of *S.* Typhimurium phage types described previously (Petrovska et al., 2016), and whole genome sequences of *S.* Typhimurium isolates from routine clinical surveillance by Public Health England are available in public databases with accession numbers reported in **Supplementary Table S2**.

Candidate SGI-4-like elements (SLEs) were identified in genome assemblies in the NCBI non-redundant sequence database (accessed June 2018) by aligning to SGI-4 excluding the *ars*, *sil*, and *pco* loci, using discontiguous megaBLAST. This analysis identified nucleotide sequences from *Edwardsiella ictaluri* (accession CP001600, 326500..380800) *Erwinia tracheiphila* (accession CP013970, 1856073..1916737), *Enterobacter cloacae* (accession CP012162, 4040884..4152906), *Enterobacter hormaechei* (accession CP010376, 3932000..4028800), *Enterobacter hormaechei* (accession CP012165, 395200..536000), *Pluralibacter gergoviae* (accession CP009450, 1604785..1778603) and *Salmonella* Cubana (accession CP006055, 4214040..4311200). For reconstruction of the phylogeny of putative SLEs, sequence were aligned using clustalW-2.1 (Larkin et al., 2007) to determine sequence identity and conserved regions. A maximum likelihood tree was constructed from aligned nucleotide sequence using RAxML (Stamatakis, 2006). The relationship of SLEs was also investigated by determining the proportion of each SLE aligned by carrying out a pair-wise comparison with discontinuous megaBLAST (Morgulis et al., 2008). Each SLE was used as the query sequence against each SLE as the subject to determine the percent of sequence that aligned and the mean percent nucleotide sequence identity.

For reconstruction of a maximum likelihood tree to investigate the relationship of the SGI-4 encoded and previously described *silRSECBAF* and *cusRSCFBA* genes, sequence extracted from 17 whole genome sequences was aligned using ClustalW-2.1 (Larkin et al., 2007) and a tree constructed using 1000 bootstraps and the GTRCAT substitution matrix with RAxML-8.0.22 software (Stamatakis, 2006). The loci were from 15 Cu homeostasis and silver resistance island (CHASRI) sequences (Staehlin et al., 2016), of which nine were chromosomal and six on plasmids, as well as the originally described *silRSECFBAP* from *S.* Typhimurium plasmid pMG101, and chromosomal *cusSRCFBA* from *E. coli* K-12.

To estimate the distribution of the *silA* gene in 926 *Salmonella* whole genome sequences, one from each eBurst group (similar to serovar groups) and therefore representing the diversity of the genus (Alikhan et al., 2018), query sequence was aligned using BLASTn with an 80% coverage threshold and 2,000 maximum hits. BIGSI (Bradley et al., 2019) was used to query the presence of the *silA* and *invA* genes in over 450,000 bacterial sequence entries in the European Nucelotide Archive (ENA; accessed on March 29, 2019 with an api url search).

Determination of SGI-4 Transfer *in vitro*

In order to provide a convenient selectable marker for the presence of SGI-4, we constructed a strain of monophasic *S.* Typhimurium S04698-09 in which the *bar* gene on SGI-4 was replaced by the *cat* gene (S04698-09 SGI-4 Δ*bar*::*cat*), conferring resistance to chloramphenicol (**Supplementary Table S1**). To provide a selectable marker for the recipient strain, we constructed a strain of *S.* Typhimurium SL1344 in which the *copA* gene was deleted and replaced by an *aph* gene, conferring resistance to kanamycin (**Supplementary Table S1**). Donors and recipients were cultured in LB broth for 18 h at 37°C with shaking. The OD_{600nm} of each culture was adjusted to 0.1 with fresh LB broth and 2.5 ml of each added to a 50ml tube and incubated for 18 h at 37°C with shaking in aerobic or anaerobic atmosphere, and in the presence or absence of 0.5 mg/l mitomycin C. The number of CFUs per ml of donors and recipients were quantified by culturing serial dilutions on LB agar supplemented with 0.03 mg/l chloramphenicol or 0.05 mg/l kanamycin, respectively. The presence of SGI-4 Δ*bar*::*cat* in recipient strains was quantified by serial dilution on LB agar supplemented with chloramphenicol and kanamycin. Transfer frequency was defined as the number of recipients containing SGI-4 Δ*bar*::*cat* as a proportion of donor cells in the culture. To determine whether transconjugant recipient strains contained SGI-4 in the same chromosomal location as the donor, the predicted right junction was amplified using primers that annealed on either side of the right junction of SGI-4 by PCR (**Supplementary Table S1**). To detect circularization of SGI-4 after excision, outward facing primers that annealed at the left and right junction of SGI-4 were used for PCR amplification (**Supplementary Table S1**). The sequence of the amplicons was determined by Sanger sequencing using the same primers (Eurofins sequencing service) and aligned to the genome of strain S04698-09.

Determination of Minimal Inhibitory Concentration (MIC) for Cu Sulfate

Fifteen bacterial strains (monophasic *S.* Typhimurium ST34 strains S04698-09, S04332-09, L00857-09, S05092-07, S06578-07,

[1]https://github.com/tseemann/snippy

L0938-09, and *S.* Typhimurium strains S01960-05, S05968-02, 11020-1996, NCTC13348, 4582-1995, S00914-05, SL1344, 9115-1996, and 6164-1997) were cultured for 18 h in LB broth at 37°C with shaking. A stock solution of 20 mM $CuSO_4$ in LB broth and 25 mM HEPES pH7 using NaOH. Serial dilutions in LB broth (pH7) were performed to generate a range of concentrations in 1mM intervals from 1 to 20 mM $CuSO_4$. 0.2 ml of LB broth buffered with 25mM HEPES pH7 and containing a range of concentrations of Cu sulfate were added to a polystyrene 96-well plate (Nunc) and incubated at 37°C for 24 h in normal atmospheric, microaerobic, or anaerobic conditions to equilibrate. After 24 h, each well was inoculated with 1×10^7 colony-forming units (CFUs) of the test bacterial strain cultured for 18 h in 10 ml LB Broth with shaking. Plates were incubated at 37°C for 24 h in normal atmospheric, microaerobic, or anaerobic conditions. The OD_{600nm} of each well was measured using a BIORAD Benchmark Plus microplate spectrophotometer. The MIC was defined as the mean concentration of Cu sulfate for which the OD_{600nm} of the culture was <0.2 from four biological replicates.

Quantitative Real Time PCR Expression Analysis

Expression relative to transcript abundance of a constitutively expressed housekeeping gene was determined as previously described (Branchu et al., 2014). Total RNA was extracted from 2 ml samples of overnight cultures of S04698-09 and S04698-09Δ*silCBA* strains grown in LB broth supplemented with 20, 100, 200 μM $CuSO_4$, or without $CuSO_4$ supplementation to mid-log phase (OD_{600nm} of 0.2) in anaerobic atmosphere (*pcoA*) or aerobic atmosphere (*copA*). Bacteria were harvested by centrifugation and RNA extracted using a FastRNA™ spin kit for microbes (MPBio) according to the manufacturer's instructions. RNA was treated with a TURBO DNA-free™ kit (ambion, life technologies™) according to the manufacturer's instructions before being reverse-transcribed using the QuantitTect® Reverse Transcription kit (Qiagen®). The resulting cDNA or serial dilutions of a known quantity of genomic DNA for generation of standard curves were amplified using a QuantiFast® SYBR® Green PCR kit (Qiagen®) with specific primers (**Supplementary Table S1**) for *copA* or *pcoA* test genes, and *rpoD* as a control housekeeping gene using the Applied Biosystems® 7500 real-time PCR system. The expression of *copA* or *pcoA* is presented relative to transcript abundance of the *rpoD* gene.

RAW264.7 Macrophage Survival Assay

Murine macrophages (RAW 264.7, ATCC, Rockville, MD) were grown in minimum essential medium (Sigma Aldrich) supplemented with 10% fetal bovine serum, L-glutamine (2 mM), and 1 × nonessential amino acids. For infection studies, 2×10^5 RAW 264.7 cells were seeded per well into 24-well plates and incubated at 37°C for 48 h. Cells were infected with S04698-09 wild-type strain, S04698-09Δ*sodCI*, S04698-09Δ*silCBApco* or S04698-09Δ*sodCI*Δ*silCBApco* from an overnight culture in LB broth at a multiplicity of infection of 20. Plates were centrifuged at 1500 rpm for 3 min followed by incubation of the cells for 30 min. The medium was then exchanged for a fresh one of the same compositions with the exception of the addition of gentamicin (100 μg/ml) to kill the extracellular bacteria. After either 2 or 24 h, the cells were washed with PBS and lysed with 1% Triton in PBS. The number of bacterial CFUs was determined by culturing lysate serial dilutions on LB agar. The data were expressed as the proportion of CFUs at 24 h relative to the CFU determined at 2 h.

Statistical Analysis

Where indicated in each figure with lines and asterisks, the Mann-Whitney *U* test was used to test the null hypothesis that randomly selected values in a sample were equally likely to be greater or smaller than from a second sample, using an alpha level of 0.05 to reject the null hypothesis.

RESULTS

SGI-4 Is a Member of a Novel Family of Integrative Conjugative Elements (ICE)

We investigated the presence of candidate genes within the SGI-4 coding for proteins capable of enabling mobilization of SGI-4 to recipient bacteria. The DNA excision, DNA processing, and conjugative transfer mechanisms we considered are common to MGEs such as ICE, also known as conjugative self-transmissible integrative (CONSTIN) elements (Hochhut and Waldor, 1999), integrative mobilizable elements (IMEs), and *cis*-mobilizable elements (CIMEs; Wozniak and Waldor, 2010). A number of ORFs that exhibited sequence similarity to DNA processing enzymes were identified on SGI-4 by sequence alignment to available databases (**Figure 1A** and **Supplementary Table S3**). Sequences of two of these exhibited similarity to site-specific tyrosine recombinases *xerC* and *xerD*, which are predicted to mediate integration and excision. However, we were unable to identify a recombination direction factor (RDF), which are known to exhibit little sequence conservation. ORF86 had sequence similarity to *traI* that encodes a relaxase protein. This relaxase protein is capable of binding to dsDNA at the origin of transfer (oriT) and inducing a single strand nick in conjunction with putative *uvrB* DNA helicase (ORF85) and *topB* topoisomerase (ORF 13), this facilitates unwinding of the DNA helix (Salyers et al., 1995). At the left end of SGI-4, *parA* and *parB* orthologs (ORFs 1 and 5) encode putative chromosome partitioning proteins. SGI-4 also contained a number of ORFs with similarity to genes encoding components of type IV secretion systems (T4SSs) involved in conjugative transfer of DNA: ORF26 of SGI-4 encoded a putative type IV coupling protein (T4CP), TraD, which initiates conjugal transfer; ORF51 encodes a putative inner membrane protein, TraG, for T4SS stabilization; ORF53 encodes a TraU ortholog involved in pilus assembly; ORF19 encodes a PilL F-type pilus protein; and ORF45 encoding an F-pilus assembly protein. These ORFS were present within several apparent operons (ORFs 14–27, 37–46, 48–53), consisting of multiple genes encoding proteins with no significant similarity to proteins in available databases but often associated with MGEs (**Figure 1A** and

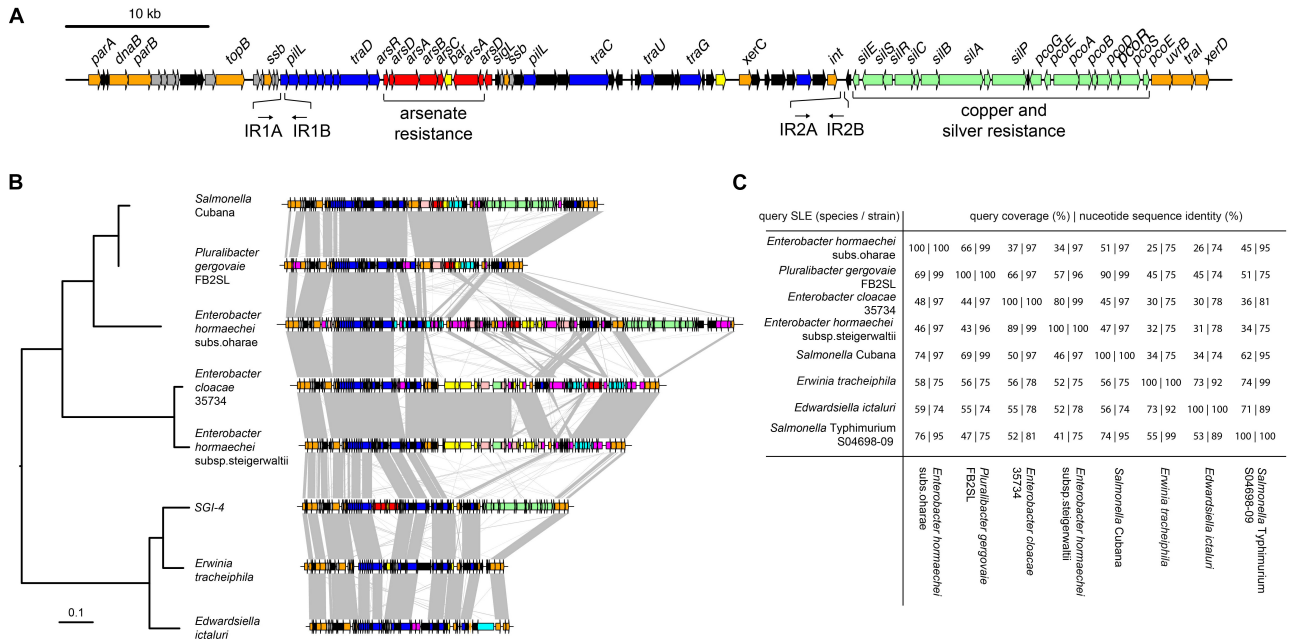


FIGURE 1 | Genetic map of *Salmonella* Genomic Island 4 (SGI-4) and phylogenetic relationship and gene synteny with SLEs. **(A)** Genetic diagram of *Salmonella* Genomic Island 4 (SGI-4); filled arrows indicate open reading frames with predicted functions based on sequence similarity for integration, excision and DNA processing (orange), type 4 secretion systems and conjugal transfer (blue), Cu and silver resistance (green), arsenic resistance (red), genes of unknown function commonly associated with ICE (gray), genes of unknown function not commonly associated with ICE (black). **(B)** Mid-point rooted maximum likelihood tree of eight SGI-4-like elements (SLEs) constructed using sequence variation in the core nucleotide sequence alignment of SLEs from diverse Enterobacteriaceae and genetic diagram of SLEs with regions of >75% nucleotide sequence identity (gray shading); filled arrows indicate open reading frames with the same designations as **(A)**, with additional functions, cadmium, cobalt, zinc or mercury resistance (pink), transposase and insertion elements (purple), and a repressor protein lexA (brown). **(C)** Pair-wise sequence comparison of SLEs indicating the percentage of sequence exhibiting >60% identity and the mean nucleotide sequence identity of the shared sequence.

Supplementary Table S3). We were unable to identify a putative *oriT* sequence in SGI-4 using oriTfinder (Li et al., 2018). However, two pairs of inverted repeats, each with 80% identity to one another, were present in ORF19 and in an intergenic region between ORF64 and ORF65.

To determine whether SGI-4 was similar to known MGEs we aligned the sequence of SGI-4 with non-redundant nucleotide sequences in the NCBI database using discontinuous megaBLAST. Seven assembled contiguous sequences in the NCBI database contained a core set of ICE genes with a modular arrangement with complete synteny to SGI-4, none of which had been described previously as MGEs (**Figure 1B**). These exhibited at least 75% nucleotide sequence identity with over 40% of SGI-4, and we therefore refer to these as SGI-4-like elements (SLEs). SLEs were inserted in the genome adjacent to a phenylalanine phe-tRNA in strains from a diverse range of Enterobacteriaceae, indicating this was the common attachment site (attB/attP). SGI-4 was most closely related to SLEs from *Erwinia tracheiphila* and *Edwardsiella ictaluri*, and more distantly related to SLEs from two *Enterobacter hormachei* strains from subspecies *oharae* and *steigerwaltii*, *Enterobacter cloacae, Pluralibacter gergoviae* and an SLE present in a strain of *Salmonella enterica* serovar Cubana (**Figures 1B,C**). All shared a number of regions of at least 75% sequence identity and synteny that encoded putative DNA processing enzymes or components of a T4SS, and were interspersed amongst apparent cargo genes involved in diverse functions capable of modifying the phenotype of the host bacterium.

Since SGI-4 encoded many of the genes normally associated with ICE, we determined whether it was capable of mobilization to a recipient *S.* Typhimurium (strain SL1344) during co-culture. To enable identification of recipients containing SGI-4 we inserted a chloramphenicol resistance gene (*cat*) in the arsenic resistance locus of SGI-4. Transfer frequency was low under aerobic culture conditions ($<1 \times 10^{-8}$ CFU per donor CFU) but was substantially increased in the presence of mitomycin C, as described previously for the SXT ICE (Beaber et al., 2004), or by culture in anaerobic conditions (**Figure 2**). Mitomycin C and anaerobiosis had an additive effect on transfer since the transfer rate in anaerobic conditions with mitomycin C was significantly higher ($p < 0.005$) than any other condition tested ($\sim 1 \times 10^{-4}$ CFU per donor CFU). PCR amplification and sequence analysis of the junction site of the donor strain indicated that, on transfer, SGI-4 inserts into the same position on the chromosome (phe-tRNA locus) as the donor strain (data not shown). Furthermore, in the monophasic *S.* Typhimurium strain S04698-09 cultured in the presence of mitomycin C, PCR amplification and sequence analysis of an amplicon generated using specific outward facing primers at each end

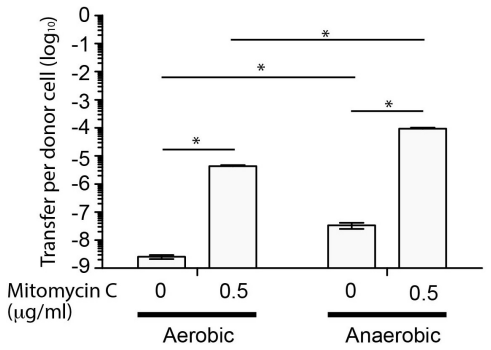


FIGURE 2 | Transfer frequency of SGI-4::*cat* to *S*. Typhimurium during co-culture is enhanced by mitomycin C and anaerobic conditions. Bars indicate the mean number of chloramphenicol resistant CFU of *S*. Typhimurium strain SL1344 per CFU of monophasic *S*. Typhimurium ST34 strain S04698-09 SGI-4 *silCBA*::*cat*, following co-culture. The mean from six biological replicates ± standard deviation are shown. *Indictaes that groups indicated by lines were significantly different ($p < 0.05$).

of SGI-4, was consistent with circularization of excised SGI-4 (data not shown).

SGI-4 Is Characteristic of the Monophasic *S*. Typhimurium ST34 Clade

The SGI-4 sequence was present in 23 of 24 monophasic *S*. Typhimurium ST34 clade strains, but absent from all other strains, in a collection of representative *S*. Typhimurium (**Figure 3**). SGI-4 sequence was also present almost exclusively in isolates of the monophasic *S*. Typhimurium ST34 clade in the whole genome sequence of 1697 *S*. Typhimurium and monophasic variant *S*. Typhimurium isolates from human clinical cases in England and Wales during 2014 and 2015 (**Supplementary Figure S1**). Within the monophasic *S*. Typhimurium ST34 clade just 38 isolates (4%) lacked the SGI-4 sequence, and these were distributed sporadically throughout the clade in 28 small clusters or individual leaves of the phylogenetic tree (**Supplementary Figure S1**). In contrast, of 797 *S*. Typhimurium isolates that were present outside the monophasic *S*. Typhimurium ST34 clade, just five contained the SGI-4 sequence, and these all shared a recent common ancestor with the monophasic *S*. Typhimurium ST34 clade.

SGI-4-Encoded *sil* Is Phylogenetically Distinct From the Chromosome-Encoded *cus* Genes of *Escherichia coli*

The nomenclature for copper RND-family efflux pump genes is confusing with *sil* and *cus* being used without reference to ancestry. A cluster of 18 ORFS on SGI-4 included 15 genes that exhibited sequence similarity to genes predicted to encode an RND-family efflux pump previously designated as either *cusRS cusCFBA* (Outten et al., 2001) or *silRSE silCBAP* (Gupta et al., 1999), and *pcoABDRSE pcoEG* (*pco* locus) involved in Cu and silver resistance. Alignment of the RND-family efflux pump genes of SGI-4 with sequences from previously characterized homologs indicated that the *cusRSCFBA* genes on the *E. coli* chromosome form a distinct outgroup from a closely related cluster that included the SGI-4 genes, that evolved from a common ancestor closely related to the *silRSE silCBAP* on pMG101 (**Supplementary Figure S2**). For this reason, we designated SGI-4 encoded RND-family efflux pump genes as *silRSE silCBAP* (*sil* locus).

The *sil* Locus Is Rare in the Genus *Salmonella*

Investigation of the distribution of the SGI-4 *silA* genes in two sequence databases indicated that these genes are rare in the genus *Salmonella*, consistent with deletion early in the evolution of the genus, and infrequent reacquisition. Alignment using BLASTn of the *silA* gene on SGI-4 with 926 *Salmonella* genomes, one from each eBurst group (largely corresponding serovar) representing all known genotypic diversity of *Salmonella* (Alikhan et al., 2018), indicated that just 16 genomes (2%) contained an ortholog. In a second search of the ENA, the *silA* gene was in 3439 sequence entries, while the *invA* gene in 51026 data entries, indicating that *silA* was present in approximately 7% of *S*. Typhimurium genomes in the database.

Enhanced Resistance of Monophasic *S*. Typhimurium ST34 to Cu *in vitro* Is Mediated by the *silCBA* Genes

In order to investigate the impact of the acquisition of SGI-4 on metal ion resistance, we compared the MICs of Cu sulfate for five strains of monophasic *S*. Typhimurium ST34 with three strains each of DT204, U288 and DT104 that lacked SGI-4, and a single monophasic *S*. Typhimurium ST34 strain that also lacked SGI-4 due to deletion (**Figure 4**). During culture in aerobic or microaerobic conditions the presence of SGI-4 had a small but significant impact on the MIC for Cu. However, under anaerobic conditions, the MICs for strains lacking SGI-4 decreased by around five-fold; in contrast, SGI-4-containing monophasic *S*. Typhimurium ST34 strains had similar MICs under both aerobic and microaerobic conditions.

SGI-4 has two clusters of genes predicted to encode an RND family efflux pump (*silCBA*), and a second that includes a multicopper oxidase system (*pcoABDE*), both previously implicated in resistance to Cu. To determine the relative role of these two loci in Cu resistance we determined the MICs of mutants of monophasic *S*. Typhimurium ST34 strain S04698-09 that had deletions of either *silCBA*, *pcoABC* or both these loci (**Figure 5**). When functional *silCBA* genes were present, deletion of *pcoABDE* genes alone had no effect on resistance to Cu under any of the conditions evaluated. Deletion of *silCBA* genes resulted in a small but significant ($p < 0.05$) decrease in the MIC for Cu under aerobic and microaerobic conditions, and a more substantial decrease under anaerobic conditions. However, even under anaerobic conditions, the effect of the *silCBA* deletion on the MIC for Cu was not comparable to the absence of SGI-4 in U288, DT104 and DT204 isolates which had a MIC that was

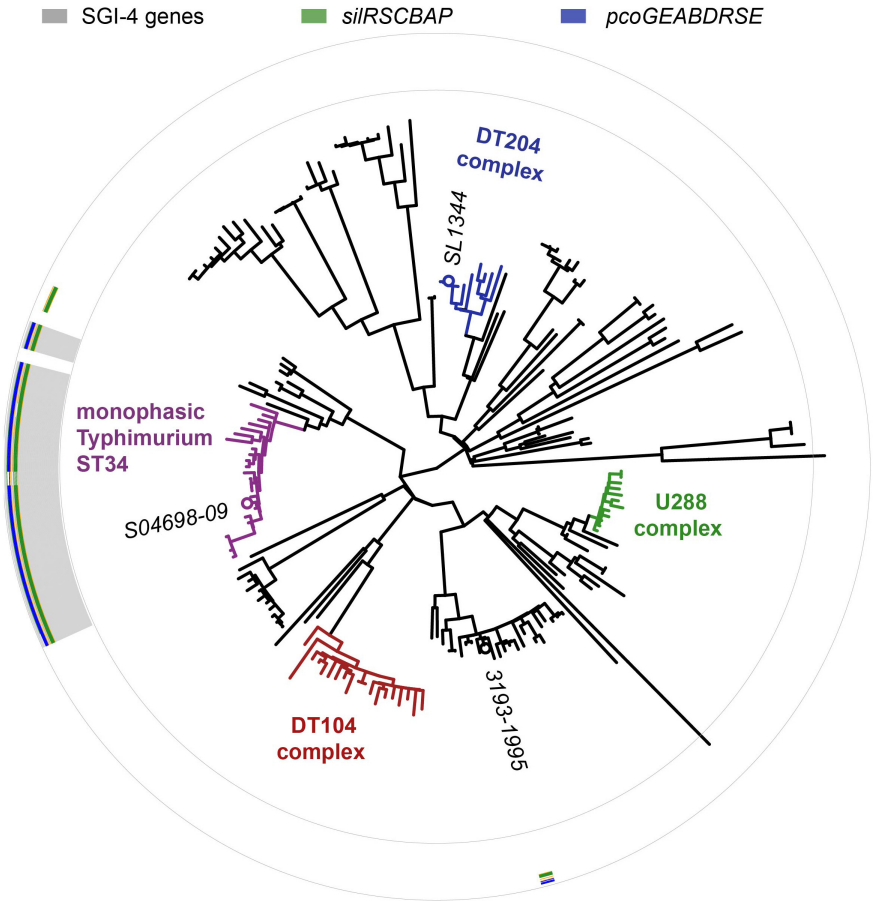


FIGURE 3 | Distribution of SGI-4 in monophasic *S.* Typhimurium ST34 and representative strains of *S.* Typhimurium. A maximum likelihood tree constructed using sequence variation in the core genome of 24 monophasic *S.* Typhimurium ST34 strains and 153 representative strains of *S.* Typhimurium with reference to the whole genome sequence of strain SL1344 (accession FQ312003). Selected strains are identified (italicized text). The presence of sequences that mapped to 87 SGI-4 ORFs from S04698-09 are represented as filled boxes in 87 concentric circles: *sil* genes (green), *pco* genes (blue), or other SGI-4 ORFs (gray).

below 5 mM Cu sulfate. Cu resistance at a similar level to these strains was only observed in monophasic *S.* Typhimurium ST34 strain S04698-09 when both *silCBA* and *pcoABDE* were deleted.

The Presence of *silCBA* Genes on SGI-4 Alters Expression of *pcoA* and the Native Cu Homeostasis Gene *copA*

To investigate any apparent redundancy of *pcoABDE* in the presence of *silCBA*, we determined the expression of *pcoA* in the presence or absence of *silCBA*. The expression of *pcoA* was not affected by increasing concentrations of Cu, unless the *silCBA* genes were deleted from monophasic *S.* Typhimurium ST34 strain S04698-09 (**Figure 6A**). In the absence of *silCBA* induction was around ten-fold that achieved in the absence of Cu in the growth medium.

To determine the impact of SGI-4 acquisition on typical Cu homeostasis in *S.* Typhimurium, we investigated the expression of the *copA* gene that encodes the native ATPase protein involved in transport of Cu from the cytoplasm into the periplasm in *Salmonella*. The expression of *copA* increased in a dose-dependent manner in the presence of Cu in *S.* Typhimurium strain SL1344, which lacked SGI-4, and in the monophasic *S.* Typhimurium ST34 strain S04698-09 that encoded the island. However, expression of *copA* was significantly higher in *S.* Typhimurium strain SL1344 than in monophasic *S.* Typhimurium ST34 strain S04698-09 during culture in media with 100 μM Cu. Deletion of *cusACFBA* genes in strain S04698-09 resulted in increased expression of *copA* approaching the expression level observed in SL1344 in 100 μM of $CuSO_4$ (**Figure 6B**).

Presence of the *silCBA* Genes Does Not Affect SodCI-Mediated Resistance to Macrophage Killing

We then tested the hypothesis that the presence of *silCBA* affected the function of SodCI, a periplasmic superoxide dismutase that uses Cu as co-factor that is supplied by the CopA ATPase Cu transporter and the periplasmic Cu binding protein CueP (Osman et al., 2013). SodCI dismutates oxygen free radicals produced by phagosome associated NADPH-dependent oxidase

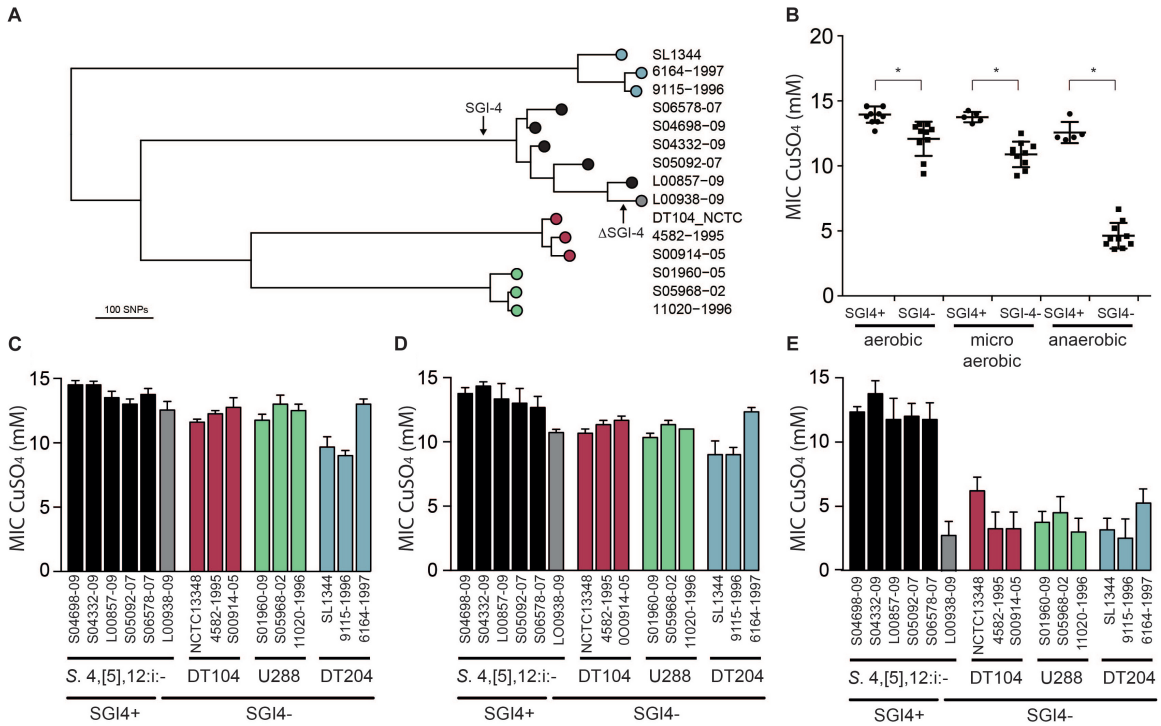


FIGURE 4 | Phylogenetic relationship and Cu sulfate MIC of representative strains of monophasic *S*. Typhimurium ST34 and *S*. Typhimurium. **(A)** A mid-point rooted phylogenetic tree constructed using sequence variation in the core genome with reference to *S*. Typhimurium strain SL1344 whole genome sequence (accession FQ312003). Leaves of the tree corresponding to representative strains are labeled as filled circles color coded for SGI-4$^+$ monophasic *S*. Typhimurium ST34 (black) and SGI-4$^-$ monophasic *S*. Typhimurium ST34 (gray), and *S*. Typhimurium DT104 complex (red), U288 complex (green), and DT204 complex (blue) are shown. **(B)** Mean MIC for Cu sulfate of SGI-4$^+$ (filled circles) and SGI-4$^-$ (filled squares) monophasic *S*. Typhimurium ST34 and *S*. Typhimurium strains during aerobic, microaerobic and anaerobic culture. Bars indicate the mean MIC for Cu sulfate for each strain ± standard deviation in aerobic **(C)**, microaerobic **(D)**, and anaerobic **(E)** atmosphere. Bar colors match the tree leaf circles in **(A)**. *Indicates that groups indicated by lines were significantly different ($p < 0.05$).

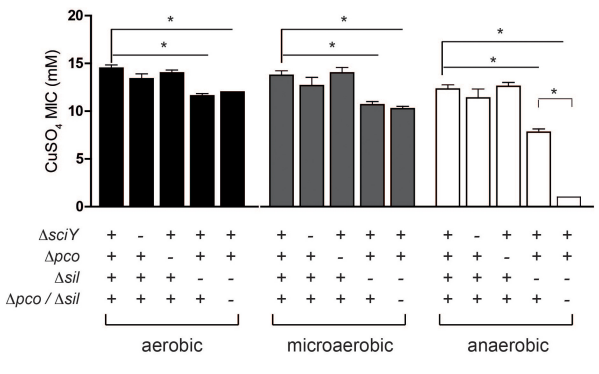


FIGURE 5 | MIC for Cu sulfate of monophasic *S*. Typhimurium ST34 strain S04698-09 containing defined deletions in *sil* and *pco* genes. Bars indicate the mean MIC ± standard deviation of four biological replicates during culture in either aerobic, microaerobic or anaerobic conditions. The genotype with respect to *sciY* (negative control), *pco* and *sil* are indicated as present (+) or deleted and replaced with a *cat* gene (–). * indicates the MIC was significantly different from wild type strain S04698-09 ($p < 0.05$).

(Phox) of macrophages, contributing to intracellular survival of *Salmonella*. We expected that the expression of *silCBA* that encoded a copper RND efflux pump on SGI-4 would result in transport of Cu from the cytoplasm and periplasm to the external milieu (Gupta, 1994), depleting the pool of Cu available in the periplasm. Consistent with this, expression of *pcoA*, a periplasmic multicopper oxidase gene whose expression is controlled by PcoR that senses copper in the periplasm, was not induced during culture in 100 μM Cu, unless the *silCBA* was deleted from SGI-4.

To test the hypothesis we infected gamma interferon-activated RAW macrophages with mutants of monophasic *S*. Typhimurium ST34 strain S04698-09 that encoded either *silCBA* and *sodCI*, or lacked one or both of these loci and determined the change in intracellular viable counts of *S*. Typhimurium in macrophages between 2 and 24 h post-inoculation (**Figure 7**). Strain S04698-09 exhibited a net replication of nearly ten-fold after 24 h in RAW macrophages, and deletion of the *sodCI* gene resulted in a small but significant ($p < 0.05$) reduction. This was consistent with previous reports that SodCI is required by the bacterium to evade the killing mechanisms of macrophages. However, deletion of the *silCBA* genes did not result in a significant decrease in net replication in RAW macrophages, but additional deletion of *sodCI* resulted in a similar decrease in net replication to that observed in the presence of the *silCBA* genes. These data are consistent with a fully functional SodCI gene in the presence of altered Cu homeostasis due to the presence of the *silCBA* genes.

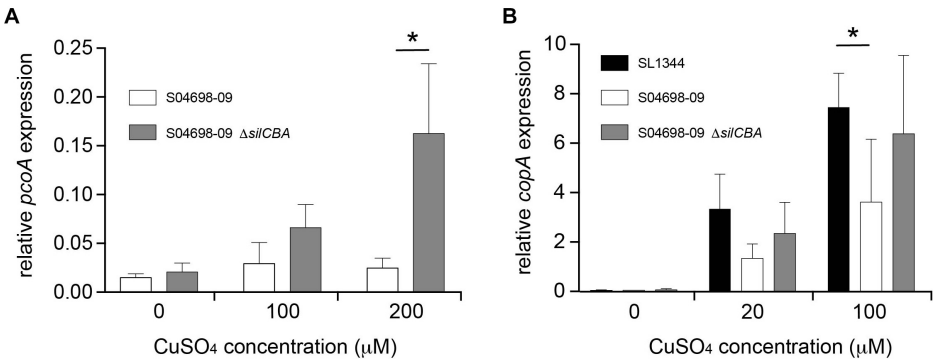


FIGURE 6 | Presence of the *silCBA* genes decreases relative transcript abundance of *pcoA* and *copA*. The abundance of *pcoA* **(A)** or *copA* **(B)** transcript relative to the housekeeping gene *rpoD* quantified by quantitative RT-PCR from total RNA prepared from mid-log phase cultures in LB broth supplemented with Cu sulfate. Bars indicate mean relative transcript from five biological replicates ± standard deviation. *Indicates that groups indicated by lines were significantly different ($p < 0.05$).

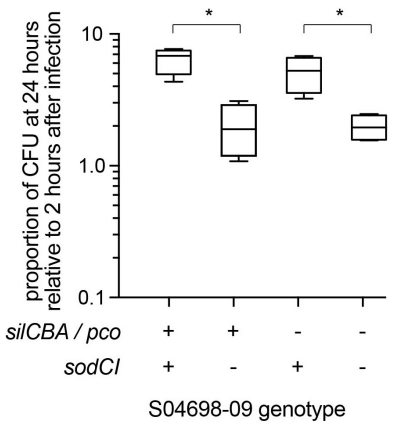


FIGURE 7 | Net replication of monophasic *S*. Typhimurium ST34 in activated RAW macrophages is *sodCI* dependent but *sil* / *pco* independent. Box and whiskers plot indicating the mean number of bacterial CFUs 24 h post-infection, relative to 2 h post infection of wild type monophasic *S*. Typhimurium strain S04698-09 and otherwise isogenic strains containing deletions of the *sodCI* gene, the *sil*/*pco* genes together, or all three loci deleted. *Indicates that groups indicated by lines were significantly different ($p < 0.05$).

DISCUSSION

For vertical transfer ICE undergo replication within the chromosome; for horizontal transfer they undergo excision from the genome (Johnson and Grossman, 2015). Horizontal transfer of an ICE involves common processes including excision from the host chromosome, circularization, conjugative transfer by a type IV secretion system (T4SS), and integration within a recipient genome at an attachment site (*attB*; Guglielmini et al., 2011). A number of ORFs exhibiting sequence similarity to genes encoding DNA processing enzymes that are predicted to be involved in excision, integration and conjugative transfer of DNA via a type IV secretion system (T4SSs) were also present on SGI-4. Together these data were consistent with the idea that SGI-4 is an ICE.

In an analysis of 1,124 taxonomically diverse complete prokaryotic genomes, 335 putative ICEs were detected based on the presence of T4SS and a relaxase gene in close proximity, indicating the widespread nature of these MGEs (Guglielmini et al., 2011). Despite the presence of multiple genes with sequence similarity to those with functions required for the transfer of ICEs, SGI-4 did not exhibit extensive similarity to previously characterized MGEs, and is currently not present in the ICEberg database (Liu et al., 2018). Seven complete genomes from diverse species of Enterobacteriaceae that were added to available databases since the analysis by Guglielmini et al. (2011), had between 75 and 99% nucleotide sequence identity with SGI-4 and between 41 and 76% shared core sets of genes with SGI-4. As is common for ICEs, core genes were interspersed with cargo genes with functions that were unrelated to mobilization but had the potential to modify the phenotype of the host bacterium (Wozniak and Waldor, 2010). In particular, genes involved in resistance to heavy metals were present on SLEs, with five containing at least one locus associated with resistance to Cu, silver, mercury or arsenic. Two SLEs from *Erwinia tracheiphila* and *Edwardsiella ictaluri* were particularly closely related to SGI-4, but contained unrelated cargo genes, highlighting the rapid divergence achieved by horizontal gene transfer in this family of ICE. Our analysis is consistent with SGI-4 evolving from a common ancestor of these ICEs by acquisition of *ars*, *sil* and *pco* cargo genes involved in resistance to metal ions.

With the exception of a small outbreak in a burns unit associated with a *S*. Typhimurium strain in the 1970s, historically, the *sil* and *pco* loci encoding resistance to silver and Cu, have rarely been associated with *Salmonella enterica* isolates. Indeed, we found that *sil* genes were present in just 2% of whole genome sequences representing the genotypic diversity of *S. enterica* and *S. bongori*, and 7% of more than 50,000 *Salmonella* whole genome sequences in the ENA. However, *sil* and *pco* genes, have been commonly associated with three lineages of monophasic *S*. Typhimurium in the past 20 years: the "European clone" (monophasic *S*. Typhimurium ST34); the "Spanish clone"; and the "Southern European clone" (Mourao et al., 2015; Mastrorilli et al., 2018). However, only 74 and

26% of the latter two, respectively, encoded the *sil* and/or the *pco* genes and these were plasmid-borne (Mourao et al., 2015), suggesting that they may be lost relatively frequently. In contrast, monophasic *S.* Typhimurium ST34 encodes *sil* and *pco* genes on the chromosome, and although on a MGE, we found these genes in 96% of the 797 clinical isolates from England and Wales that we evaluated. Furthermore, loss of SGI-4 and the *silCBA* and *pco* genes was sporadic, and mostly as singleton taxa on the phylogenetic tree, consistent with the absence of selection for their loss (Petrovska et al., 2016).

Given the general paucity of genes encoding Cu/silver RND efflux pumps in *Salmonella enterica* and the evolutionary history of the genus, the apparent lack of a selective advantage for the loss of *silCBA* genes may indicate an important role for these genes in the monophasic *S.* Typhimurium ST34 clone. Deletion of *cusCFBA* in the *Salmonella* ancestor is likely to have had a profound effect on Cu distribution in the cell because *Salmonella* has no alternative mechanism to remove Cu from cell. While *E. coli* can remove Cu from the cell entirely, *Salmonella* transports Cu from the cytoplasm, where it is especially toxic, to the periplasm, where it is detoxified by the Cu oxidase CueO or bound to the Cu binding protein, CueP, the major reservoir of Cu in *Salmonella* (Osman et al., 2010). CueP is thought to play an important role in supplying Cu to superoxide dismutase SodCI, important for survival in the face of the oxidative burst generated by macrophage. The evolution of *Salmonella* pathogenesis must, therefore, have proceeded in the context of fundamentally altered Cu homeostasis. The reintroduction of the *silCBA* locus on SGI-4 did indeed appear to alter the Cu distribution and levels in *Salmonella* as indicated by the lack of *pcoA* expression in the presence of the RND-family of efflux genes, which are regulated in response to Cu levels in the periplasm. The consequences of altered Cu levels in monophasic *S.* Typhimurium ST34 is not clear but, at least in RAW macrophages *in vitro*, the dependency of SodCI Cu as a cofactor did not affect resistance to macrophage killing. This is consistent with the report that limitation of Cu in the periplasm by deletion of the *copA* and *golT* genes, and the supply of Cu to SodCI by deletion of the *cueP* gene in S. Typhimurium had no effect on systemic infection in the murine model of infection (Fenlon and Slauch, 2017).

Strong selection pressure may be important for retention of SGI-4 since our data is consistent with its presence altering Cu distribution and levels. Monophasic *S.* Typhimurium ST34 is primarily associated with pigs, and it has been suggested that the success of this clone may, in part, have been driven by the extensive use of Cu as a growth promotor in pig rearing (Mourao et al., 2015; Petrovska et al., 2016). Consistent with this idea, the presence of SGI-4 is correlated with enhanced resistance to Cu; this appears to be particularly apparent under anaerobic conditions which are similar to those encountered by *Salmonella* in the host intestinal tract. *Salmonella* strains that lacked SGI-4, exhibited MICs for Cu of around 2–3 mM under anaerobic conditions, which is approximately 15% of the values achieved under aerobic conditions. Importantly, this is in the range of the concentrations of Cu found in pig manure effluent and sludge on farms where Cu-supplemention of diet was common (Nicholson et al., 1999; Bolan et al., 2003; Cang et al., 2004). The presence of SGI-4, however, increased the MIC for Cu (under anaerobic conditions) by approximately 500%, elevating it to levels above the concentrations likely to be encountered on pig farms (Nicholson et al., 1999; Bolan et al., 2003; Cang et al., 2004). The *sil* genes encoded on SGI-4 were entirely responsible for the observed increase in MICs for Cu, since deletion of *pco* genes alone did not result in a decrease in the MICs. The *pco* locus encodes a multiCu oxidase which is thought to detoxify the more damaging Cu^{+} to Cu^{2+} ions by oxidation (Lee et al., 2002). Monophasic *S.* Typhimurium ST34 encodes a native Cu oxidase, CueO, and the presence of this protein may, in part, mask the activity of the *pco* locus. However, the *pco* genes were capable of enhancing resistance to Cu in the absence of *silCFBA*. Furthermore, since there was no evidence for loss of *pco* genes during clonal expansion of the monophasic *S.* Typhimurium ST34 clade, it remains possible that the Pco system may play a more prominent role in Cu resistance under environmental conditions that we did not evaluate in our *in vitro* experiments.

AUTHOR CONTRIBUTIONS

RK, PB, and OC conceived the study. PB and OC performed the experiments and generated the data. TD and LP provided materials and expert advice. RK, PB, and OC drafted the manuscript. All authors analyzed the data, provided critical input into the final manuscript, and approved the final version.

FUNDING

RK was supported by research grants BB/N007964/1 and BB/M025489/1, and by the BBSRC Institute Strategic Programme Microbes in the Food Chain BB/R012504/1 and its constituent project(s) BBS/E/F/000PR10348 and BBS/E/F/000PR10349. OC was supported by a BBSRC DTP studentship (BB/M011216/1).

ACKNOWLEDGMENTS

This manuscript has been released as a Pre-Print at *bioRxiv* (Branchu et al., 2019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2019.01118/full#supplementary-material

FIGURE S1 | Distribution of SGI-4 in monophasic *S.* Typhimurium ST34 and representative strains of *S.* Typhimurium. A maximum likelihood tree constructed using sequence variation in the core genome of 1814 monophasic *S.* Typhimurium ST34 strains or *S.* Typhimurium isolates with reference to the whole genome sequence of strain SL1344 (accession FQ312003). The presence of sequence in each genome that mapped to 787 SGI-4 ORFs from S04698-09 are represented as filled boxes in 87 concentric circles: *sil* genes (green), *pco* genes (blue), or other SGI-4 ORFs (gray).

FIGURE S2 | A maximum likelihood tree showing the relationship of *cus* and *sil* loci.

TABLE S1 | Oligonucleotide primers used for the construction of mutant strains of *Salmonella*, determination of genotype, or qRT-PCR.

TABLE S2 | Isolate name, accession number and the figure(s) in which the data were used to construct maximum likelihood trees.

TABLE S3 | Gene model for SGI-4 including gene designation, annotation, SLE core genome and nucleotide sequence of predicted coding sequences.

REFERENCES

Alikhan, N. F., Zhou, Z., Sergeant, M. J., and Achtman, M. (2018). A genomic overview of the population structure of *Salmonella*. *PLoS Genet.* 14:e1007261. doi: 10.1371/journal.pgen.1007261

Andres-Barranco, S., Vico, J. P., Marin, C. M., Herrera-Leon, S., and Mainar-Jaime, R. C. (2016). Characterization of *Salmonella enterica* serovar typhimurium isolates from pigs and pig environment-related sources and evidence of new circulating monophasic strains in Spain. *J. Food Prot.* 79, 407–412. doi: 10.4315/0362-028X.JFP-15-430

Animal and Plant Health Agency [APHA] (2017). *Salmonella in livestock production in Great Britain, 2016*. Available at: https://www.gov.uk/government/publications/salmonella-in-livestock-production-in-great-britain-2016 (accessed October 31, 2017).

Antunes, P., Mourao, J., Pestana, N., and Peixe, L. (2011). Leakage of emerging clinically relevant multidrug-resistant *Salmonella* clones from pig farms. *J. Antimicrob. Chemother.* 66, 2028–2032. doi: 10.1093/jac/dkr228

Arguello, H., Sorensen, G., Carvajal, A., Baggesen, D. L., Rubio, P., and Pedersen, K. (2013). Prevalence, serotypes and resistance patterns of *Salmonella* in Danish pig production. *Res. Vet. Sci.* 95, 334–342. doi: 10.1016/j.rvsc.2013.04.001

Beaber, J. W., Hochhut, B., and Waldor, M. K. (2004). SOS response promotes horizontal dissemination of antibiotic resistance genes. *Nature* 427:72.

Bolan, N. S., Khan, M. A., Donaldson, J., Adriano, D. C., and Matthew, C. (2003). Distribution and bioavailability of copper in farm effluent. *Sci. Total Environ.* 309, 225–236. doi: 10.1016/S0048-9697(03)00052-4

Bonardi, S. (2017). *Salmonella* in the pork production chain and its impact on human health in the European Union. *Epidemiol. Infect.* 145, 1513–1526. doi: 10.1017/S095026881700036X

Bradley, P., den Bakker, H. C., Rocha, E. P. C., McVean, G., and Iqbal, Z. (2019). Ultrafast search of all deposited bacterial and viral genomic data. *Nat. Biotechnol.* 37, 152–159. doi: 10.1038/s41587-018-0010-1

Branchu, P., Bawn, M., and Kingsley, R. A. (2018). Genome variation and molecular epidemiology of *Salmonella* Typhimurium pathovariants. *Infect. Immun.* 86:e0079-18. doi: 10.1128/IAI.00079-18

Branchu, P., Charity, O. J., Bawn, M., Thilliez, G., Dallman, T. J., Petrovska, L., et al. (2019). SGI-4 in monophasic *Salmonella* Typhimurium ST34 is a novel ICE that enhances resistance to copper. *bioRxiv* doi: 10.1101/518175

Branchu, P., Matrat, S., Vareille, M., Garrivier, A., Durand, A., Crepin, S., et al. (2014). NsrR, GadE, and GadX interplay in repressing expression of the *Escherichia coli* O157:H7 LEE pathogenicity island in response to nitric oxide. *PLoS Pathog* 10:e1003874. doi: 10.1371/journal.ppat.1003874

Brown, N. L., Barrett, S. R., Camakaris, J., Lee, B. T., and Rouch, D. A. (1995). Molecular genetics and transport analysis of the copper-resistance determinant (pco) from *Escherichia coli* plasmid pRJ1004. *Mol. Microbiol.* 17, 1153–1166.

Campos, J., Mourao, J., Marcal, S., Machado, J., Novais, C., Peixe, L., et al. (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, 843–845. doi: 10.1093/jac/dkv409

Cang, L., Wang, Y., Zhou, D., and Dong, Y. (2004). Heavy metals pollution in poultry and livestock feeds and manures under intensive farming in Jiangsu Province, China. *J. Environ. Sci.* 16, 371–374.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

Datsenko, K. A., and 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, 6640–6645.

Espariz, M., Checa, S. K., Audero, M. E., Pontel, L. B., and Soncini, F. C. (2007). Dissecting the *Salmonella* response to copper. *Microbiology* 153(Pt 9), 2989–2997. doi: 10.1099/mic.0.2007/006536-0

Fenlon, L. A., and Slauch, J. M. (2017). Cytoplasmic copper detoxification in *Salmonella* can contribute to sodc metalation but is dispensable during systemic infection. *J. Bacteriol.* 199:e00437-17. doi: 10.1128/JB.00437-17

Fookes, M., Schroeder, G. N., Langridge, G. C., Blondel, C. J., Mammina, C., Connor, T. R., et al. (2011). *Salmonella bongori* provides insights into the evolution of the *Salmonella*e. *PLoS Pathog.* 7:e1002191. doi: 10.1371/journal.ppat.1002191

Franke, S., Grass, G., Rensing, C., and Nies, D. H. (2003). Molecular analysis of the copper-transporting efflux system CusCFBA of *Escherichia coli*. *J. Bacteriol.* 185, 3804–3812.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*

Grass, G., and Rensing, C. (2001). CueO is a multi-copper oxidase that confers copper tolerance in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 286, 902–908. doi: 10.1006/bbrc.2001.5474

Guglielmini, J., Quintais, L., Garcillán-Barcia, M. P., De La Cruz, F., and Rocha, E. P. (2011). The repertoire of ICE in prokaryotes underscores the unity, diversity, and ubiquity of conjugation. *PLoS genet.* 7:e1002222. doi: 10.1371/journal.pgen.1002222

Gupta, A. (1994). Multidrug-resistant typhoid fever in children: epidemiology and therapeutic approach. *Pediatr. Infect. Dis. J.* 13, 134–140.

Gupta, A., Matsui, K., Lo, J.-F., and Silver, S. (1999). Molecular basis for resistance to silver cations in *Salmonella*. *Nat. Med.* 5:183.

Hauser, E., Tietze, E., Helmuth, R., Junker, E., Blank, K., Prager, R., et al. (2010). Pork contaminated with *Salmonella enterica* serovar 4,[5],12:i:-, an emerging health risk for humans. *Appl. Environ. Microbiol.* 76, 4601–4610. doi: 10.1128/AEM.02991-09

Hochhut, B., and Waldor, M. K. (1999). Site-specific integration of the conjugal *Vibrio cholerae* SXT element into prfC. *Mol. Microbiol.* 32, 99–110.

Hopkins, K. L., de Pinna, E., and Wain, J. (2012). Prevalence of *Salmonella enterica* serovar 4,[5],12:i:- in England and Wales, 2010. *Euro surveill.* 17, 10–16.

Hugas, M., and Beloeil, P. (2014).). Controlling *Salmonella* along the food chain in the European Union - progress over the last ten years. *Euro Surveill.* 19:20804.

Johnson, C. M., and Grossman, A. D. (2015). Integrative and conjugative elements (ICEs): what they do and how they work. *Ann. Rev. Genet.* 49, 577–601. doi: 10.1146/annurev-genet-112414-055018

Kroger, C., Dillon, S. C., Cameron, A. D., Papenfort, K., Sivasankaran, S. K., Hokamp, K., et al. (2012). The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proc. Natl. Acad. Sci. U.S.A.* 109, E1277–E1286. doi: 10.1073/pnas.1201061109

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Lee, S. M., Grass, G., Rensing, C., Barrett, S. R., Yates, C. J., Stoyanov, J. V., et al. (2002). The Pco proteins are involved in periplasmic copper handling in *Escherichia coli*. *Biochem. Biophys. Res. Commun.* 295, 616–620.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv* doi: 1303.3997v2

Li, X., Xie, Y., Liu, M., Tai, C., Sun, J., Deng, Z., et al. (2018). oriTfinder: a web-based tool for the identification of origin of transfers in DNA sequences of bacterial mobile genetic elements. *Nucleic Acids Res.* 46, W229–W234. doi: 10.1093/nar/gky352

Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., et al. (2018). ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res.* 47, D660–D665. doi: 10.1093/nar/gky1123

Macomber, L., and Imlay, J. A. (2009). The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 8344–8349. doi: 10.1073/pnas.0812808106

Mastrorilli, E., Pietrucci, D., Barco, L., Ammendola, S., Petrin, S., Longo, A., et al. (2018). A comparative genomic analysis provides novel insights into

the ecological success of the monophasic *Salmonella* serovar 4,[5],12:i. *Front. Microbiol.* 9:715. doi: 10.3389/fmicb.2018.00715

Mather, A. E., Lawson, B., de Pinna, E., Wigley, P., Parkhill, J., Thomson, N. R., et al. (2016). Genomic Analysis of *Salmonella enterica* serovar typhimurium from wild passerines in england and wales. *Appl. Environ. Microbiol.* 82, 6728–6735. doi: 10.1128/AEM.01660-16

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., et al. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature* 413:852. doi: 10.1038/35101614

McHugh, G. L., Moellering, R. C., Hopkins, C. C., and Swartz, M. N. (1975). *Salmonella* typhimurium resistant to silver nitrate, chloramphenicol, and ampicillin. *Lancet* 1, 235–240.

Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., and Schaffer, A. A. (2008). Database indexing for production MegaBLAST searches. *Bioinformatics* 24, 1757–1764. doi: 10.1093/bioinformatics/btn322

Mourao, J., Novais, C., Machado, J., Peixe, L., and Antunes, P. (2015). Metal tolerance in emerging clinically relevant multidrug-resistant *Salmonella enterica* serotype 4,[5],12:i:- clones circulating in Europe. *Int. J. Antimicrob. Agents* 45, 610–616. doi: 10.1016/j.ijantimicag.2015.01.013

Nicholson, F. A., Chambers, B. J., Williams, J. R., and Unwin, R. J. (1999). Heavy metal contents of livestock feeds and animal manures in England and Wales. *Biores. Technol.* 70, 23–31. doi: 10.1016/S0960-8524(99)00017-6

Osman, D., Patterson, C. J., Bailey, K., Fisher, K., Robinson, N. J., Rigby, S. E., et al. (2013). The copper supply pathway to a *Salmonella* Cu,Zn-superoxide dismutase (SodCII) involves P(1B)-type ATPase copper efflux and periplasmic CueP. *Mol. Microbiol.* 87, 466–477. doi: 10.1111/mmi.12107

Osman, D., Waldron, K. J., Denton, H., Taylor, C. M., Grant, A. J., Mastroeni, P., et al. (2010). Copper homeostasis in *Salmonella* is atypical and copper-CueP is a major periplasmic metal complex. *J. Biol. Chem.* 285, 25259–25268. doi: 10.1074/jbc.M110.145953

Outten, F. W., Huffman, D. L., Hale, J. A., and O'Halloran, T. V. (2001). The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*. *J. Biol. Chem.* 276, 30670–30677. doi: 10.1074/jbc.M104122200

Petrovska, L., Mather, A. E., AbuOun, M., Branchu, P., Harris, S. R., Connor, T., et al. (2016). Microevolution of monophasic *Salmonella* Typhimurium during epidemic, United Kingdom, 2005–2010. *Emerg. Infect. Dis.* 22:617. doi: 10.3201/eid2204.150531

Pontel, L. B., and Soncini, F. C. (2009). Alternative periplasmic copper-resistance mechanisms in Gram negative bacteria. *Mol. Microbiol.* 73, 212–225. doi: 10.1111/j.1365-2958.2009.06763.x

Rabsch, W. (2007). "*Salmonella* Typhimurium Phage Typing for Pathogens," in *Salmonella, Methods and Protocols*, 394 Edn, eds H. Schatten and A. Eisenstark (Totowa, NJ: Humana Press), 177–212.

Rabsch, W., Tschape, H., and Baumler, A. J. (2001). Non-typhoidal salmonellosis: emerging problems. *Microb. Infect.* 3, 237–247.

Rensing, C., and Grass, G. (2003). *Escherichia coli* mechanisms of copper homeostasis in a changing environment. *FEMS Microbiol. Rev.* 27, 197–213.

Salyers, A. A., Shoemaker, N. B., Stevens, A. M., and Li, L.-Y. (1995). Conjugative transposons: an unusual and diverse set of integrated gene transfer elements. *Microbiol. Rev.* 59, 579–590.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Staehlin, B. M., Gibbons, J. G., Rokas, A., O'Halloran, T. V., and Slot, J. C. (2016). Evolution of a heavy metal homeostasis/resistance island reflects increasing copper stress in enterobacteria. *Genome Biol. Evol.* 8, 811–826. doi: 10.1093/gbe/evw031

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690.

Stoyanov, J. V., Hobman, J. L., and Brown, N. L. (2001). CueR (YbbI) of *Escherichia coli* is a MerR family regulator controlling expression of the copper exporter CopA. *Mol. Microbiol.* 39, 502–511.

Threlfall, E. J. (2000). Epidemic *Salmonella* typhimurium DT 104 - a truly international multiresistant clone. *J. Antimicrob. Chemother.* 46, 7–10. doi: 10.1093/Jac/46.1.7

Wozniak, R. A., and Waldor, M. K. (2010). Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. *Nat. Rev. Microbiol.* 8:552. doi: 10.1038/nrmicro2382

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.