Impact of the accessory genome upon the biology of African *Salmonella enterica* serovar Typhimurium ST313

Thesis submitted in accordance with the requirements of
the University of Liverpool for the degree of Doctor of Philosophy

by

Xiaojun Zhu

March 2022

Abstract

Impact of the accessory genome upon the biology of African *Salmonella enterica* serovar Typhimurium ST313

Author: **Xiaojun Zhu**

Invasive non-typhoidal *Salmonella* disease (iNTS) is caused by certain variants of *Salmonella* enterica serovar Typhimurium, and is a particular clinical problem in sub-Saharan Africa. A novel sequence type of *S.* Typhimurium, ST313, is responsible for about two thirds of human bloodstream infections. Previously, RNA-seq based comparative transcriptomic analysis of ST313 and the well-studied ST19 sequence type identified genes and sRNAs that were differentially expressed in 16 infection-relevant growth conditions, but many of these genes remain to be studied at a functional level. This thesis is focused on the investigation of the links between the accessory genome and differentially-expressed genes of *S.* Typhimurium ST313 and ST19.

Several ST313 isolates from the UK were recently discovered that are closely related to ST313 lineages from Africa, and cause gastroenteritis but not bloodstream infection. The genome-based analysis of the UK ST313 isolates revealed several homologs of novel prophages BTP1 and BTP5 in African ST313. An experimental approach showed that these homologs of BTP1 have a similar host range to the African ST313-derived BTP1 phage, and a range of different levels of spontaneous induction.

Recent phylogenetic analysis of African ST313 identified a novel sub-lineage, L2.2 represented by strain D37712. RNA-seq based transcriptomic analysis showed high expression of SPI2 genes by L2.2 in a non-inducing (NonSPI2) growth condition compared with ST313 lineage L2. This finding was confirmed by a GFP transcriptional fusion-based strategy and flow cytometry. A series of competition experiments revealed that the novel ST313 sub-lineage L2.2 has greater fitness than ST313 L2 in the NonSPI2 growth condition.

The comparative transcriptomic data of ST313 identified repression of the essential chromosomal gene *cysS*, coupled with high level of expression of a paralog of *cysS* gene carried by the novel plasmid pBT1. The expression of the chromosomal *cysS* promoter was found to be repressed by pBT1, and modulated by a novel pBT1-encoded genetic region between *pBT1-0051* and *traL*.

By providing new information about the accessory genome of *Salmonella* ST313, as well as the relative fitness of Lineage 2/Lineage 2.2, this thesis contributes to a better understanding of the evolution and success of *Salmonella* ST313 in Africa.

Acknowledgements

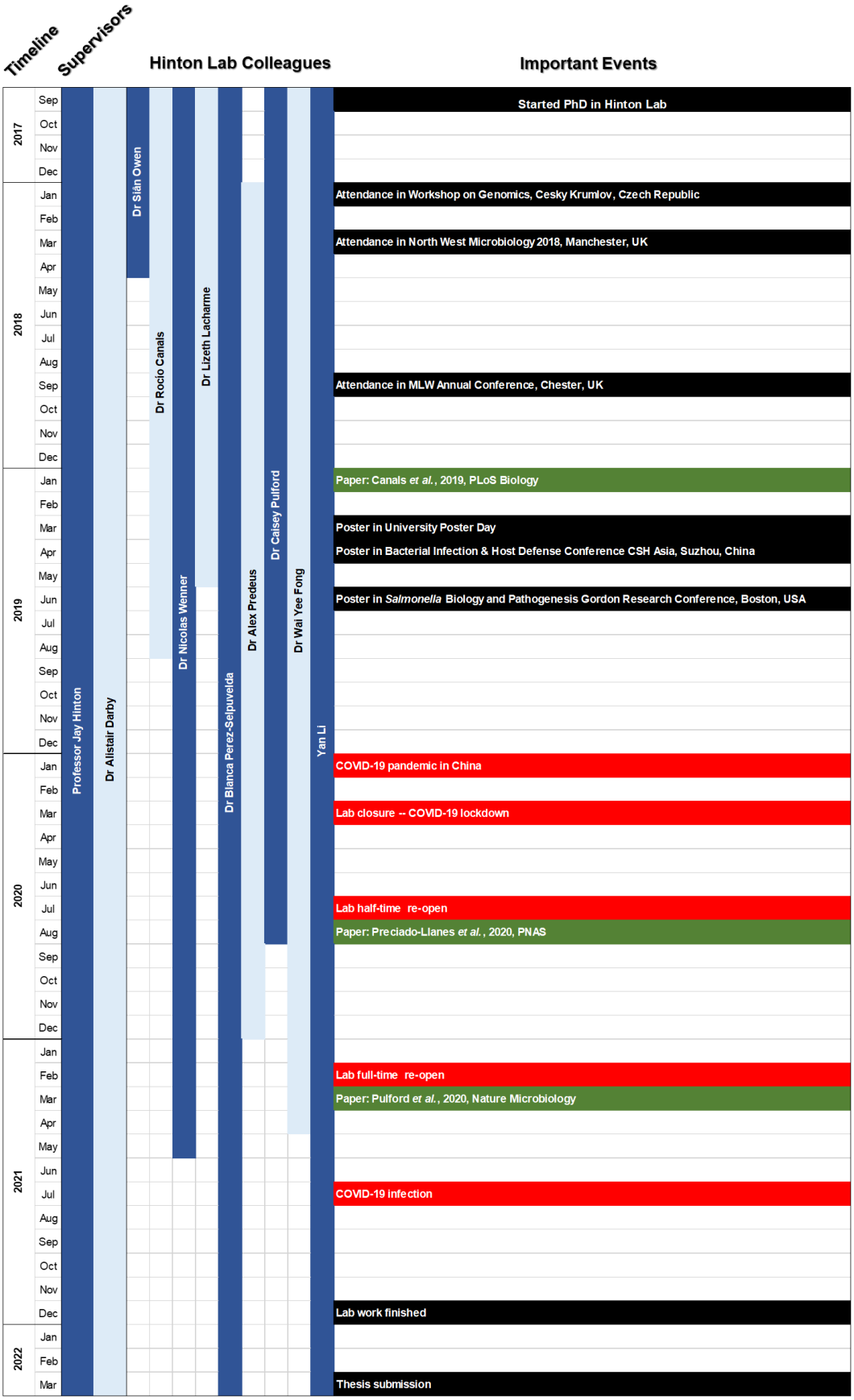
First, I owe Professor Jay Hinton, my primary supervisor, a debt of gratitude for welcoming me to his lab for my PhD project. Jay has guided and supported me throughout my undergraduate studies, Master's project and my PhD. He provided me with invaluable scientific advice and psychological support, and I learnt a lot from him about how to be an excellent scientist. I also thank Dr. Alistair Darby, my second supervisor, for his support and supervision. I thank my internal assessors, Dr. Heather Alison and Dr. John Bollback, for their comments and suggestions during and after our assessor panel meetings.

I am grateful to all past and present postdocs and colleagues in the Hinton Lab, namely Dr. Rocío Canals, Dr. Nico Wenner, Dr. Siân Owen, Dr. Lizeth Lacharme-Lora, Dr. Will Rowe, Dr. Blanca Perez-Sepulveda, Dr. Alex Predeus, Dr. Ben Kumwenda, Dr. Caisey Pulford, Dr. Wai Yee Fong, Yan Li, Wenqing Hu, Ella Rodwell, Arthur Bowers-Barnard and Hermione Webster. In particular, I want to express my gratitude to Nico for his constant assistance and supervision in my PhD project, to Siân for her help and support with my phage-related analyses, and to Rocío for teaching me to interpret transcriptomic and genomic data.

I would like to thank all past and present PIs, postdocs and colleagues in Lab H and IVES (IIB) in the Biosciences Building over the years. You are all wonderful colleagues and I really enjoyed the academic environment in the lab. I was particularly fortunate to have such a conscientious and resourceful lab manager Paul Loughnane, who always found the materials I needed and looked after equipment for my project.

Finally, I would like to express my gratitude to my girlfriend, family and all of my friends for their unwavering support and friendship throughout my PhD, particularly during the COVID-19 outbreaks. I am grateful that you are always there to support me during such a difficult time.

My PhD Journey



Publications and Conferences

Co-authored publications during the duration of the PhD

Canals, R., Hammarlöf, D. L., Kröger, C., Owen, S. V., Fong, W. Y., Lacharme-Lora, L., **Zhu, X.**, Wenner, N., Carden, S. E., Honeycutt, J., Monack, D. M., Kingsley, R. A., Brownridge, P., Chaudhuri, R. R., Rowe, W. P. M., Predeus, A. V., Hokamp, K., Gordon, M. A. and Hinton, J. C. D. (2019) Adding function to the genome of African *Salmonella* Typhimurium ST313 strain D23580. *PLoS Biology*, **17**(1), e3000059. doi:10.1371/journal.pbio.3000059.

Preciado-Llanes, L., Aulicino, A., Canals, R., Moynihan, P. J., **Zhu, X.**, Jambo, N., Nyirenda, T. S., Kadwala, I., Gerós, A., Owen, S. V., Jambo, K., Kumwenda, B., Veerapen, N., Besra, G. S., Gordon, M. A., Hinton, J. C. D., Napolitani, G., Salio, M. and Simmons A. (2020) Evasion of MAIT cell recognition by the African *Salmonella* Typhimurium ST313 pathovar that causes invasive disease. *PNAS*, **117**(34), 202007472. doi:10.1073/pnas.2007472117

Pulford, C. V., Perez-Sepulveda, B. M., Canals, R., Bevington, J. A., Bengtsson, R. J. , Wenner, N., Rodwell, E. V., Kumwenda, B., **Zhu, X.**, Bennett, R. J., Stenhouse, G.E., De Silva, P. M., Webster, H. J., Bengoechea, J. A., Dumigan, A., Tran-Dien, A., Prakash, R., Banda, H. C., Alufandika, L., Mautanga, M. P., BowersBarnard, A., Beliavskaia, A. Y., Predeus, A. V., Rowe, W., Darby, A., Hall, N., Weill, F. X., Gordon, M. A., Feasey, N. A., Baker, K. S. and Hinton, J. C. D. (2020). The stepwise evolution of *Salmonella* Typhimurium ST313 responsible for bloodstream infection in Africa. *Nature Microbiology*, **6**(3), 1-12. doi:10.1038/s41564-020-00836-1

Yang, M., Wenner, N., Dykes, G., Li, Y., **Zhu, X.**, Sun, Y., Huang, F., Hinton, J. C. D. and Liu, L. (2022). Biogenesis of a bacterial metabolosome for propanediol utilization. *Nature Communications*, **13**, 2920. doi:10.1038/s41467-022-30608-w

.

Contents

List of Figures

List of Tables

List of Abbrevations

ACT	Artemis comparison tool
AnTc	Anhydrotetracycline
Ap	Ampicillin
BLAST	Basic Local Alignment Search Tool
bp	Base pair
CFU	Colony forming units
Cm	Chloramphenicol
Contig	Contiguous sequence
CRISPR	Clustered regularly interspaced short palindromic repeats
DNA	Deoxyribonucleic acid
ESP	Early stationary phase
FC	Fold change
FDR	False discovery rate
FLP	Flippase recombinase
FRT	Flippase recognition target
GFP	Green Fluorescent Protein
Gm	Gentamicin
HGT	Horizontal gene transfer
HIV	Human immunodeficiency virus
InSPI2	SPI2-inducing condition
iNTS	Invasive nontyphoidal *Salmonella*
Km	Kanamycin
LB	Lysogeny broth
MDR	Multi-drug resistant
MGE	Mobile genetic element
MLST	Multi-locus sequence typing
NCBI	National Center for Biotechnology Information
NoO2	Anaerobic growth condition

NonSPI2	Non-SPI2 inducing contion
NTS	Nontyphoidal *Salmonella*
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PFU	Plaque forming units
PP	Peyer’s patches
RDAR	Red, dry and rough
RFP	Red Fluorescent Protein
RNA	Ribonucleic acid
RNA-seq	RNA sequencing
SCV	*Salmonella*-containing vacuole
SNP	Single nucleotide polymorphism
SPI	*Salmonella* Pathogenicity Island
sRNA	Small RNA
ST	Sequence type
TA system	Toxin/antitoxin system
Tc	Tetracycline
TIS	Transposon insertion sequencing
Tn	Transposon
TPM	Transcipts per million
tRNA	Transfer RNA
T3SS	Type 3 secretion system
TSS	Transciption start site
UV	Ultraviolet
WT	Wild type

List of Appendices

Chapter 1.

Introduction

1.1 *Salmonella* genus

Salmonella is a genus of Gram-negative, facultatively anaerobic rod-shaped bacteria that belongs to the family Enterobacteriaceae (Eng *et al.*, 2015; Lamas *et al.*, 2018). This bacterium was first discovered by Theobald Smith, an American microbiologist, and then named by his supervisor Daniel E. Salmon (Eng *et al.*, 2015). *Salmonella* shares many similarities with other members of the family *Enterobacteriaceae*, such as *Escherichia coli*. However, *Salmonella* differs from other bacterial genera because it does not ferment lactose but can ferment glucose (Cox *et al.*, 2000). It is estimated that *Salmonella* shared a common ancestor with *Escherichia* and *Shigella* around 100 million years ago (Doolittle *et al.*, 1996).

The *Salmonella* genus contains two species, *S. bongori* and *S. enterica*. *S. bongori* is mainly found in cold-blooded animals, and shares a closer ancestor with *Escherichia* than with *S. enterica*. It was suggested that *S. bongori* was an intermediate between *Salmonella* and *E. coli* (Fookes *et al.*, 2011). The *S. enterica* species is divided into 6 subspecies: (I) *S. enterica*, (II) *S. salamae*, (IIIa) *S. arizonae*, (IIIb) *S. diarizonae*, (IV) *S. houtenae* and (VI) *S. indica* (Lamas *et al.*, 2018). The evolution of *Salmonella* is shown in Figure 1.1 (Alikhan *et al.*, 2018). The *Salmonella* subspecies are further classified using a method called serotyping, and the taxonomic level below subspecies are called serovars. This method follows the White-Kauffmann-Le Minor Scheme to distinguish *Salmonella*, and relies upon two main antigenic determinants on the cell surface: the bacterial lipopolysaccharide

(somatic O antigen) and the flagella (H antigen) (Grimont and Weill, 2007). The H antigen of *Salmonella* is in two forms, phase 1 and phase 2, which are synthesised by the flagella genes *fliC* and *fljB*, respectively (McQuiston *et al.*, 2008).

The formula of a serotype is written as “O antigen: H antigen phase 1: H antigen phase 2”. For example, *Salmonella enterica* serovar Typhimurium is written as “1,4,[5],12:i:1,2”, which means this strain carries the O antigen factor 1, 4, 5 and 12 (the square brackets means that the factor may not be present), the H antigen i (phase 1) and the H antigen 1 and 2 (phase 2). To date, over 2,500 serovars have been identified and 59% of them belong to subspecies *S. enterica* (Brenner *et al.*, 2000; Grimont and Weill, 2007). Less than 100 serovars cause the majority of human infections (CDC, 2015). To aid the readability of this thesis, the abbreviation *S.* Typhimurium is used in place of Sal*monella enterica* serovar Typhimurium throughout.

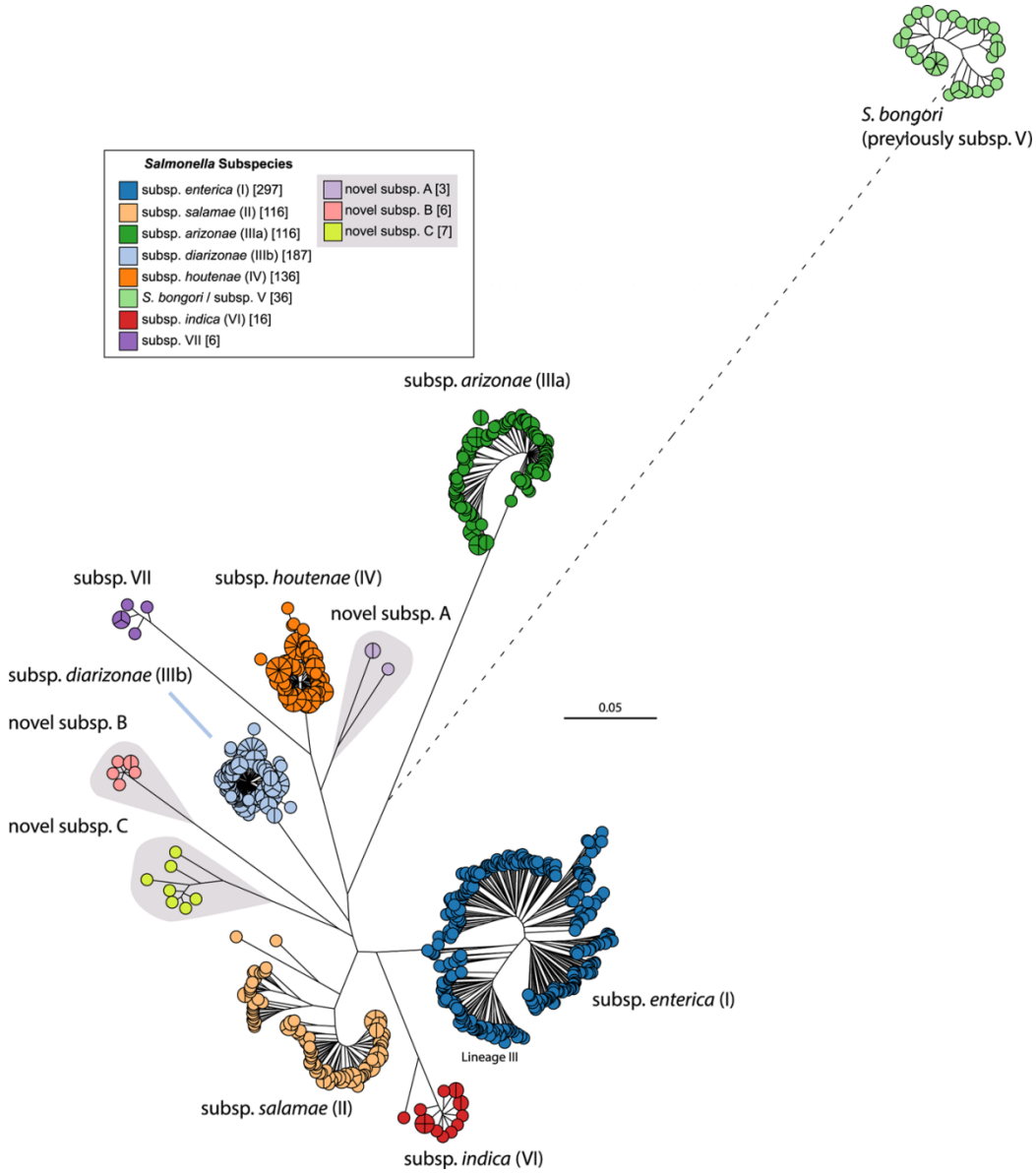


Figure 1.1. The evolution of the *Salmonella* genus. The grapetree represents a maximum-likelihood phylogeny of core SNPs from 926 representative genomes of *S. enterica* plus *S. bongori*. Pie charts at the tips represent different genomes and the slices in a pie chart mean multiple genomes are located at the same tip. The figure is modified from Alikhan *et al.*, 2018

The Multi-Locus Sequence Typing (MLST) was developed by Martin Maiden, Mark Achtman and Brian Spartt in 1998 (Martin *et al.*, 1998), and is an excellent and cost-effective approach for discriminating between *Salmonella* isolates (Achtman *et al.*, 2012; Ashton *et al.*, 2016). Initially, this method involves the sequencing of seven

highly-conserved housekeeping genes from *Salmonella* isolates: *aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA* and *thrA*. More recently, whole genome sequences are used to extract the same information (Yang *et al.*, 2022). Each time a new combination of the seven alleles is identified, a new Sequence Type (ST) designation is made which is recorded at the PubMLST website (https://pubmlst.org/organisms/salmonella-spp). Two examples of sequence types that are relevant to this thesis are the well-studied *S.* Typhimurium variant ST19 that causes human gastroenteritis around the world, and the novel African *S.* Typhimurium variant ST313, that is described in Section 1.6.

To put the different *Salmonella* typing methods into perspective, Figure 1.2 shows a timeline of the development of the key characterisation methods (Tang *et al.*, 2019). Since the first development of serotyping in 1934, MLST became widely used to differentiate *Salmonella* serovars in 2012. More recently, whole genome sequencing and CRISPR-typing have become the standard approaches for distinguishing *Salmonella* isolates (Tang *et al.*, 2019).

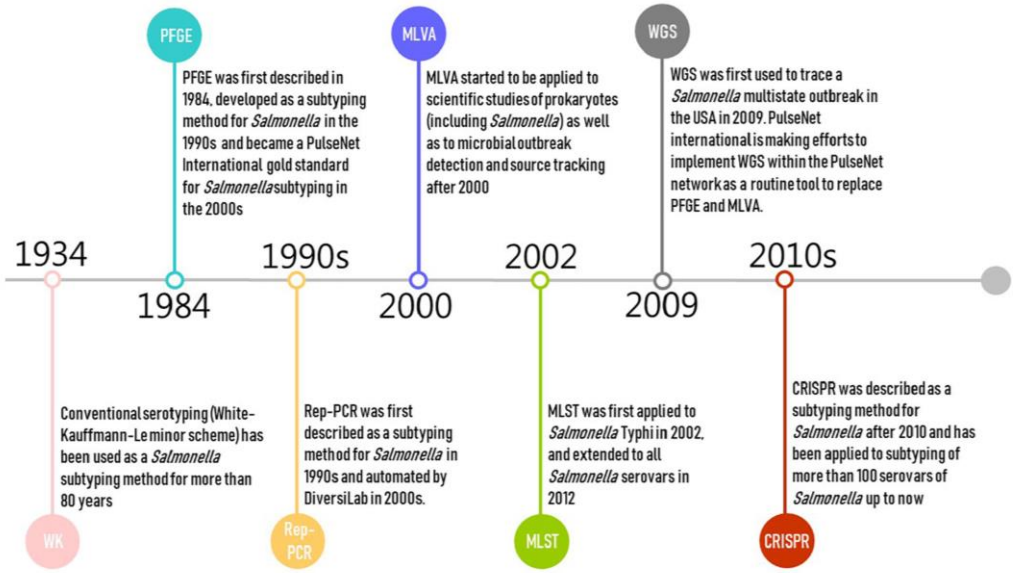


Figure 1.2. Timeline of the development of selected molecular subtyping and characterisation methods for *Salmonella*. The figure is modified from Tang *et al.*, 2019.

1.2 Salmonellosis is a zoonotic disease that causes human infections

Salmonellosis is a zoonotic disease, meaning that it can be transmitted between animals and humans (Stevens *et al.*, 2009). Human infections can originate from pets (Dróżdż *et al.*, 2021), farm animals (Rukambile *et al.*, 2019) or food (De Cesare, 2018).

Overall, *Salmonella* genus has a broad host range. *S. bongori* and *S. enterica* subspecies other than subspecies *enterica* are most frequently isolated from the environment or from cold-blooded animals such as venomous snakes and other reptiles (Brenner *et al.*, 2000; Pulford *et al.*, 2019). The *S. enterica* serovars that infect animals have been classified according to host range as host-restricted (e.g. *S.* Typhi, which only infects human), host-adapted (e.g. *S.* Dublin, which infects

cattles and human) or host-generalist (e.g. *S.* Typhimurium, which infects a broad range of species, including humans, livestock, rodents and birds) (Uzzau *et al*., 2000). Previous research identified that over 200 *S.* Typhi genes are non-functional pseudogenes, most of which are functional in *S.* Typhimurium, which may help to explain the difference between the host range of different *Salmonella* serovars (de Jong *et al.*, 2012; Gal-Mor *et al.*, 2014).

Salmonella pathogens are also categorised according to clinical syndromes: typhoidal *Salmonella*, which causes invasive systemic disease, and non-typhoidal *Salmonella* (NTS), which generally causes gastroenteritis. Typically, the host range of *Salmonella* and the diseases they cause are linked (Figure 1.3). The most host-restricted *Salmonella* is more likely to cause invasive diseases and results in a high mortality, while more host-generalist *Salmonella* causes self-limiting gastroenteritis.

The concept of *Salmonella* pathovariants has been proposed recently, and builds on the differing host ranges of closely-related pathogens such as *S.* Typhimurium ST19 and ST313 (Branchu *et al.*, 2018).

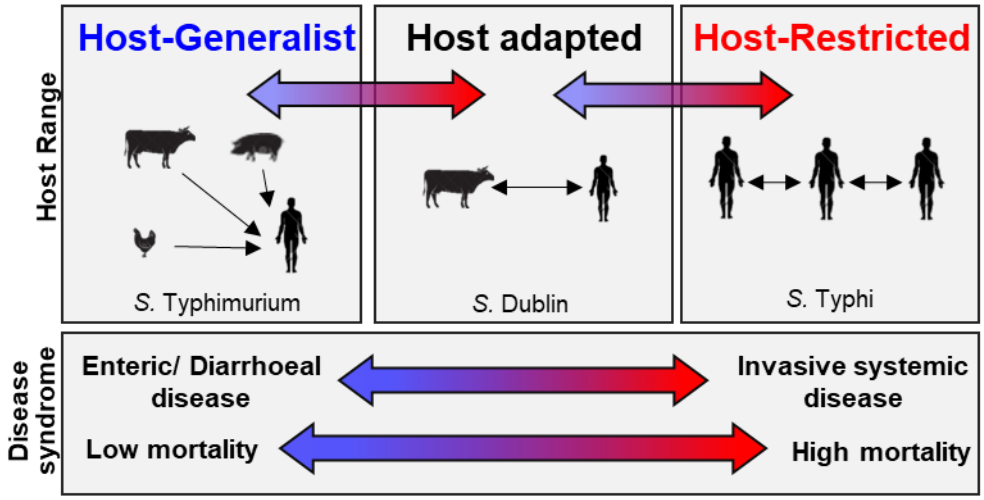


Figure 1.3. The host adaptation and disease syndrome paradigm. The top row shows Salmonella infection models for different groups of serovars classified by host range, with an example serovar shown at the bottom. The lower row shows the relationship between host range and pathogenicity. The figure is modified from S. Owen *Pers. Comm.*

1.3 The epidemiology of *Salmonella* infections

Salmonella infection is a major public health issue worldwide (Eng *et al.*, 2015). The infection of *S.* Typhi and *S.* Paratyphi causes invasive systemic disease which is called typhoid fever or enteric fever (and paratyphoid fever). Typhoidal disease results in sustained fever, chill, abdominal pain and hepatosplenomegaly, and sometimes involves diarrhoea or constipation (Buckle *et al.*, 2012). Typhoid and paratyphoid are common in countries with low-income and poor water quality (Mogasale *et al.*, 2014). In 2017, there was an estimated 14.3 million cases of typhoid and paratyphoid fevers, resulting in approximately 135,900 deaths (Stanaway *et al.*, 2019b).

Non-typhoidal *Salmonella* (NTS) pathovariants are generally not host-specific and cause self-limiting gastroenteritis in humans. The broad host range of NTS leads to frequent human infections via the food chain. The gastroenteritis is characterised by symptoms such as non-bloody diarrhoea, vomiting, nausea and headache. People with a weakened immune system such as infants, elders and HIV-infected patients (Eng *et al.*, 2015) are more likely to be infected by NTS. *S.* Typhimurium and *S.* Enteritidis are the two most common NTS serovars involved in the global epidemiology of this desease. In 2010, it was estimated that there was 93.8 million cases of gastroenteritis and 155,000 deaths caused by NTS and over 80 million of these cases were foodborne (Majowicz *et al.*, 2010).

1.4 *Salmonella* virulence factors

The infection biology of *Salmonella* pathogens has been studied intensively since the 1950s (dos Santos *et al*., 2019; Tsolis *et al.*, 2011), and a large number of virulence genes have been identified. One key virulence factor of *S.* Typhi is the virulence capsular polysaccharide, known as the Vi antigen. The Vi antigen is a major virulence factor that modulates pro-inflammatory signaling pathways involved in host denfense and associated with typhoid fever (Schadich *et al.*, 2016). The Vi antigen is found only in *S.* Typhi, Paratyphi and Dublin, and is absent from other serovars that cause gastroenteritis (Jajere, 2019).

Many of the virulence factors are associated with *Salmonella* pathogenicity islands (SPIs). The development of whole genome sequencing in 1995 (Fleischmann *et al.*, 1995), led to the discovery of pathogenicity islands as groupings of virulence genes on the bacterial chromosome (Hacker *et al.*, 2000). Other factors are located elsewhere on the chromosome or are associated with accessory genome components such as plasmids and prophages (Fàbrega and Vila, 2013).

The best characterised virulence factors are the effector proteins that are translocated via the SPI1 and SPI2 Type 3 Secretion System (T3SS), and interact directly with the host cell machinery (Galán, 2021). The importance of the SPI1 and SPI2 was found in 1995 (Mills *et al.*, 1995; Shea *et al.*, 1996), and both SPI1 and SPI2 encode a T3SS. Since then, a total of 12 *Samonella enterica* SPIs have been identified (Sabbagh *et al.*, 2010). The SPI1-encoded T3SS delivers effectors proteins that trigger cytoskeleton rearrangement of host cells and facilitate the invasion of the intestinal epithelium (Lara-Tejero *et al.*, 2009). SPI2 is essential for systemic virulence of *Salmonella* (Waterman *et al.*, 2003). The SPI2-encoded T3SS is active intracellularly, within *Salmonella*-containing vacuoles (SCVs) and translocates effectors to the host cell cytosol and to the SCV membranes to promote survival in macrophages (Ilyas *et al.*, 2017).

1.5 *Salmonella* Pathogenesis

It can be argued that infection of mammalian hosts by *S.* Typhimurium ST19 is the best understood infection model in the world (Tsolis *et al.*, 2011). The first experiments involving inoculation of *S.* Typhimurium into mice were done in 1982 (Loeffler, 1982). Infection by NTS such as *S.* Typhimurium is a complex process which can involve three major stages: colonization of the gastrointestinal tract, invasion of epithelial cells and survival in immune cells (Ilyas *et al.*, 2017). Briefly, after oral infection through contaminated food or water, the bacterium enters the stomach and faces an acidic environment. After entering the small intestine, *Salmonella* competes for nutrients and encounters colonization resistance mediated by the resident microbiota (Patel and McCormick, 2014). *S.* Typhimurium then invades both enterocytes and M cells (specialized epithelial cells) of the Peyer's patches (PPs) in the intestinal epithelium, with the help of the SPI1-encoded T3SS or flagella (Hiyoshi *et al.*, 2018). Invasion of the host cells lead to the re-arrangement of cytoskeleton to form membrane ruffles that engulf *Salmonella* into *Salmonella*-containing vacuoles (SCVs). Intracellular replication of *Salmonella* occurs in the SCV (de Jong *et al.*, 2012). Invasion of epithelial cells by *S.* Typhimurium involves expression of flagella and the SPI1 T3SS which triggers an immune response that results in phagocytosis by neutrophils, dendritic cells and macrophages. *Salmonella*-containing neutrophils can migrate back to the intestinal lumen to cause gastroenteritis (Fàbrega and Vila, 2013).

Infection by *S.* Typhi results in a severe systemic disease. During the systemic infection, *S.* Typhi enter the lymphatic pocket through the SCVs followed by phagocytosis by macrophages. This intracellular lifestyle can protect *Salmonella* from the immune system, and allow replication (Petersen and Miller, 2019). The infected macrophages transport *S.* Typhi to the lymph nodes, and then into liver and spleen via bloodstream (Fàbrega and Vila, 2013).

Multiple mechanisms have been identified that moderate intestinal inflammation by typhoidal Salmonella serovars and result in systemic diseases. The repression of flagella and SPI1 T3SS gene expression in *S.* Typhi allows the pathogen to evade the innate immune response (Hiyoshi *et al.*, 2018). Another important *S.* Typhi virulence factor is the Vi antigen described in Section 1.4, which inhibits complement activation to avoid the cytokine response (Schadich *et al.*, 2016).

Unlike typhoidal *Salmonella* serovars, most NTS pathogens are unable to evade the innate immune system, and therefore cause inflammation in the human gut. However, some NTS bacteria cause invasive systemic disease (iNTS), specifically in immunocompromised patients (Feasey *et al.*, 2012; Gal-Mor *et al.*, 2014; Shah *et al.*, 2017). Compared with NTS strains, iNTS strains possess several pathogenic characteristics that are associated with invasive diseases. Reduced expression of SPI-1 and flagella associated genes are found in iNTS strains, which lead to the reduced neutrophil recruitment and an attenuated inflammatory response in

macrophages (Carden *et al.*, 2015). Recent transcriptomic comparison of NTS and iNTS strains identified a single nucleotide polymorphism (SNP) in the *pgtE* promoter. This SNP results in increased resistance of iNTS strains to human serum, with implications for bloodstream infection (Hammarlöf *et al.*, 2018). Further details of iNTS pathogenesis remain under investigation (Gilchrist *et al.*, 2019).

1.6 Invasive Non-typhoidal *Salmonella* ST313 in sub-Saharan Africa

In recent years, iNTS has become the most common bloodstream infection in sub-Saharan Africa (Marchello *et al.*, 2019). In 2017, the Global Burden of Disease study estimated that there were 535,000 cases of iNTS infection which lead to 77,500 deaths worldwide (Stanaway *et al.*, 2019a). Most cases of iNTS occurred in sub-Saharan Africa in countries with low incomes and a high prevalence of HIV, malaria and malnutrition (Piccini and Montomoli, 2020). The iNTS has been confirmed as a HIV/AIDS related disease (Feasey *et al.*, 2012). Both HIV and malaria attenuate the human immune response, which allows S. Typhimurium to grow intracellularly. Malnutrition may change the integrity and acidity of the intestinal barrier (Piccini and Montomoli, 2020). The intestinal barrier is a complex environment that involves the human gut microbiome, and plays an important role in immune sensing against pathogens (Baumgart and Dignass, 2002).

In sub-Saharan Africa, *S.* Typhimurium and *S.* Enteritidis are the two major serovars associated with iNTS bloodstream infections (Uche *et al.*, 2017). *S.* Typhimurium

causes two-thirds of iNTS disease, and by using whole genome sequencing, it was found that the *S.* Typhimurium sequence type ST313 was the leading cause of iNTS (Kingsley *et al.*, 2009). Subsequent studies identified two distinct phylogenetic lineages of ST313: Lineage 1 (L1) and Lineage 2 (L2) (Okoro *et al*., 2012). It was estimated that L1 arose in South-West Africa in about 1960 and that L2 emerged in about 1977 in Malawi (Okoro *et al.*, 2012). More recently, a new Lineage 3 of *S.* Typhimurium ST313 (Pulford *et al.*, 2021) and also two sub-lineages L2.1 (Van Puyvelde *et al.*, 2019) and L2.2 (Msefula *et al.*, 2012) have been identified as causal agents of iNTS in Africa. Among these lineages, L2 is the most studied and L2 strain D23580 (Accession Number: FN424405) is now used as a reference strain for *S.* Typhimurium ST313.

Antibiotic resistance and accumulated loss-of-function mutations are two main reasons for the success of ST313 in Africa (Pulford *et al.*, 2021). *S.* Typhimurium ST313 carries the virulence plasmid pSLT combined with different antibiotic resistance genes (Kingsley *et al.*, 2009). L1 and L2 had different antibiotic resistance and it was estimated that clonal replacement of ST313 L1 by L2 occurred due to the acquisition of chloramphenicol resistance by L2 (Kingsley *et al.*, 2009; Okoro *et al.*, 2012). Compared to ST313 L2, L3 strains contain pseudogenes including *ratB,* which is associated with gut persistence in the mouse infection model, and *pipD,* which contributes to macrophage persistence in the mouse model for ST19 (Pulford *et al.*, 2021). Apart from the ST313 lineages identified in Africa, *S.* Typhimurium

ST313 isolates have also been found in Brazil (Almeida *et al.*, 2017) and the United Kingdom (Ashton *et al.*, 2017). A thorough epidemiological analysis proved that the UK-ST313 isolates caused a self-limiting gastroenteritis, and not an invasive systemic disease (Ashton *et al.*, 2017).

The genome of the ST313 L2 reference strain D23580 was used for a phylogenetic analysis which indicated that ST313 is closely related to the characterised *S.* Typhimurium ST19, responsible for gastroenteritis worldwide (Kingsley *et al.*, 2009). Comparative genomic investigation of D23580 and the ST19 reference strain LT2 (McClelland *et al.*, 2001) identified key differences between the two strains (Kingsley *et al.*, 2009). Compared with LT2, D23580 had 44 novel pseudogenes and deletions. One example of the pseudogenes is *katE*, which is required for resistance against peroxide (Singletary *et al.*, 2016). Another mutation generates a *bcsG* pseudogene. BcsG is a cellulose biosynthetic enzyme that is required for the RDAR (red, dry and rough) morphotype (Singletary *et al.*, 2016), which is responsible for biofilm formation (Romling *et al.*, 1998).

Comparative genomic investigation showed that the accessory genome of ST313 L2 is distinct from ST19. A core genome contains genes present in all strains, whilst the accessory genome varies between particular strains. An accessory genome contains chromosomal regions and also mobile genetic elements (MGEs) including plasmids and prophages, which play an important role in the evolution and adaption of

bacteria (Shintani, 2017). This thesis focuses on MGEs of the accessory genome of *S.* Typhimurium ST313.

D23580 carries a distinct set of MGEs from LT2. D23580 carries two novel prophages BTP1 and BTP5, and three novel plasmids pBT1, pBT2 and pBT3 (BT stands for Blantyre Type; Blantyre is a major city in Malawi; Kingsley *et al.*, 2009). BTP1 contains a *gtrC* gene, which is involved in O antigen modification (Kintz *et al.*, 2015). There is also a novel BTP1-encoded *bstA* gene, which defends host D23580 cells against attack by a range of exogenous phage (Owen *et al.*, 2021). The other novel MGEs of D23580 have not been well characterised, but the plasmid biology and phage biology is discussed below.

1.7 Plasmids

Plasmids are circular or linear extra-chromosomal DNA molecules found in most bacteria and they are able to autonomously replicate in a host cell (Jackson *et al.*, 2011; Orlek, *et al.*, 2017). Plasmids were named in 1952 by the American molecular biologist Joshua Lederberg and then first characterised in the early 1960s (Rodríguez-Beltrán *et al.*, 2021). Plasmids can differ in size, GC content and copy number. The average size of bacterial plasmids from the GenBank database is 80 kb and the average GC content is about 44% (Shintani *et al.*, 2015). Plasmids can be classified into two groups based on their copy numbers in the host cell: LCPs (low copy number plasmids) and HCPs (high copy number plasmids). Typically, LCPs

have a larger size and process a functional conjugative system, while HCPs are small and not conjugable (Rodríguez-Beltrán *et al.*, 2021).

Horizontal gene transfer (HGT) is the exchange of genetic material between organisms (Emamalipour *et al.*, 2020). Plasmids play a significant role by driving the HGT process for bacteria; plasmids contain a large amount of DNA and frequently carry genes that have important survival functions for bacteria (Jackson *et al.*, 2011). Plasmids normally do not encode essential bacterial genes, but contribute to the adaption of the host to stressful environments or provide antibiotic resistance (Carroll and Wong, 2018).

Plasmid carriage is often beneficial to the host cell (Kottara *et al.*, 2018). Benefits can include antibiotic resistance, virulence, carbon source metabolism and tolerance to metals (Carroll and Wong, 2018). Resistance to aminoglycosides, β-lactams and tetracyclines are the most common resistances encoded by plasmids (Pal *et al.*, 2015). Nowadays HGT is driving increased dissemination of *Klebsiella* multidrug-resistant (MDR) plasmids (Pedersen *et al.*, 2020), and the transfer of MDR plasmids containing the genetic code for NDM-1 (New Delhi metallo-beta-lactamase 1) is resulting in the global emergence of “superbugs” (Sun *et al.*, 2019). Some plasmids carry transposons (Tn) which are discrete DNA elements that are able to move within a cell. For example, transposons encoding mercury resistance in

Pseudomonas transfer between distinct species via plasmids (Rodríguez-Beltrán *et al.*, 2021).

Virulence is another important property encoded by bacterial plasmids. Virulence plasmids are normally low copy number and are able to mediate conjugation to drive the transfer of themselves (Pilla and Tang, 2018). Virulence plasmids are abundant in enteric pathogens. Enterohaemorrhagic *E. coli* (EHEC) serotype O157:H7 contains the pO157 virulence plasmid, which is associated with haemolytic uraemic syndrome (HUS) and promotes adhesion of the bacteria onto intestinal epithelial cells (Johnson and Nolan, 2009). *Salmonella* subspecies I contains the pSLT virulence plasmid which carries the *spvABCDR* locus which plays a key role in pathogenesis (Guiney and Fierer, 2011). This virulence region is inserted into the *Salmonella* chromosome in certain subspecies (Silva *et al.*, 2017). Specifically, the *spvABCD* operon encodes proteins which enhance intra-macrophage replication of *Salmonella* and contributes to systemic infection (Pilla and Tang, 2018).

S. Typhimurium ST313 L2 strain D23580 carries three novel plasmids pBT1, pBT2 and pBT3 and a derivative of the virulence plasmid called pSLT-BT (Kingsley *et al.*, 2009). The plasmid pSLT-BT in D23580 contains an inserted Tn*21*-like element that carries the chloramphenicol resistance gene *cat* (Kingsley *et al.*, 2009; Pulford *et al.*, 2021).

1.8 Bacteriophages

Bacteriophages (phages) are viruses that infect bacteria. Phages were first discovered by an English biologist Frederick William Twort in 1915 and their ability to kill bacteria was discovered by a French microbiologist Felix d'Herelle in 1917 (Clokie *et al.*, 2011). Phages are categorised into 4 types based on the genetic material carried: single stranded DNA (ssDNA), double stranded DNA (dsDNA), single stranded RNA (ssRNA) and double stranded RNA (dsRNA). Tailed dsDNA phages in the order *Caudovirales* account for 96% of all known phages (Zinke *et al.*, 2021).

Normally phages have two distinct lifecycles: lytic and lysogenic. In a lytic cycle, phages infect cells and then kill the host cells. During a lysogenic cycle, instead of killing the host, phages integrate their sequence into the host genome and the integrated sequence is called a prophage. Prophage sequences remain stable on the *Salmonella* chromosome through hundreds of thousands of generations, and are an important component of the accessory genome (Bobay *et al.*, 2013).

Phages that have both a lytic and lysogenic cycle and are called temperate phages. One well-studied temperate phage is the Lambda phage. By making the lytic-lysogenic decision, temperate phages are able to switch between two lifecycles (Shao *et al.*, 2019). In Lambda phages, the lytic-lysogenic decision depends on the expression of CII protein (Ptashne, 2004). The CII protein is normally degraded in an

environment suitable for phage replication, resulting in a lytic cycle. The decision to enter the lysogenic cycle is thought to reflect poor environmental conditions, when the accumulation of CII activator protein triggers the synthesis of CI repressor which blocks expression of phage replication genes (Shao *et al.*, 2019). Subsequently, the phage sequence is integrated into host chromosome.

The prophages component of the accessory genome is able to contribute to bacterial virulence. One hypothesis was that phages need sufficient numbers of host cells to ensure the survival and so provide benefits to the host such as an enhancement of virulence or fitness (Brüssow, 2007; Rankin *et al.*, 2011; Wahl *et al.*, 2018). For example, the prophage Gisfy-2 in *S.* Typhimurium carries the *sodC* gene which encodes a virulence factor that detoxifies superoxide within host cells (Figueroa-Bossi and Bossi, 1999). In some cases, prophages can repress host genes. The CI repressor of Lambda phage down-regulates the growth rate of *E. coli* by regulating the host *pckA* gene which is required for gluconeogenesis (Chen *et al.*, 2005).

The *S.* Typhimurium ST313 strain D23580 carries five prophages: Gifsy-1, Gifsy-2, ST64B, BTP1 and BTP5 (Kingsley *et al.*, 2009). Gifsy-1, Gifsy-2 and ST64B have been previously studied in *S*. Typhimurium ST19 (Owen *et al.*, 2017). Both Gifsy-1 and Gifsy-2 contribute to the virulence of ST19 (Ho *et al.*, 2002; Stanley *et al.*, 2000; Wahl *et al.*, 2018). However, these three prophages Gifsy-1, Gifsy-2 and ST64B are

inactivated in D23580 due to mutations in the prophage genomes (Owen *et al.*, 2017).

BTP1 and BTP5 are two novel prophages found in D23580. Functions of some BTP1 genes were mentioned in Section 1.6, but the involvement of other BTP1 genes in *Salmonella* virulence or fitness is still unknown. BTP5 has some functional properties, but no plaques were detectable from a D23580 supernatant when plated on a lawn of phage-free D23580 (Owen *et al*, 2017). The function of BTP5 prophage remains unknown.

1.9 Bacterial fitness and competition

S. Typhimurium is a host-generalist bacterium that infects humans, chicken and other mammals. For a successful infection process, the bacteria need to survive in relatively high temperature environments (42°C in the chicken) and at different pH values (pH 2 in gastric acid and pH 9 in chicken eggs). To cause invasive disease, the pathogen must also survive oxidative stress and exposure to reactive nitrogen species with macrophages (Gogoi *et al*., 2019). In addition, *Salmonella* must compete with other bacteria to achieve growth during the infection process. The term “fitness” is defined as the ability to survive and grow in particular environments (Laurent *et al.*, 2001). Bacteria with higher fitness is tend to have a higher growth rate and a faster generation time to produce more offspring than bacteria with lower fitness (Laurent *et al.*, 2001;Wiser and Lenski, 2015).

As MGEs are able to transfer between strains and species and usually carry antibiotic resistance genes, the accessory genome plays major role in the modulation of bacterial fitness (Jackson *et al.*, 2011). One example is STM2239 which is encoded by a prophage remnant, and is a Q antiterminator protein that both interacts with RNA polymerase, and is required for optimal fitness in a host (Tomljenovic-Berube *et al.*, 2013). To survive antibiotic therapy, bacteria acquire antibiotic resistance genes by HGT to improve their fitness. However, resistance can impose fitness costs in the absence of antibiotics. MDR plasmids are the most common MGEs that increase bacterial fitness. It was recently reported that some MDR plasmid resistance genes can be cost-free, while the fitness cost imposed by a plasmid can be caused by other resistance genes (Rajer and Sandegren, 2022). Specifically, the authors discovered that only 2 of the 13 antibiotic resistance genes on *Klebsiella* plasmid pUUH239.2 imposted a fitness cost (Rajer and Sandegren, 2022).

Bacterial fitness is measured by several different approaches. One common method is to measure the growth rate of a pure bacterial culture under different growth or stress conditions (Lofton *et al.*, 2015; Wiser and Lenski, 2015). To measure the fitness cost of antibiotic resistance, minimum inhibitory concentrations (MIC) are normally measured because an increased MIC is negatively correlated with relative fitness (Melnyk *et al.*, 2014; Wiser and Lenski, 2015). A third approach for fitness

measurement involves pair-wise competition experiments, where two bacterial strains are mixed and cultured together.

The competitive index identifies bacteria that have the greatest selective advantage (Beuzón and Holden, 2001). Some fitness experiments involve long term co-cultivation, and involve daily measurements to search for the fitness decay of individual bacterial species (Heilbron *et al.*, 2014). Long-term competition experiments result in the replacement of one bacterial line by another member of the community (Wiser and Lenski, 2015), which can help to explain the evolution of bacteria species.

1.10 Thesis aims

S. Typhimurium ST313 is causing a significant public health issue in sub-Saharan Africa. The ST313 L2.2 sub-lineage (B. Kumwenda, *Pers. Comm*; Msefula *et al.*, 2012) has become increasingly prevalent in Malawi, Africa. The recent success of ST313 L2.2 has prompted researchers to question how this pathogen has become such a prominent cause of human bloodstream infection.

Comparative genomics showed that pBT1 was absent from the L2.2 reference strain D37712, and revealed the presence of the novel pBT4 plasmid (B. Kumwenda *Pers. Comm.*). RNA-seq-based transcriptomics was used to do differential gene expression analysis between L2 reference strain D23580 and ST19 reference strain

4/74 (Canals *et al.*, 2019b), and between L2 reference strain D23580 and L2.2 reference strain D37712 (B. Kumwenda *Pers. Comm.*; Kröger *et al.*, 2013). The resulting data provide the possibility of identify gene expression signatures that are linked to the accessory genome of these important *Salmonella* strains.

As described earlier, MGEs such as plasmids and prophages play an important role in the evolution, virulence and fitness of bacteria, and so the accessory genome of ST313 strains was investigated in some detail.

The aims of this PhD project were as follows:

- Compare the accessory genomes of ST313 isolates from Africa and the United Kingdom and study the prophage repertoire of UK ST313.
- Identify phenotypic and fitness differences that distinguish ST313 L2 from L2.2, based on the transcriptomic data for strain D37712.
- Use the transcriptomic data for strain D23580 as inspiration for the investigation of the function of novel plasmid pBT1. Investigate an example of regulatory cross-talk between the core genome and the accessory genome of ST313 L2.

Chapter 2.
Methods and materials

2.1 Reagents and chemicals

Reagents and chemicals used in this thesis, along with supplier information, are listed in Table 2.1.

2.2 Media and antibiotics

Recipes for all media used in this thesis are shown in Table 2.1. Water used to make up media is RO water from the PURELAB flex Dispenser. LB media was sterilized by autoclaving before use. For the minimal M9, InSPI2 and NonSPI2 media, each ingredient was prepared and sterilized by autoclaving or by filtration through a 0.22 μm filter unit (Merck Millipore). All carbon source solutions were sterilized by 0.22 μm filters before adding to media.

Media was supplemented with the following antibiotics: Ampicillin (Ap), 100 μg/ml; Chloramphenicol (Cm), 25 μg/ml; Kanamycin (Km), 50 μg/ml; Gentamicin (Gm), 20 μg/ml; Tetracycline (Tc), 20 μg/ml. The stock and final concentrations of all antibiotics are listed in Table 2.1. All antibiotics were made in H_2O, Ethanol or Methanol as detailed in Table 2.1, and were stored at -20°C before use.

Agar plates were made by aliquoting 25 ml of autoclaved molten agar into Petri dishes. Concentrations of agar used were 1.5% unless otherwise stated. Plates were left at room temperature to dry for one day and stored in sealed plastic bags at 4°C

before use. To make agar plates with antibiotics, stock solutions of antibiotics were added to the molten agar at 50°C before the agar was poured.

Table 2.1. Reagents and chemicals used in this thesis

Use	Chemical/Reagent	Supplier	Catalogue Number	Stock Concentration	Final Concentration
LB media	Tryptone	Appleton Woods	MN649	N/A	10 g/L
	NaCl	Sigma-Aldrich	S3014	N/A	5 g/L
	Bacto-Yeast Extract	Appleton Woods	DM832	N/A	5 g/L
	Bacto-Agar	Appleton Woods	214010	N/A	15 g/L
InSPI2 media	MES (pH 5.8)	Sigma-Aldrich	M8250	400 mM	80 mM
	Tricine	Sigma-Aldrich	T5817	400 mM	4 mM
	$FeCl_3$	Sigma-Aldrich	236489	100 mM	100 µM
	K_2SO_4	Sigma-Aldrich	P0772	376 mM	376 µM
	K_2HPO_4/KH_2PO_4 (pH 5.8)	N/A	N/A	100 mM	0.4 mM
	NaCl	Sigma-Aldrich	S3014	2.5 M	50 mM
	Glucose	Sigma-Aldrich	G8270	20% w/v	0.4% w/v
	NH_4Cl	Sigma-Aldrich	A9434	1.5 M	15 mM
	$MgSO_4$	Sigma-Aldrich	83266	1 M	1 mM
	$CaCl_2$	Sigma-Aldrich	C3306	1 M	0.01 mM
	Micronutrients Solution	N/A	N/A	10,000 X	1 X
NonSPI2 media	MOPS (pH 7.4)	Sigma-Aldrich	M1254	400 mM	80 mM
	Tricine	Sigma-Aldrich	T5817	400 mM	4 mM
	$FeCl_3$	Sigma-Aldrich	236489	100 mM	100 µM

	K_2SO_4	Sigma-Aldrich	P0772	376 mM	376 µM
	K_2HPO_4/KH_2PO_4 (pH 7.4)	N/A	N/A	100 mM	25 mM
	NaCl	Sigma-Aldrich	S3014	2.5 M	50 mM
	Glucose	Sigma-Aldrich	G8270	20% w/v	0.4% w/v
	NH_4Cl	Sigma-Aldrich	A9434	1.5 M	15 mM
	$MgSO_4$	Sigma-Aldrich	83266	1 M	1 mM
	$CaCl_2$	Sigma-Aldrich	C3306	1 M	0.01 mM
	Micronutrients Solution	N/A	N/A	10,000 X	1 X
K_2HPO_4/KH_2PO_4 (pH 5.8/7.4)	K_2HPO_4	Sigma-Aldrich	60353	1 M	Mixed until pH 5.8 or 7.4
	KH_2PO_4	Sigma-Aldrich	P8416	1 M	
Micronutrients Solution	Na_2MoO_4	Sigma-Aldrich	737860	100 µM	10 nM
	Na_2SeO_3	Sigma-Aldrich	S5261	100 µM	10 nM
	H_3BO_3	Sigma-Aldrich	B6768	40 µM	4 nM
	$CoCl_2$	Sigma-Aldrich	C8661	3 mM	300 nM
	$CuSO_4$	Sigma-Aldrich	C1297	1 mM	100 nM
	$MnCl_2$	Sigma-Aldrich	529680	8 mM	800 nM
	$ZnSO_4$	Sigma-Aldrich	Z4750	10 µM	1 nM
Green plate media	Tryptone	Appleton Woods	MN649	N/A	8 g/L
	NaCl	Sigma-Aldrich	S3014	N/A	5 g/L
	Bacto-Yeast Extract	Appleton Woods	DM832	N/A	1 g/L
	Bacto-Agar	Appleton Woods	214010	N/A	11 g/L

	Methyl Blue	Sigma-Aldrich	M6900	N/A	0.1 g/L
	Alizarin Yellow	Sigma-Aldrich	206709	N/A	0.6 g/L
	Glucose	Sigma-Aldrich	G8270	20% w/v	0.7% w/v
M9 media	M9 salts (5 X)	Sigma-Aldrich	M6030	5 X	1 X
	Glucose	Sigma-Aldrich	G8270	20% w/v	0.4% w/v
	$MgSO_4$	Sigma-Aldrich	83266	1 M	2 mM
	$CaCl_2$	Sigma-Aldrich	C3306	1 M	0.1 mM
M9 salt (5 X) (Sigma-Aldrich)	KH_2PO_4	Sigma-Aldrich	N/A	N/A	15 g/L
	NaCl	Sigma-Aldrich	N/A	N/A	2.5 g/L
	Na_2HPO_4	Sigma-Aldrich	N/A	N/A	33.9 g/L
	NH_4Cl	Sigma-Aldrich	N/A	N/A	5 g/L
Antibiotics	Ampicillin sodium salt	Melford	A0104	100 mg/ml (H_2O)	100 µg/ml
	Chloramphenicol	Sigma-Aldrich	C0378	20 mg/ml (ethanol)	20 µg/ml
	Gentamicin sulphate	Melford	G0124	40 mg/ml (H_2O)	20 µg/ml
	Kanamycin monophosphate	Melford	K0126	50 mg/ml (H_2O)	50 µg/ml
	Tetracycline	Sigma-Aldrich	87128	20 mg/ml (methanol)	20 µg/ml

Supplements	L-Arabinose	Sigma-Aldrich	A3256	20% w/v (H_2O)	0.2% w/v
	Congo Red	Sigma-Aldrich	C6767	N/A	40 µg/ml
	H_2O_2 (30% w/w)	Merck Millipore	107209	N/A	20% w/w
	Anhydrotetracycline hydrochloride	Merck Millipore	37919	1 mg/ml	500 ng/ml
Reagents	Ethanol	Sigma-Aldrich	E7023	N/A	N/A
	Isopropanol	Sigma-Aldrich	19516	N/A	N/A
	Methanol	Sigma-Aldrich	E7023	N/A	N/A
	DEPC-treated water	Fisher Scientific	10514065	N/A	N/A
	Formaldehyde solution	Fisher Scientific	10160052	N/A	N/A
	Glycerol	Sigma-Aldrich	49767	N/A	N/A
	Phosphate buffered saline (PBS)	Fisher Scientific	11503387	1 X	1 X

2.3 Strain list and general culture conditions

All bacterial strains used in this thesis are listed in Table 2.2. Bacterial overnight cultures were made by inoculating a single colony from agar plates into 5 ml LB in a glass universal bottle and incubating for about 16 hours at 37°C with shaking at 220 rpm. For growth in flasks, 25 µl of overnight culture was added into 25 ml liquid media in a 250 ml sterilized glass Erlenmeyer flask (to make a 1:1,000 dilution), and then grown at 37°C with shaking at 220 rpm. For growth in minimal media (M9, InSPI2 or NonSPI2), 1 ml of LB-grown overnight culture was harvested by centrifugation at 12,000 x g at room temperature for 1 minute. Cell pellets were washed in PBS or minimal media three times and sub-inoculated 1:500 into 25 ml minimal media in a 250 ml sterilized glass Erlenmeyer flask.

To make a frozen stock of bacteria, 900 µl of overnight culture was mixed with 600 µl of autoclaved 50% (v/v) glycerol in a 2 ml cryovial tube (STARLAB) and then stored at -80°C. To prepare isolated bacterial colonies, one speck of frozen stock was streaked onto agar plates and incubated at 30°C or 37°C overnight. The resulting plates were stored in sealed plastic bags at 4°C for a maximum of 2 weeks.

Table 2.2. Strain list

Name	Description	Reference
Salmonella Typhimurium		
Wild type		
4/74	Reference strain for *S.* Typhimurium ST19	Rankin and Taylor (1966)
A130	Reference strain for *S.* Typhimurium ST313 Lineage 1 from Sub-Saharan Africa	Kingsley *et al.*, (2009)
D23580	Reference strain for *S.* Typhimurium ST313 Lineage 2 from Sub-Saharan Africa	Kingsley *et al.*, (2009)
D37712	Reference strain for *S.* Typhimurium ST313 Lineage 2.2 from Sub-Saharan Africa	Msefula *et al.* (2012)
U2	Reference strain for *S.* Typhimurium ST313 from the UK	Ashton *et al.* (2017)
U8	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U15	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U34	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U36	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U37	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U38	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
U76	*S.* Typhimurium ST313 strain from the UK	Ashton *et al.* (2017)
4/74 derivatives		
JH4301	Nal^R pBT1*::aph*	R. Canals *Pers. Comm.*
JH4302	Nal^R pBT1 Δ*cysS*^{pBT1}*::aph*	R. Canals *Pers. Comm.*
SZS001	Nal^R pBT1 pP_{cysS}-*gfp*	This study
SZS002	Nal^R pBT1-Δ*cysS*^{pBT1}*::aph* pBT1-0061^{mut} pP_{cysS}-*gfp*	This study
SZS003	Nal^R pBT1-Δ*cysS*^{pBT1}*::aph* pBT1-0081^{mut} pP_{cysS}-*gfp*	This study
SZS004	Nal^R	This study

	pBT1-Δ*cysS*^{pBT1}*::aph* pP_{cysS}-*gfp*	
D23580 derivatives		
JH3949	ΔΦ (ΔBTP1 ΔBTP5 ΔGifsy-1 ΔGifsy-2 ΔST64B)	Owen *et al.* (2017)
SZS005	ΔΦ [BTP1] lysogen	This study
SZS006	ΔΦ [BTP1 ^{UK1}] lysogen	This study
SZS007	ΔΦ [BTP1 ^{UK3}] lysogen	This study
JH4298	Δ*cysS*^{chr}	Canals *et al.* (2019b)
JH4299	Δ*cysS*^{pBT1}	Canals *et al.* (2019b)
JH4300	ΔpBT1	Canals *et al.* (2019b)
JH4235	SSS18 Km^{S}	R. Canals *Pers. Comm.*
SZS008	SSS18 Km^{S}, *ssaG-gfp*$^{+}$	This study
SNW88	ΔpBT3	N. Wenner *Pers. Comm.*
SZS009	ΔpBT2, ΔpBT3	This study
SZS010	ΔpBT2, ΔpBT3, ΔpSLT-BT (Abbreviation as pBT1^{+})	This study
SZS011	ΔpBT1 ΔpBT2 ΔpBT3 ΔpSLT-BT (Abbreviation as Δplasmid)	This study
JH3794	Km^{R}	Jared *et al.* (2020)
SZS012	pMRE-Tn7-145 (Tn7-mScarlet-I)	This study
SZS013	pMRE-Tn7-152 (Tn7-sGFP2)	This study
SZS014	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::cat*	This study
SZS015	ΔpBT1 ΔpBT2 ΔpBT3 ΔpSLT-BT pP_{cysS}-*gfp*	This study
SZS016	ΔpBT1 ΔpBT2 ΔpBT3 ΔpSLT-BT pP_{cysS}^{mut}-*gfp*	This study
SZS017	ΔpBT2, ΔpBT3, ΔpSLT-BT pP_{cysS}-*gfp*	This study
SZS018	ΔpBT2, ΔpBT3, ΔpSLT-BT pP_{cysS}^{mut}-*gfp*	This study
SZS019	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*pBT1-0081::aph*	This study
SZS020	ΔpBT2, ΔpBT3, ΔpSLT-BT P_{tetA}-*pBT1-0081-aph*	This study
SZS021	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::cat* P_{tetA}-*pBT1-0081-aph*	This study
SZS022	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::cat* Δ*pBT1-0061-traL::aph*	This study
SZS023	ΔpBT2, ΔpBT3, ΔpSLT-BT	This study

	Δ*cysS*^{pBT1}*::cat* Δ*pBT1-0071-traL::aph*	
SZS024	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::cat* Δ*pBT1-0081-traL::aph*	This study
SZS025	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::cat* Δ*traL::aph*	This study
SZS026	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::aph* pP_{cysS}-*gfp*	This study
SZS027	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*pBT1-0081::aph* pP_{cysS}-*gfp*	This study
SZS028	ΔpBT2, ΔpBT3, ΔpSLT-BT P_{tetA}-*pBT1-0081-aph* pP_{cysS}-*gfp*	This study
SZS029	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::*FRT P_{tetA}-*pBT1-0081-aph* pP_{cysS}-*gfp*	This study
SZS030	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::*FRT P_{tetA}-*traL-aph*	This study
SZS031	ΔpBT2, ΔpBT3, ΔpSLT-BT Δ*cysS*^{pBT1}*::*FRT P_{tetA}-*traL-aph* pP_{cysS}-*gfp*	This study
D37712 derivatives		
SZS032	*ssaG-gfp*$^{+}$	This study
JH4232	Km^{R}	N. Wenner *Pers. Comm.*
SZS033	pMRE-Tn7-145 (Tn7-mScarlet-I)	This study
SZS034	pMRE-Tn7-152 (Tn7-sGFP2)	This study
U2 derivatives		
SZS035	[BTP1] lysogen	This study
SZS036	[BTP1^{UK1}] lysogen	This study
SZS037	[BTP1^{UK3}] lysogen	This study
SZS038	[P22] lysogen	This study

2.4 General bacterial methods

2.4.1 Cell counting

Optical density (OD) was used to assess bacterial growth by light scattering. Measurements were taken at a wavelength of 600 nm using a JENWAY 6705 Spectrophotometer. When the OD_{600} value was higher than 1, a 1:2, 1:10 or 1:20 dilution was made before measurement to give a more accurate assessment of cell density.

To measure colony forming unit (CFU), several 10-fold serial dilutions of cultures were made, until a factor of 10^{-8} in LB media and then 100 µl of each dilution was spread on LB agar plate for overnight incubation at 37°C. For each dilution, triplicate plates were prepared. Colony numbers were determined using a colony counter (Stuart) to establish CFU per ml.

2.4.2 Growth curves

To determine bacterial growth curves, cultures were cultured in 96-well plates and scanned using two different machines: the FLUOstar Omega microplate reader (BMG LABTECH) or the Growth Profiler (EnzyScreen, https://www.enzyscreen.com/growth_profiler.htm) as detailed below.

Plates and lids used for FLUOstar Omega were from Greiner (Product code: Plate, 655161; Lid, 656171). Overnight cultures were diluted to OD_{600}=1 and a further

1:100 dilution was inoculated into a 96-well plate to generate a starting OD_{600} of 0.01. The total volume of media in each well for the Omega plate reader was 200 µl and for Growth Profiler was 250 µl. The 96-well plates were incubated at 37°C with shaking at 220 rpm in both Omega plate reader and Growth Profiler for 24 or 48 hours. Measurement was taken every 30 minutes.

Growth Profiler is a machine that can incubate 10 96-well plates with shaking at the same time. The plates and lids used for Growth Profiler were from EnzyScreen (Product code: Plate, CR1496dg; Lid, CR1296c). Growth Profiler machine reads the green pixel value of each well and converts the value into optical density based on a calibration curve. Calibration curves were based on the green values of a series of 10 dilutions of the bacterial culture from OD_{600}=20 to OD_{600}=0.01. For different media, different calibration curves were made. The calibration curves used in this thesis were as follows:

LB:

$$OD_{600}=0.0234 \times (GV - 18.143) + 7.3E^{-7} \times (GV - 18.143)^{3.39} + 1.56E^{-13} \times (GV - 18.143)^{6.58},$$

Minimal media:

$$OD_{600}=0.0234 \times (GV - 17.725)^{1.03} + 7.6E^{-7} \times (GV - 17.725)^{3.31} + 1.56E^{-13} \times (GV - 17.725)^{6.58}.$$

(GV: Green value)

Each growth curve was made by the mean value of six replicates (six wells). Standard Deviation for each group of replicates was calculated and the values were all below 5% for the significance. Error bars were not shown in this thesis. Growth rate for each replicate of a growth curve was calculated using Microsoft Excel. The growth curve was transformed into log scale for Y-axis (OD_{600}), and then the maximum growth rate was determine by comparing the slope between 6 or 8 timepoints of the growth curve. T-test and ANNOVA were used to show significance for each sample.

For some of growth curves in this thesis, the lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 (Hersch *et al.*, 2021). The calculation was performed in GraphPad Prism 8 using nonlinear regression.

2.4.3 Swimming Motility assay

About 25 mL 0.3% LB agar was aliquoted into Petri dishes and left to set and dry at room temperature overnight. Overnight cultures of 3 µL were spotted onto the plates and left at room temperature for 30 minutes to dry before incubation at 37°C for 8 hours. The diameter of migration for triplicates was measured at 4 and 8 hours. The level of bacterial swimming motility was assessed by the average of the migration halo.

2.4.4 RDAR phenotypic assay

RDAR (red, dry and rough) colonial phenotype (Singletary, *et al.,* 2016) was assessed using LB plates without NaCl and supplemented with 40 µg/ml Congo red (Sigma-Aldrich, C6277). Two µl of overnight culture was spotted onto LB-Congo red plates and incubated at room temperature for 3 days. The colonies were photographed using the ImageQuantTM LAS 4000 biomolecular imager (GE Healthcare). A positive RDAR phenotype appeared as a red colony with filamentous texture, whereas the negative phenotype was a smooth and yellow colony.

2.4.5 Microscopy

The EVOS FL cell imaging system (Thermo Fisher) was used in this thesis to visualize cells using transmitted light cube, GFP light cube (470 nm excitation, 525 nm emission) and RFP light cube (530 nm excitation, 593 nm emission) with an EVOS 40X flurorite coverslip-corrected objective (Thermo Fisher, AMEP4699) and an Olympus 100X super-apochromat, coverslip-corrected oil objective (Thermo Fisher, AMEP4733). One µl of cell culture was spotted onto the middle of a glass microscopy slide and covered with a cover slip. When using 100X oil objective, immersion oil (Sigma-Aldrich, 56822) was dropped onto the cover slide before imaging.

2.4.6 Flow cytometry

Bacteria tagged with the GFP or mScarlet gene was incubated in InSPI2 or NonSPI2 media at 37°C with shaking at 220 rpm for approximately 8 hours until OD_{600}=0.3. Cells were harvested from 1 ml of culture and washed with sterile PBS buffer by centrifugation at 12,000 x g for 2 minutes. The pellets were resuspended in 1 ml PBS with 4% formaldehyde (Fisher Scientific, 10160052) for 1 hour to fix the samples. The fixed cells were then washed with PBS for three times to remove formaldehyde and stored at 4°C for flow cytometric analysis (FACS).

The FACS measurements were performed on a BD FACSAria Flow Cytometer. The voltages set for forward angle light scatter (FSC), side scatter (SSC) and fluorescence were 568V, 389V and 464V, respectively. The threshold was set at 200 for FSC, and the fluorescence of 10,000 cells was measured. FITC filter (488 nm, 530/30) was used to measure fluorescence intensity of GFP, while PE Yell-Grn filter (561 nm, 582/15) was used to measure fluorescence intensity of mScarlet. FACS data was further analyzed using FlowJo VX software.

2.4.7 Competition index assay

Two groups of strains (Strain A and B) were used to measure relative fitness by competition index assay using the antibiotic method. Five µl overnight cultures for each group were inoculated into 25 ml LB media in a 250 ml glass Erlenmeyer flask to generate a final 1:1,000 dilution. The 50:50 mixture of the pair of strains was

incubated overnight for approximately 16 hours at 37°C with shaking at 220 rpm. One ml of the resulting mixed culture was serial-diluted to 10^{-8}, and 100 µl from each dilution was spread on two groups of LB plates supplemented with different antibiotics. After overnight incubation at 37°C, the CFU was counted from plates that contained between 10 to 200 colonies. The competitive index (Beuzón and Holden, 2001) was calculated as the ratio between CFU of Strain A and CFU of strain B. The experiments were repeated in triplicate for each pair of strains.

2.4.8 Evolution experiments

For strains that grew poorly on a specific carbon source, experimental evolution was used to select spontaneous mutants capable of utilizing the carbon source. For evolution experiments in liquid media, approximately 1,000 starter cells from an overnight culture were inoculated into 250 µl M9 minimal media with glucose as the sole carbon source, in a 10 wells of Growth Profiler plate. The bacteria were incubated at 37°C with shaking at 220 rpm for 2 days, streaked to single colonies twice on LB plates and then streaked on M9 glucose plates determine whether isolated large colonies appeared or not.

For evolution experiments conducted on agar plates, 10^8 starter cells from overnight cultures were spread on M9 glucose plates and incubated for 2 days at 37°C until large colonies appeared.

The candidate spontaneous mutants were streaked on LB plates for another 2 passages, and tested again by streaking on an M9 plate to confirm their positive phenotypes. Subsequently the positive candidates were sent to MicrobesNG (UK) for whole genome sequencing, according to the company's instruction. SNPs and other mutations were identified by a comparative genomic approach using the genome of the wild type strain (Section 2.8.2).

2.5 Phage-related methods

2.5.1 Phage propagation

To obtain pure lysates of individual phages, and to avoid contamination by native prophages, the phages were propagated on D23580 ΔΦ (strain JH3949, Table 2.2), where all prophages have been previously deleted (Owen *et al.*, 2017). Phage lysates were collected by centrifugation (12,000 x g, 2 min) followed by syringe filtration through a 0.22 μm filter. Phage lysates were stored at 4°C before use.

2.5.2 Phage enumeration and plaquing

The double layer technique was used for phage enumeration and plaque isolation. For spontaneously induced phages, phage numbers were calculated as plaque forming units (PFU) per ml of phage lysate. One mL overnight culture was filter-sterilized by a 0.22 μm syringe filter. The supernatants were serially diluted in LB. Four mL of 0.4% soft LB agar was seeded with 100 μl overnight culture of the indicator strain. After 5 min drying, 10 μl of phage dilutions were dropped onto the

cell lawn. Plates were then inverted and incubated overnight at 37°C after drying.

2.5.3 Phage lysogen isolation

Host strains were incubated in 5 ml LB at 37°C with shaking at 220 rpm until exponential phase. The cultures and target phage were plated using the double layer technique described Section 2.5.2. After overnight growth, the turbid centers from plaques were purified on green indicator plates (Table 2.1; Maloy, 1990). Green plates contain a pH indicator that is originally pale green, and becomes dark green when phage infection causes cell lysis that releases organic acids. Candidate lysogens were identified by PCR (Table 2.3) and then cross streaked against target phage onto green plates to confirm resistance to phage infection.

2.6 Molecular biology techniques

2.6.1 Primers and plasmids

All primers and plasmids used in this thesis are shown in Table 2.3 and Table 2.4 respectively. Primers were ordered from MWG Eurofins (Germany) and were HPSF (High-Purity and Salt-Free) purified. Plasmids DNA was prepared using the ISOLATE II Plasmid Mini Kit (Bioline). Plasmid isolation followed the protocol from Bioline Kit, except that DEPC-treated H_2O (Fisher Scientific) was used instead of elution buffer at the last step. The presence or absence of target genes was determined by Sanger sequencing conducted by GATC Biotech (Germany) using

LIGHTRUN Tubes. A NanoDrop ND-1000 Spectrophotometer (Thermo Fisher) was used to measure all DNA concentrations in this thesis.

Protocol for isolation of plasmids with the Bioline ISOLATE II Plasmid Mini Kit*	
Harvest cells	Prepare 5 ml overnight culture. Centrifuge for 5 minutes at 11,000 x g and discard supernatant.
Lyse cells	Add 500 µl Resuspension Buffer P1 and mix. Add 500 µl Lysis Buffer P2 and invert the tube for 6 to 8 times. Incubate at room temperature for 5 minutes. Add 600 µl Neutralization Buffer P3 and invert the tube for 6 to 8 times.
Clarification	Centrifuge at 11,000 x g for 10 minutes.
Bind DNA	Place a ISOLATE II Plasmid Mini Spin Column onto a collection tube. Pipette 700 µl of sample supernatant onto the column and centrifuge at 11,000 x g for 1 minute and discard flow-through.
Wash membrane	Add 500 µl Wash Buffer PW1 (pre-heated at 50°C), centrifuge at 11,000 x g for 1 minute and discard flow-through. Add 600 µl Wash Buffer PW2 (supplemented with ethanol), centrifuge at 11,000 x g for 1 minute and discard flow-through.
Dry membrane and elute DNA	Centrifuge at 11,000 x g for 2 minutes and discard the residual liquid. Place the column onto a 1.5 ml Eppendrof tube. Add 30 µl DEPC-treated water, incubate at room temperature for 1 minute and centrifuge for 1 minute at 11,000 x g.

*Bioline kit product code: BIO-52057

Table 2.3. Primer list

Name	Sequence (5'→3')	Procedure	Template	Purpose
BTP1_JP_fw	CGTGAACACACCCTTCTCAG	Initial denaturation: 95°C 120 s Denaturation: 95°C 15 s 35X Annealing: 55°C 30 s 35X Extension: 72°C 30 s per kb 35X Final Extension: 72°C 300 s	*Salmonella* genome	To confirm absence of BTP1 prophage
BTP1_JP_rv	ACTCATGGCGCATGGTAAAC			
BTP1UK1_fw	CCTCAGGATGGTACGTAACT			To confirm absence of BTP1^{UK1} prophage
BTP1UK1_rv	GACAACACCGATAATGCGGC			
BTP1UK3_fw	GAGCTGTTCGTCCTGTACAA			To confirm absence of BTP1^{UK3} prophage
BTP1UK3_rv	GCTCGGTGGTTCTCGTGTAC			
P22_F	ACTCAAGATTAGACAAGCCGC			To confirm absence of P22 prophage
P22_R	CTTGGTTCGACTTCGGGATC			
ssaG_ex_fw	GCTCAACCCAGAATCGATGA			To confirm insertion of *gfp*$^{+}$-*aph* region
ssaG_ex_rv	GCGCTGGTAACATCGCATGT			
cysSpBT1_ex_fw	GGATTCGTACACCCGCCTC			To confirm absence of *cysS*^{pBT1} in pBT1
cysSpBT1_ex_fw	TCGGTGGCCGTTCGTCTCG			
cysSChr_ex_fw	GGGGCACATATAGGGGCTTG			To confirm absence of *cysS*^{Chr} in chromosome
cysSChr_ex_rv	GGCAATACCCGGGTGTGAAC			
pBT1-81_ex_fw	CAACTCCGCTTCGACAACT			To confirm absence of *pBT1-0081* in pBT1
pBT1-81_ex_rv	GGCCGATACCATTCCGTAT			
traL_ex_fw	CATTTCCGACTGGTGCATAT			To confirm absence of *traL* in pBT1
traL_ex_rv	GCTGGTAAATGGATGCGCTG			
pBT1-61_ex_fw	CCTACAGCTACCATCGATTC			To confirm absence of *pBT1-0061* to *traL* region in pBT1
pBT1-71_ex_fw	GGCTCCGGTATTTGTGATCA			To confirm absence of *pBT1-0071* to *traL* region in pBT1
Test_pBT2_fw	GGCTGAATGGTGTGAATTTCC			To confirm absence of pBT2 plasmid

Test_pBT2_rv	GAAAGCTACTGCGATTGTTTCC			
Test_pSLT-BT_fw	ACATAATGCAGAATAAAAGATCGCT			To confirm absence of pSLT-BT plasmid
Test_pSLT-BT_rv	AGTGTCCTGCCCAAAGCCTCTG			
glmS_up	CGTTCCGCTGCAACTGCTGG			To confirm the insertion of Tn7-mScarlet/sGFP2 module
glmS_down	ATCCACCTGAATAACCTGGG			
PcysSchr_Fw	ACGTCCCGGGTCGCCAGCGTACCACGGGTA			To amplify promoter region of *cysS* in chromosome
PcysSchr_Rv	TTTTTCTAGACGAAGATTCCGTATAGACGT			
ssaG_fw	GTGGAATCATTGCTAAAATCTGAAGTTATTAGCGACGATGTTCTAGATTTAAGAAGGAGATATAC	Initial denaturation: 98°C 30 s Denaturation: 98°C 10 s 35X Annealing: 53°C 30 s 35X Extension: 72°C 30 s per kb 35X Final Extension: 72°C 300 s	pNAW52 plasmid	To amplify *gfp*⁺-*aph* region from pNAW52 plasmid
ssaG_rv	TTAACGCCCGCAAACATGATTTCCAGCAGCAACCGTCGAAGTGTAGGCTGGAGCTGCTTC			
tetR_pBT1-81_fw	TTCGGCTATCAGTTTCATTTTATATGCACCAGTCGGAAATGTCAGTAATGCATATGAATATCCTCCTTAG		pNAW55 plasmid	To amplify *aph-tetR*-P*tetA* module from template plasmid pNAW55 to construct an inducible *pBT1-0081*
tetR_pBT1-81_rv	GTGACGATTTTCTTACTCATTGCATCCTCTTTACCAGACGTGAACTTCATTTCACTTTTCTCTATCACTG			
tetR_traL_fw	ATTGTGGTATATCATACTTCATCTTACAACATAACTATCAGATTTATATGCATATGAATATCCTCCTTAG			To amplify *aph-tetR*-P*tetA* module from template plasmid pNAW55 to construct an inducible *traL*

tetR_traL_rv	GCTCCTGAGGCTATAATGGTTAAGAACACGATTACATTAAGTGTTTTCATTTCACTTTTCTCTATCACTG			
cysSpBT1_del_fw	GGCTCTATTGATTAAGTAAACGAACTACTATCAGGAATGTCATAAGTGTAGGCTGGAGCTGCTTC	Initial denaturation: 98°C 30 s Denaturation: 98°C 10 s 35X Annealing: 53°C 30 s 35X Extension: 72°C 30 s per kb 35X Final Extension: 72°C 300 s	pKD3 or pKD4 plasmid	For λ Red recombination to delete *cysS*^{pBT1} from the plasmid pBT1
cysSpBT1_del_fw	AGCTAAACAGGCTTACCTATAGCGATAACGGTACTATTCAATTAACATATGAATATCCTCCTTAG			
cysSChr_del_fw	ATCATACGGAGATTATTACCCACACACGTCTATACGGAATCTTCGTGTAGGCTGGAGCTGCTTC			For λ Red recombination to delet *cysS*^{Chr} from the chromosome
cysSChr_del_rv	AAAAAACAGGCCGTTCTATGCAAATAGAGCGGCCTGTAAAAGCGCCATATGAATATCCTCCTTAG			
pBT1-81_del_fw	CGGCTATCAGTTTCATTTTATATGCACCAGTCGGAAATGTCAGTAGTGTAGGCTGGAGCTGCTTC			For λ Red recombination to delete *pBT1-0081* from the plasmid pBT1
pBT1-81_del_rv	ACCCTTCTCTGTGCGGGACACAGAGAAGGTTACTGGTTAAGACGACATATGAATATCCTCCTTAG			
traL_del_fw	TATATCATACTTCATCTTACAACATAACTATCAGATTTATGTGTAGGCTGGAGCTGCTTC			For λ Red recombination to delete *traL* from the plasmid pBT1

traL_del_rv	AAACTAAGAATAAAGGCCGGAAAACGGCCTTTATTTATAACATATGAATATCCTCCTTAG			
pBT1-61_del_fw	AGGTTCTTCTTCAGTGTCGTAAAAAAAAGCAGACATTAGCGTGTAGGCTGGAGCTGCTTC			For λ Red recombination to delete *pBT1-0061* to *traL* from the pBT1
pBT1-71_del_fw	ATGGTTTCCAGTGATTTGATTGAGGCCGATACCATTCCGTGTGTAGGCTGGAGCTGCTTC			For λ Red recombination to delete *pBT1-0071* to *traL* region from the pBT1
tn5ext	GCTTTTAAGGTTTAACGGTTG	Initial denaturation: 95°C 120 s Denaturation 1: 95°C 15 s 6X Annealing 1: 30°C 30 s 6X Extension 1: 72°C 90 s 6X Denaturation 2: 95°C 15 s 30X	*Salmonella* genome	For arbitrary PCR round 1 for Tn*5* insertion site sequencing
ARB1	GGCCACGCGTCGACTAGTACNNNNNNNNNNGATAT			

		Annealing 2: 50°C 30 s 30X Extension 2: 72°C 90 s 30X Final Extension: 72°C 300 s		
tnpRL17-1	AACAAGCCAGGGATGTAACG	Initial denaturation: 95°C 120 s Denaturation: 95°C 15 s 30X Annealing: 50°C 30 s 30X Extension: 72°C 90 s 30X Final Extension: 72°C 300 s	Product from arbitrary PCR 1	For arbitrary PCR round 2 for Tn*5* insertion site sequencing
ARB2	GGCCACGCGTCGACTAGTAC			

Table 2.4. Plasmid list

Plasmid	Description*	Reference
pKD3	Template plasmid carrying FRT-flanked chloramphenicol resistance cassette; Cm^R	Datsenko and Wanner (2000)
pKD4	Template plasmid carrying FRT-flanked kanamycin resistance cassette; Km^R	Datsenko and Wanner (2000)
pKD46	λ Red recombination plasmid, arabinose-inducible	Datsenko and Wanner (2000)
pCP20-Gm	Plasmid carrying FLP recombinase to remove antibiotic resistance cassette from pKD3/pKD4; Gm^R	Doublet *et al.* (2008)
pSIM*5-tet*	λ Red recombination plasmid, temperature-inducible; Tc^R	Koskiniemi *et al.* (2011)
pZEP08	Template plasmid to amplify *gfp+*	Hautefort *et al.* (2003)
pP_{cysS}-*gfp*	pZEP08-*PcysS*$^{chr(WT)}$-*gfp*$^+$-*cat*	This study
pP_{cysS}^{mut}-*gfp*	pZEP08-*PcysS*$^{chr(mut)}$-*gfp*$^+$-*cat*	This study
pMRE-Tn7-145	Tn7-mScarlet-I delivery plasmid; Gm^R, Cm^R	Schlechter *et al.* (2018)
pMRE-Tn7-152	Tn7-sGFP2 delivery plasmid; Km^R, Cm^R	Schlechter *et al.* (2018)
pNAW52	pEMG plasmid carrying *gfp*$^+$-FRT-*aph*-FRT module	N. Wenner *Pers. Comm.*
pNAW55	*aph-tetR*-P_{tetA} delivery plasmid	Owen *et al.* (2021)
pNAW136	Cas9-CRISPR plasmid with an unstable origin of replication *oriR*K2; Km^R	Owen *et al.* (2021).
pNAW139	Cas9-CRISPR plasmid with a protospacer targeting the pBT2 plasmid of D23580; Km^R	Owen *et al.* (2021).
pNAW168	Cas9-CRISPR plasmid with a protospacer targeting the pSLT-BT plasmid of D23580; Km^R	Owen *et al.* (2021)
pNAW169	Cas9-CRISPR plasmid with a protospacer targeting the pBT1 plasmid of D23580; Km^R	Owen *et al.* (2021)
pRL27	Tn*5* transposon delivery plasmid; Km^R	Larsen *et al.* (2002)

*Relevant antibiotic resistances are indicated by ^{R}: Cm, chloramphenicol; Km, kanamycin; Gm, Gentamicin; Tc, Tetracycline

2.6.2 Polymerase chain reaction (PCR)

Polymerase chain reaction was used to amplify DNA fragments. The thermo cycler used in this thesis was Eppendorf Mastercycler Pro S. For PCR to check presence or absence of target gene, MyTaq Red Mix polymerase mastermix (Bioline) was used. To make a final volume of 20 µl for PCR reaction, 10 µl of mastermix together with 1 µl of each primer (final concentration of 0.5 µM each) was added to 8 µl DEPC-treated H_2O (Fisher Scientific). Either 100 ng of template DNA was added, or a bacterial colony was collected with an inoculation loop and used directly.

For DNA cloning and recombineering, Phusion High-Fidelity DNA polymerase (New England Biolabs, M0530) was used to give a high degree of accuracy in DNA replication. In a volume of 50 µl for PCR reaction, 0.5 µl of DNA (100 ng) was mixed with 2.5 µl of each primer (final concentration of 0.2 µM each), 10 µl of 5M Betaine (Sigma-Aldrich, B0300), 20 µl of Phusion HF buffer (New England Biolabs, B0518S), 1.5 µl of DMSO (New England Biolabs, 12611S), 1 µl of dNTP Mix (10mM; BIOLINE, DM-717209), 0.5 µl of Phusion polymerase and 22 µl of DEPC-treated H_2O (Fisher Scientific). All the solutions were stored in -20°C before use.

Initial denaturation time for each PCR reaction depended upon the template DNA that was used. Detailed procedure and template used for each PCR reation is listed in Table 2.3. For example, for extracted or purified DNA, initial denaturation time was 2 minutes at 95°C. Initial denaturation was followed by 35 cycles of denaturation for

15 seconds at 95°C, annealing for 30 seconds at 55°C and extension for approximately 30 seconds per kilobase at 72°C. The cycling was followed by a final extension step at 72°C for 5 minutes. The PCR products were then held at 4°C in the Mastercycler, or stored at -20°C prior to analysis by gel electrophoresis.

2.6.3 Agarose gel electrophesis

Agarose gel electrophesis was used to analysis PCR products. Gels were 1% w/v agarose (Molecular grade, Bioline) within 1X TAE buffer. One litre 50X TAE buffer contained 242 g/L Tris-base, 57.1 ml/L Acetic acid, 100 ml/L sodium EDTA (0.5 µM) and milliQ water. Following mixing, the agarose gel was heated in a microwave oven, and cooled to about 50°C. Midori Green Advanced DNA Stain (NIPPON Genetics EUROPE GmbH) was then added to the gel to a final concentration of 30 µl/L. Gel electrophesis was run in BioRad PowerPac at a voltage of 120 V for approximately 40 minutes to separate DNA. DNA bands were compared with a HyperLadder 1kb (Bioline). Separated DNA fragments were visualized under UV light using a GeneGenius Bio Imaging System or on a blue light illuminator.

To extract DNA fragments from gels for DNA cloning, the band of interest was excised using a scalpel blade on a transilluminator (Clare Chemical, DR-46B). DNA was extracted from the slice of gel, or directly from PCR products, using the ISOLATE II PCR and Gel Kit (Bioline) following the manufacturer's instructions except DEPC-treated H_2O (Fisher Scientific) was used instead of elution buffer.

Protocol for DNA extraction from PCR and gel with the Bioline ISOLATE II PCR and Gel Kit*	
Dissolve gel slice	Place gel slice into a 1.5 ml Eppendorf tube. Add 500 µl Binding Buffer CB. Incubate at 50°C for 10 minutes and vortex until gel is fully dissolved.
Bind DNA	Place ISOLATE II PCR and Gel Column in a collection tube and load sample. Centrifuge at 11,000 x g for 1 minute and discard flow-through.
Wash membrane	Add 700 µl Wash Buffer CW. Centrifuge at 11,000 x g for 1 minute and discard flow-through. Repeat this step one more time.
Dry membrane and elute DNA	Centrifuge at 11,000 x g for 2 minutes and discard the residual liquid. Place the column onto a 1.5 ml Eppendrof tube. Add 20 µl DEPC-treated water, incubate at room temperature for 1 minute and centrifuge for 1 minute at 11,000 x g.

*Bioline kit product code: BIO-52060

2.6.4 Genomic DNA extraction

Wizard Genomic DNA Purification Kit (Promega, A1120) was used to isolate genomic DNA from bacteria. In brief, 1 ml of overnight culture was lysed by 600 µl of nuclei lysis solution (Promega, A7941), then incubated in a heat block (TECHNE Dri-Block DB-2D) at 80°C for 5 minutes. Following cooling down to room temperature, 200 µl of protein precipitation solution (Promega, A795A) was added and the lysates were then incubated on ice for 5 minutes. The lysates were centrifuged at 4°C for 5 minutes and the supernatant was transferred to a new tube. The centrifugation step repeated once and then the supernatant was mixed with 600 µl of isopropanol. The mixture was centrifuged again and all the supernatant was discarded. DNA pellets were washed with 1 ml 70% ethanol and air-dried at room

temperature. 50 µl of DEPC-treated H_2O (Fisher Scientific) was added to the pellets to gently dissolve the DNA by incubation in a heat-block at 65°C for 1 hour with shaking at 950 rpm. The resulting genomic DNA was stored at -20°C.

To obtain high quality genomic DNA for illumina DNA sequencing at MicrobesNG (UK), the ZYMO Research Quick-DNA Miniprep Plus Kit was used, following the manufacturer's instructions:

Genomic DNA extraction by ZYMO Research Quick-DNA Miniprep Plus Kit	
Harvest cells	Prepare 1 ml overnight culture Centrifuge at 12,000 x g for 1 minute and discard flow-through Resuspend cell pellets in 200 µl PBS
Lyse cells	Add 200 µl BioFluid & Cell Buffer (Red) Add 20 µl Proteinase K Mix thoroughly and incubate at 55°C for 10 minutes
Bind DNA	Add 420 µl Genomic Binding Buffer and mix thoroughly Place Zymo-Spin™ IIC-XL Column in a collection tube and load sample Centrifuge at 12,000 x g for 1 minute and discard flow-through
Pre-Wash	Add 400 µl DNA Pre-Wash Buffer to the column in a new collection tube Centrifuge at 12,000 x g for 1 minute and discard flow-through
Wash membrane	Add 900 µl DNA Wash Buffer Centrifuge at 12,000 x g for 1 minute and discard flow-through Repeat this step one more time
Dry membrane and elute DNA	Centrifuge at 12,000 x g for 2 minutes and discard the residual liquid Place the column onto a 1.5 ml Eppendorf tube Add 30 µl DEPC-treated water, incubate at room temperature for 1 minute and centrifuge for 1 minute at 12,000 x g

*ZYMO kit product code: D4069

2.6.5 Plasmid transformation by electroporation

Preparation of electro-competent cells for electroporation followed the method from Edwards *et al.* (1999) to increase the efficiency of transformation. 250 µl of bacterial overnight culture were inoculated into 25 ml LB without salt, supplemented with appropriate antibiotics, and incubated at 30°C with shaking at 220 rpm until OD_{600} was between 0.4 and 0.5. Cells were then incubated on ice for 10 minutes, and then concentrated by centrifugation at 7,000 x g for 15 minutes. Cell pellets were washed three times in ice-cold distilled water by centrifugation, and then resuspended in 200 µl of ice-cold 10% (v/v) glycerol solution for storage at -80°C before use.

For electroporation reactions, 50 µl of electro-competent cells were mixed with approximately 500 ng of PCR products or plasmids and incubated on ice for 5 minutes prior to transfer to a pre-chilled sterile 2mm Electroporation Cuvette (Geneflow, E6-0060). Electroporation reactions were performed in the Bio-Rad MicroPulser (Parameters: Manual, 2.5 kV). Cells were resuspended in 1 ml of LB media at room temperature and transferred to a 1.5 ml Eppendorf tube to incubate at 37°C for 1 hour with shaking at 220 rpm for recovery.

To transform temperature-sensitive plasmids such as pSIM*5-tet* used for Lambda Red recombination, the recovery step was performed at 30°C for 2 hours. After recovery, cells were centrifuged at 12,000 x g for 2 minutes and resupended in 100 µl LB prior to spreading on an LB plate with appropriate antibiotics for an overnight

incubation at 37°C to achieve isolated colonies. For each experiment, a cuvette of cells without DNA were electroplated and plated on the antibiotic agar plate as a negative control.

2.6.6 Construction of the pZEP08-P_{cysS}^{chr}-gfp^{+} transcriptional reporter fusion

Plasmid pZEP08 was used to construct a plasmid-encoded transcriptional fusion. The plasmid was extracted from *E. coli* DH5α pZEP08 using the Bioline ISOLATE II Plasmid Mini Kit described in Section 2.6.1. The plasmids were digested by restriction enzyme SmaI (New England Biolabs, R0141) and XbaI (New England Biolabs, R0145). For a total 20 µl reaction, 2 µg pZEP08 plasmid was mixed with 2 µl Tango buffer 10X (Thermo Fisher, BG5), 1 µl SmaI and 1 µl XbaI. The reaction was first incubated at 30°C for SmaI digestion for 1.5 hours and then at at 37°C for XbaI digestion for 1.5 hours. The product was analyzed by electrophoresis on 1% agarose gel following protocol from Section 2.6.3. Two bands (A: ~6.7 kB; B: ~1 kB) were visible under blue light and the larger band A (~6.7 kB) was excised from the gel and purified as described in Section 2.6.3 to elute the DNA into 20 µl DEPC-treated water (Fisher Sicentific). The concentration of DNA was measured using the NanoDrop instrument.

To amplify the $cysS^{chr}$ promoter region from the *Salmonella* chromosome, primers PcysSchr_Fw and PcysSchr_Rv (Table 2.3) were used for a PCR with Phusion High-Fidelity DNA polymerase (New England Biolabs). The PCR products was

analyzed by electrophoresis on 1% agarose and then the DNA fragment of interest was cut and purified following protocol described in Section 2.6.3 to elute in 16 µl DEPC-treated water (Fisher Scientific). The PCR product was further mixed with 2 µl Tango buffer 10X (Thermo Fisher, BG5), 1 µl SmaI and 1 µl XbaI for a totally 20 µl for digestion (30°C for SmaI digestion for 1.5 hours and then at at 37°C for XbaI digestion for 1.5 hours). The digested product was purified using Bioline ISOLATE II PCR and Gel Kit and the DNA concentration was measured using the NanoDrop instrument.

For plasmid ligation, approximately 100 ng of digested pZEP08 plasmid and about 3 fold molar excess of digested PCR fragments was mixed together with 2 µl DNA ligase buffer 10X (New England Biolabs, B0202) and 1.5 µl T4 DNA Ligase (New England Biolabs, M0202) to make a final volume of 20 µl. The mixture was incubated at room temperature overnight for the ligation reaction. The ligation products were purified using Bioline ISOLATE II PCR and Gel Kit and transformed to *E. coli* Top10 by electroporation described in Section 2.6.5. The transformation mix was spread on LB Cm plate for overnight incubation at 37°C. Colonies were checked under blue light for expression of the *gfp*$^+$ gene.

To measure the fluorescence intensity of GFP$^+$, cells were incubated in 25 mL LB or M9 glucose media in a 250 ml sterilized glass Erlenmeyer flask to OD_{600} = 2 at 37°C with shaking at 220 rpm. 200 µl of cell culture was then transfer into a Greiner

96-well plate and the fluorescence intensity was read by FLUOstar Omega microplate reader (BMG LABTECH; wavelength: excitation: 485 nm, emission: 520 nm). The results were normalized as below:

Relative fluorescence intensity = (Read of culture – Read of media) / OD_{600}

The measurement was done in triplicate for each experiment.

2.6.7 Phage transduction

The high efficient transducing phage P22 HT 105/1 *int-201* (Schmieger, 1972) was used to perform phage transduction in this thesis. To prepare P22 phage lysates, the donor strains were incubated in 5 ml LB at 37°C with shaking at 220 rpm for 1.5 hours and then 5 µl of P22 HT 105/1 *int-201* was added to the donor culture for a further 4-hour incubation. One ml of the donor strains was centrifuged (12,000 x g, 2 min) and the supernatant was filtered through a 0.22 µm filter (STARLAB) to yield the phage lysate. Phage lysates were stored at 4°C before use.

For transduction, the recipient strains were grown in LB with shaking at 220 rpm at 37°C overnight. 200 µl of the recipient overnight culture was mixed with 2.5 µl of phage lysate from the donor strain and incubated for 1 hour at 37°C with shaking at 220 rpm. For D23580, which is highly resistant to P22 infection (Kintz *et al.*, 2015), the volume of phage lysate added was increased from 2.5 µl to 100 µl. After 1 hour incubation, the recipient culture were centrifuged into 100 µl and spread on LB agar with Kanamycin to select transductants. After one more passage on LB Km plate,

transductants were streaked on green indicator plates (Maloy, 1990) to select phage-free colonies for further use.

2.7 Mutant construction

2.7.1 Construction of mutants by Lambda Red recombination

The Lambda Red recombination plasmids used in this thesis were pSIM*5-tet* for the D23580 background (Koskiniemi *et al.*, 2011) and pKD46 for other strains (Datsenko and Wanner, 2000). To replace a gene by an antibiotic resistance construct, primers were designed with flanking ends corresponding to the precise target position and used to amplify the Chloramphenicol (*cat*) or Kanamycin (*aph*) resistance gene flanked by the FRT sites from plasmid pKD3 or pKD4 respectively. To insert DNA fragments of interest (*e.g.* a reporter gene) with an antibiotic resistance gene, primers were designed to amplify the sequence of interest together with the antibiotic resistance gene flanked by the FRT sites from appropriate plasmid given in Table 2.4. PCR was performed using Phusion polymerase described in Section 2.6.2. The PCR products were separated by gel electrophoresis and extracted from argarose gel to store at -20°C for further use.

Where pKD46 was used, electro-competent cells were prepared following the methods described in Section 2.6.5, but the LB without salt used for cell incubation was supplemented with 10 mM L-arabinose to induce recombinase expression. To prepare electro-competent cells for transformation using pSIM*5-tet*, after an

incubation at 30°C until OD_{600}=0.4, the cells were transferred to a waterbath at 42°C for 15 minutes to induce recombinase expression. Electroporation was performed for electro-competent cells to recombine PCR fragments as described (Section 2.6.5). Recombinants were selected on LB Km plates and the insertion of target genes was confirmed by PCR and gel electrophoresis using external primers. Each genetic construct was subsequently transduced to a clean background using phage P22 HT 105/1 *int-201* (Schmieger, 1972), as described in Section 2.6.7.

To reduce the potential polar effects upon downstream genes, the Kanamycin resistance locus was removed using plasmid pCP20. The pCP20 plasmid encodes a temperature-sensitive FLP recombinase that recognizes the FRT sites. As the Kanamycin resistance gene is flanked by two FRT sites, expression of the FLP recombinase results in replacement of the Kanamycin gene by an 82-85 nt scar (Datsenko and Wanner, 2000). The Lambda Red recombinants were transformed with 500 ng pCP20 by electroporation and selected on LB plates with appropriate antibiotics. Colonies were restreaked for another passage on LB plates at 42°C overnight to cure pCP20 plasmid. The removal of the antibiotic gene was confirmed by PCR using external primers.

2.7.2 Transposon mutagenesis

A transposon is a DNA sequence that is able to translocate within a bacterial genome. The suicide plasmid pRL27 which carries the Tn*5* transposon with a

Kanamycin resistance gene (Larsen *et al.*, 2002) was used to generate a Tn*5* transposon library. Electro-competent cells were prepared as described in Section 2.6.5, except growing at 45°C instead of 37°C to maximize competence (N. Wenner *Pers. Comm.*).

A preliminary test was performed by transforming 3 mg pRL27 plasmid into 50 µl competent cells, streaking on LB Km plates and counting CFU after overnight incubation. The CFU number was used to predict the number of transformants that would be generated by a single transposon mutagenesis. For each transposon mutagenesis experiment, 3 mg pRL27 plasmid was transformed into 50 µl competent cells to obtain totally about 250,000 transformants. The transformants were spread on 10 M9 glucose plates supplemented with Kanamycin. Each plate contained 25,000 Tn*5* transformants. After 1 or 2 days incubation at 37°C, large colonies were streaked for another passage on an M9 glucose plate and further transduced into a clean genetic background, as described in Section 2.6.7.

Arbitrary PCR method (O'Toole and Kolter, 1998) was performed to map Tn*5* transposon insertion sites. Genomic DNA of the candidates was extracted as described in Section 2.6.4, and then PCR amplified using a forward primer (NW_319) matching the Tn5 transposon sequence, and a reverse primer (NW_320) which has a random sequence downstream of a specific tag. The PCR products were purified and amplified for the second time using a forward primer (NW_318)

that matches the Tn*5* sequence downstream of NW_319, and a reverse primer (NW_321) that matches the specific tag of NW_320. The final products were purified and sent for Sanger sequencing at GATC Biotech (Germany). Tn*5* insertion sites were identified by comparing the resulting sequences with the wild type D23580 genome (Accession number: FN424405.1) using BLASTn.

2.7.3 Plasmid deletion by Cas9-CRISPR plasmids

To obtain D23580 which only carried pBT1 and no other plasmids (SZS010, Table 2.2) or D23580 which lacked all plasmids (SZS011), the strain D23580 ΔpBT3 (SNW88) was used as a background. The Cas9-CRISPR-based method was used to cure these plasmids (Lauritsen *et al.*, 2017; Owen *et al.*, 2021). The Cas9 plasmids used in this thesis were based on the plasmid pNAW136 (Owen *et al.*, 2021), which has a Cas9-CRISPR module (Jiang *et al.*, 2013), an unstable origin of replication *ori*RK2 and the *aph* Kanamycin resistance gene (Owen *et al.*, 2021). Due to the unstable origin of replication, the plasmid will be lost from bacteria after several passages on LB plate. Three 30 bp anti-plasmid protospacers targeting pBT1, pBT2 and pSLT-BT, respectively, were ligated into pNAW136 with T4 DNA ligase to generate plasmids pNAW169 (anti-pBT1), pNAW168 (anti-pSLT-BT) and pNAW139 (anti-pBT2)

To cure each plasmid in D23580, 250 ng of the relevant Cas9-plasmid was transformed into 50 µl competent cells following the protocol from Section 2.6.5. The

transformants were streaked on an LB Km plate for overnight growth at 37°C. After another passage on LB Km plate, the transformant was tested by PCR to confirm the loss of Cas9-CRISPR plasmid. Primers used for confirmatory PCR were: pBT1: pBT1-81_ex_fw x pBT1-81_ex_fw; pBT2: Test_pBT2_fw x Test_pBT2_rv; pSLT-BT: Test_pSLT-BT_fw x Test_pSLT-BT_rv. Confirmed transformants were streaked once more on LB plates and then LB Km plates were used to ensure that colonies were kanamycin sensitive and to double-check the loss of the Cas9 plasmid.

2.8 Bioinfomatic analysis

2.8.1 Genome sequence list

The three reference strains 4/74, D23580 and D37712 used in this thesis were genome sequenced, and the annotated sequences are available online. Details of accession number and database of each strain are given below.

Name	Accession Number	Serovar	Database
4/74	CP002487.1	*S.* Typhimurium	GenBank
D23580	FN424405.1	*S.* Typhimurium	GenBank
D37712	CP060165.1	*S.* Typhimurium	GenBank

The sources of all the other genomic data used in this thesis are listed in Appendix I.

2.8.2 Sequence analysis

Online BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) was used for nucleotide sequence (BLASTn) and protein sequence (BLASTp) comparisons using the standard database. Sequence identity was calculated as follows: Query coverage x Query identity. Genomic comparison was visualised with the Artemis Comparison

Tool (ACT; Carver *et al.*, 2008).

The genomes of mutant isolates from evolution experiments (Section 2.4.8) were sequenced by MicrobesNG (UK). SNIPPY v4.4.0 was used to analyze the mutations and SNPs of evolved isolates (https://github.com/tseemann/snippy) by Yan Li from the Hinton Lab. SNIPPY used Freebayes (Garrion and Marth, 2012) as the variant caller. The default parameter of minimal coverage was 10, and minimal fraction was 0. For the assemblies from GenBank, SNIPPY used contigs as the input. Recombinant regions were detected and removed from the alignment using Gubbins v2.4.1 (Croucher *et al.*, 2014).

In Chapter 3, Yan Li constructed the phylogenetic tree with RAxML-NG v0.9.0 (Kozlov *et al.*, 2019) and included 293 genomes of *S.* Typhimurium with the substitution model GRT + G (Section 3.3.2). *S.* Typhi CT18 (Parkhill *et al.*, 2001) was used as the outgroup to root the tree. The phylogenetic tree was visualized on iTol (Letunic and Bork, 2007; https://itol.embl.de/).

In Chapter 6, ShinyCircos was used to build a Circos plot for sequence alignment (Yu *et al.*, 2017; https://github.com/YaoLab-Bioinfo/shinyCircos) by Yan Li. The sequences used and their accession numbers are listed in Appendix IV. Yan Li also built a maximum likelihood tree with FastTree (Price *et al.*, 2010; http://www.microbesonline.org/fasttree/) using a GTR + CAT model. The tree was

mid-point rooted.

2.8.3 Transcriptomic analysis

RNA-seq based transcriptomic data for ST19 strain 4/74 and African ST313 strain D23580 and D37712 were collected previously (B. Kumwenda, *Pers. Comm.*; Canals *et al.*, 2019a; Hammarlöf *et al.*, 2018). *S.* Typhimurium RNA-seq data were generated from 16 infection-relevant conditions (Kröger *et al.*, 2013) and from intra-macrophage replication in murine RAW264.7 macrophages (Canals *et al.*, 2019b). Details of all the growth conditions are given in Table 2.5. The transcriptomic data for 4/74 and African ST313 strain D23580 are available online at SalcomD23580 (http://bioinf.gen.tcd.ie/cgi-bin/salcom_v2.pl?_HL). A global transposon insertion profiling approach was used to identify essential genes in the genome of D23580 (Canals, *et al.*, 2019a). The data is available online and visualized in a Dalliance genome browser (Down *et al.*, 2011; https://hactar.shef.ac.uk/D23580/).

Table 2.5. RNA-seq growth conditions used in *S.* Typhimurium D23580 (Canals *et al.*, 2019b)

Growth Condition	Description*
Early exponential phase (EEP)	Growth in Lennox broth to OD_{600}=0.1
Mid exponential phase (MEP)	Growth in Lennox broth to OD_{600}=0.3
Late exponential phase (LEP)	Growth in Lennox broth to OD_{600}=1.0
Early stationary phase (ESP)	Growth in Lennox broth to OD_{600}=2.0
Late stationary phase (LSP)	Growth in Lennox broth to OD_{600}=2.0 + 6 h
Low temperature 25°C	Growth in Lennox broth to OD_{600}=0.3 at 25°C
NaCl shock	Growth in Lennox broth to OD_{600}=0.3, then addition of NaCl to a final conc. of 0.3 M for 10 min
Bile shock	Growth in Lennox broth to OD_{600}=0.3, then addition of bile (Sigma-Aldrich, Cat. S9875) to a final concentration of 3% for 10 min
Low Fe^{2+} shock	Growth in Lennox broth to OD_{600}=0.3, then addition of 2,2'-dipyridyl (Sigma-Aldrich, Cat. D21630) to a final concentration of 0.2 mM for 10 min
Anaerobic shock	Growth in Lennox broth to OD_{600}=0.3 (50 ml), then filled into 50 ml closed Falcon tube and incubated without agitation for 30 min at 37°C
Anaerobic growth (NoO2)	Static growth in Lennox broth to OD_{600}=0.3 in a completely filled and closed 50 ml Falcon tube
Oxygen shock	Static growth in Lennox broth to OD_{600}=0.3 in a completely filled and closed 50 ml Falcon tube; then 15 min aerobic growth (baffled flask, 250 rpm)
NonSPI2	Growth in PCN medium to OD_{600}=0.3 (pH 7.4, 25 mM Pi)
InSPI2	Growth in PCN medium to OD_{600}=0.3 (pH 5.8, 0.4 mM Pi)
Peroxide shock (InSPI2)	Growth in PCN to OD_{600}=0.3, then addition of H_2O_2 (Merck, Cat. 822287) to final conc. of 1 mM H_2O_2 for 12 min
Nitric oxide shock (InSPI2)	Growth in PCN medium to OD_{600}=0.3 (pH 5.8, 0.4 mM Pi): then addition of 250 µM Spermine NONOate (Sigma-Aldrich, Cat. S150) for 20 min

*Media details taken from Kröger *et al.*, (2013)

2.8.4 Prophage annotation

Prophages in the *S.* Typhimurium ST313 isolates from UK were identified by comparison against the genome of the representative African ST313 strain D23580 (Accession number: FN424405) using BLASTn (The search was performed in 2019).

Prophage locations were confirmed by uploading the bacterial genome sequence to the online phage search tool PHAST (Zhou *et al.*, 2011). Genomic comparison between D23580 and UK-ST313 involved the Artemis Comparison Tool (ACT; Carver *et al.*, 2008) and sequence identity was assessed by BLASTn. For annotation the prophage ORFs of UK-ST313 were characterized by uploading the prophage sequence to RAST server (Aziz *et al.*, 2008). The RAST server provides automated gene calling and functional annotation for bacterial and archaeal genome (Aziz *et al.*, 2008). Details of the annotations are available in Appendix II. Genomes of the BTP1-like prophages were compared to BTP1 using BRIG (Alikhan *et al.*, 2011). Prophage-encoded protein sequences were visualized and compared in Clustal Omega (Sievers and Higgins, 2014).

2.9 Statistical analysis

Statistical analysis was done using GraphPad Prism 8 and Microsoft Excel. One-way ANOVA and t-tests were used to determine significant differences between datasets.

Chapter 3.
Identification of UK and African variants of the BTP1 and BTP5 prophages

3.1 Introduction

To date, three lineages of *S.* Typhimurium ST313 have been found in sub-Saharan Africa: Lineage 1 (L1), Lineage 2 (L2) and Lineage 3 (L3) (Kingsley et al., 2009; Pulford *et al.*, 2021). In addition, the azithromycin-resistant sub-lineage 2.1 (L2.1) has been identified (Van Puyvelde *et al.*, 2019) in the Democratic Republic of Congo.

Antimicrobial resistance was associated with the evolution of ST313 in Africa (Tsai and Coombes, 2021). It appears that clonal replacement of ST313 L1 by L2 occurred due to the acquisition of chloramphenicol resistance by L2 (Okoro *et al.*, 2012). ST313 L3 is a phylogenetic intermediate between L1 and L2 with pseudogenes in important virulence gene regions (Pulford *et al.*, 2021). The majority of the African ST313 isolates that have been identified by genome sequencing belonged to L2, and so L2 is the focus of this project. *S.* Typhimurium ST313 isolate D23580 has been selected as the representative strain of Lineage 2 (Canals *et al.*, 2019a; Canals *et al.*, 2019b; Kingsley *et al.*, 2009) and has been studied throughout this thesis.

Previously, a comparative genomic analysis of strains D23580 and LT2 (the *S.* Typhimurium ST19 reference strain) identified two novel prophages BTP1 and BTP5 (named BT for Blantyre, Malawi, their location of origin and P1 or P5 for prophage 1 or 5) as part of the accessory genome (Kingsley *et al.*, 2009; Owen *et al.*, 2017). BTP1 is a P22-like prophage that binds to the same receptor as P22, namely bacterial lipopolysaccharide (Kintz *et al.*, 2015). BTP1 encodes GtrC, an

acetyltransferase that modifies the *S.* Typhimurium O-antigen to prevent superinfection by BTP1 (Kintz *et al.*, 2015; Owen *et al.*, 2017). Specifically, the GtrC enzyme acetylates a rhamnose residue on the *S.* Typhimurium O-antigen of ST313 (Kintz *et al.*, 2015). Another BTP1-encoded protein, BstA, is responsible for a phage-defense system that has a unique self-immunity system that differentiates between closely-related viruses (Owen *et al.*, 2021). In *S.* Typhimurium ST313 isolate D23580, the BTP1 prophage is spontaneously induced at a remarkably high level (Owen *et al.*, 2017).

In Gram-negative bacteria, induction of Lambdoid phages is normally activated by the SOS response to cellular stress. The SOS response activates the RecA protein to stimulate degradation of the CI repressor. The mechanism behind the high level of spontaneous induction of BTP1 is not understood (Owen *et al.*, 2017). In contrast, prophage BTP5 is a novel type of prophage that has never been found to form plaques on D23580, and its biological function is unknown (Owen *et al.*, 2017).

Over the five years, ST313 isolates have been identified in Brazil and the United Kingdom; the majority of isolates were associated with gastroenteritis and not bloodstream infection, and were antibiotic-susceptible (Almeida *et al.*, 2017; Ashton *et al.*, 2017; Seribelli *et al.*, 2020). Recent phylogenetic analysis showed that ST313 isolates from the UK (hereafter referred to as UK-ST313) are more closely related to ST313 L3 than L2 (Pulford *et al.*, 2021). Following comparison between the

genomes of UK-ST313 and D23580, a key difference in the accessory genome was identified, namely that the BTP1 and BTP5 prophages were absent from UK-ST313 (Figure 3.1; Ashton *et al.*, 2017). A careful assessment of Figure 3.1 revealed that some UK-ST313 isolates carry prophages that share mosaic homology to BTP1 and BTP5. The hypothesis was that the differences in prophage complement provides an important distinction between invasive African *S.* Typhimurium ST313 lineages and UK-ST313.

Another idea that promoted the research strategy was the extensive prophage diversity that has previously been reported in genomes of a range of bacteria (Canchaya *et al.*, 2003), including *Salmonella* isolates (Kropinski *et al.*, 2007). Genomic investigations of the *Salmonella* genus have identified hundreds of genetically-distinct prophages in recent years, contributing an important level of diversity to the accessory genome (Bobay *et al.*, 2013).

In this chapter, a bioinformatic approach was used to study the analogous BTP1 and BTP5 regions of the 64 UK-ST313 isolates from Public Health England (details of the UK-ST313 isolates are given in Appendix III). The phenotypic differences between UK-ST313 and African L2 were then investigated, with a focus on phage biology.

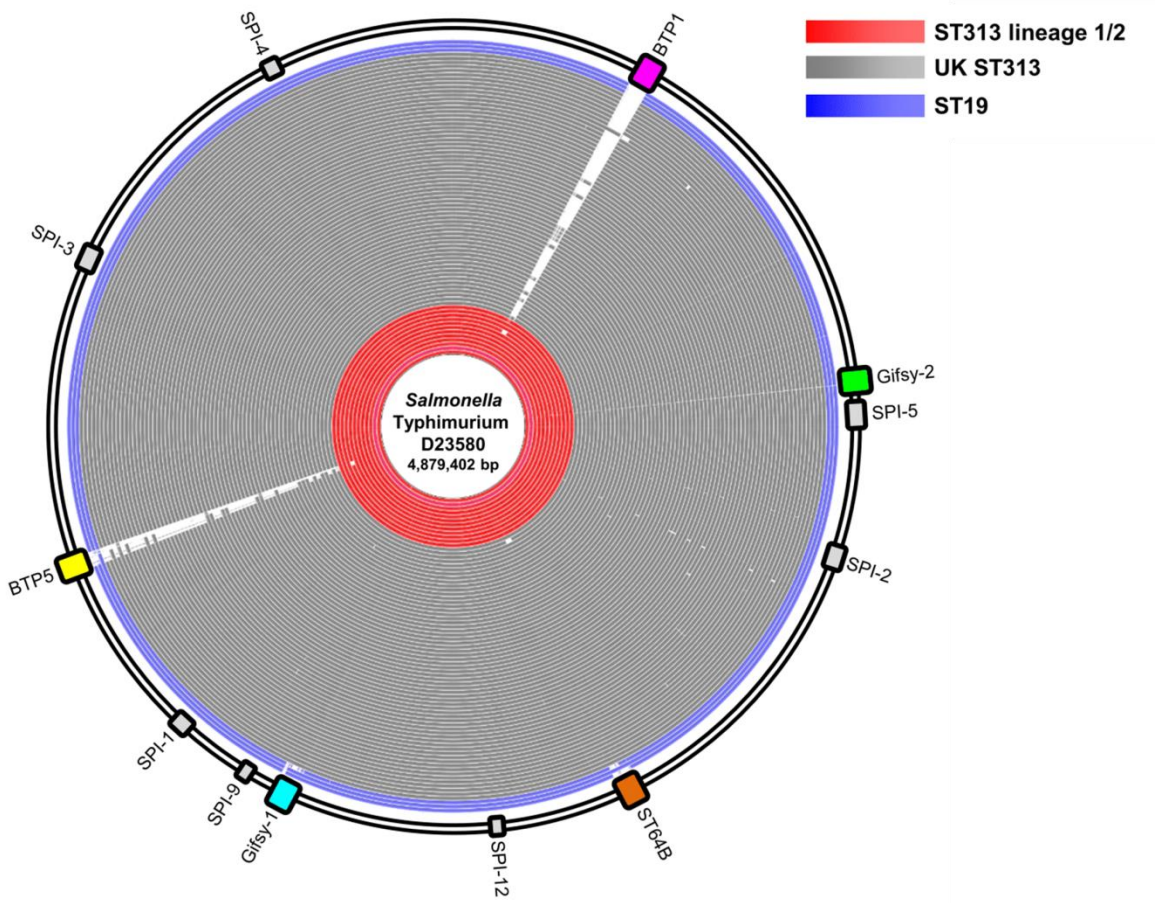


Figure 3.1. UK-ST313 isolates lack BTP1 and BTP5 but contain some homologous genes. The image was generated from 103 genomes using the BLAST Ring Image Generator (BRIG), following comparison with *S.* Typhimurium D23580. The similarities of 76 UK-ST313 isolates (grey), 24 African ST313 isolates (red) and 3 reference ST19 strains (4/74, LT2 and DT104) (blue) versus D23580 (innermost ring) are shown. Gene presence or absence is indicated in red for ST313 L1/L2, and in grey for UK-ST313. The outermost BLAST ring shows the position of D23580 prophages (coloured blocks) and pathogenicity islands (grey blocks). The figure was generated by Siân Owen and is reproduced from Owen *et al.*, 2017.

3.1.1 Acknowledgement of the bioinformatics assistance that underpinned aspects of Chapter 3

I acknowledge my colleagues Siân Owen, Yan Li and Caisey Pulford for their assistance with some of the bioinformatic analyses presented in this chapter. I am grateful to Siân Owen and Yan Li for constructing phylogenetic trees (Figure 3.2 and 3.6) and to Caisey Pulford for comparing the UK BTP1-like prophage with prophages carried by ST313 L3 (Figure 3.5).

3.2 Annotation of Prophages of ST313 isolates from UK

The published genome alignment between *S.* Typhimurium ST313 D23580 and UK-ST313 isolates suggested that UK-ST313 isolates carry prophages with mosaic homology to BTP1 and BTP5 (Ashton *et al.*, 2017). To search for the presence of BTP1-like and BTP5-like prophages in UK-ST313, draft genome assemblies of a total of 64 UK-ST313 isolates were compared with the genome of D23580 using BLASTn (Section 2.8.2). It was noted that all the draft genome assemblies of UK-ST313 consist of a number of contigs. To identify precise prophage locations, the phage *attR* and *attL* sites (the two terminal sites of a prophage; Hoess and Landy, 1978) were identified using PHAST (Zhou *et al.*, 2011). The sequence between the two attachment sites was extracted and considered to be the prophage sequence. The analysis reported below was performed in 2019. Since then, the principle of searching for prophages located between *attR* and *attL* sites on bacterial genomes has become widely adopted; the Prophage Tracer software tool uses a similar

approach, and was published in 2021 (Tang *et al.*, 2021). Using the criterion of >50% nucleotide similarity with BTP1 or BTP5, the prophages I found were classified as BTP1-like or BTP5-like.

In most cases, individual prophages were located in a single contig of the genome. Overall, the 64 UK-ST313 isolates carried seven genetically-distinct prophages. These included 4 BTP1-like prophages (denoted $BTP1^{UK1}$, $BTP1^{UK2}$, $BTP1^{UK3}$ and $BTP1^{UK4}$) and 3 BTP5-like prophages (denoted $BTP5^{UK1}$, $BTP5^{UK2}$ and $BTP5^{UK3}$). BTP1-like prophages were carried by 13 of the 64 isolates, while BTP5-like prophages were found in 21 isolates. Three UK-ST313 isolates (U28, U31 and U52) carried both BTP1-like and BTP5-like prophages. All the BTP1-like and BTP5-like prophages were analysed by RAST (Aziz *et al.*, 2008) to characterize the prophage-encoded open reading frames (ORFs) and the detailed annotation for each prophage is described in the following sections. The gene complement and biology of the BTP1-like phages is discussed in Section 3.3. The gene complement and biology of the BTP5-like prophages is discussed in Section 3.4.

The phylogenetic tree of the UK-ST313 isolates is shown in context of the prophage complement in Figure 3.2. The fact that the BTP1-like prophages and the BTP5-like prophages are only carried by a subset of UK-ST313 is consistent with the significant level of genetic diversity of these isolates that was reported previously (Ashton *et al.*, 2017).

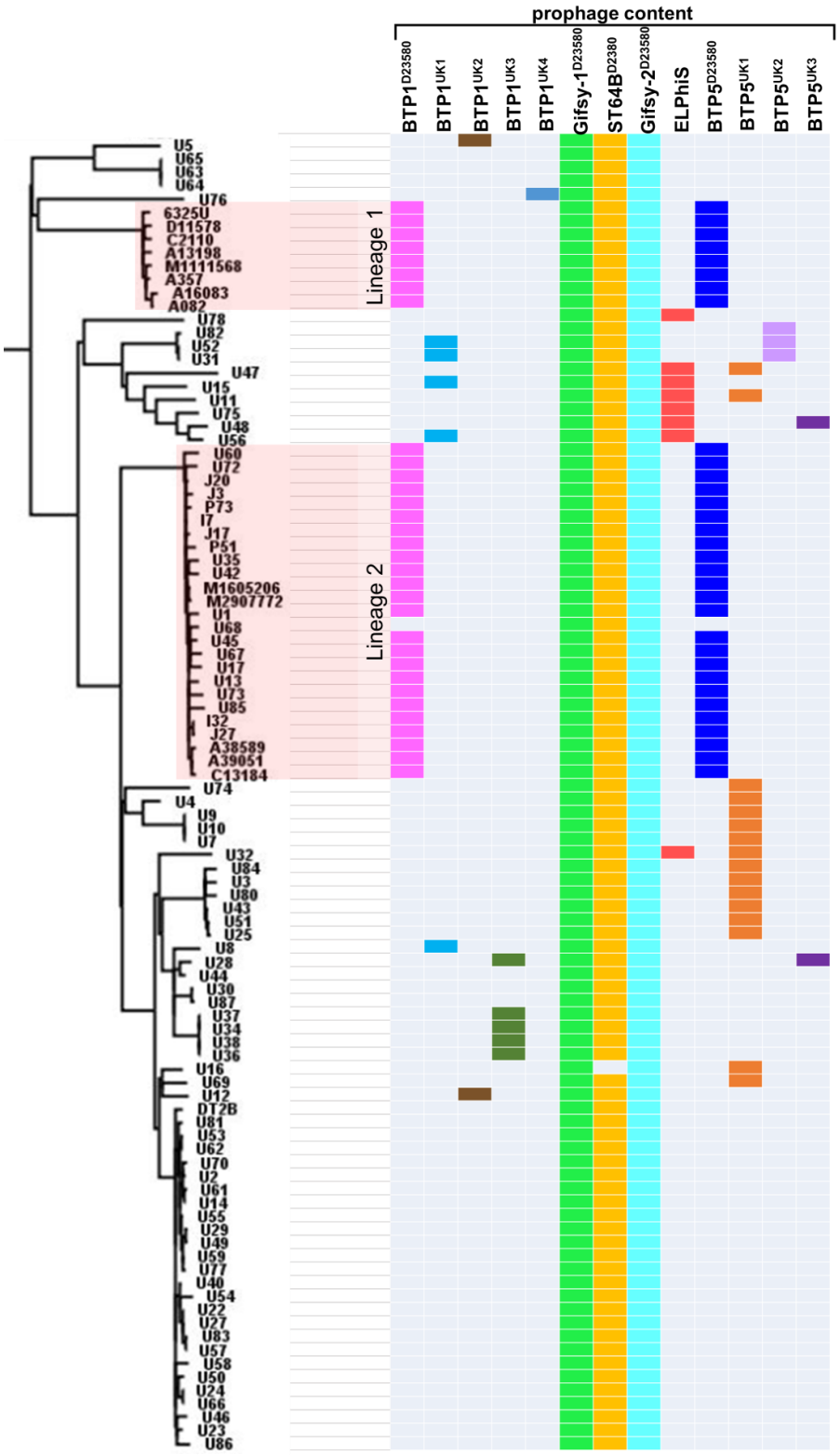


Figure 3.2. The phylogenetic tree of all *S.* Typhimurium ST313 isolated in the UK by Public Health England between 2012 and 2014, annotated to show prophage content. All the isolates were from Owen *et al.*, 2017. Isolates with pink background are ST313 L1 or L2, and isolates with white background are UK-ST313. The heatmap was generated using Microsoft Excel. The pale grey colour means absence of the prophage, while other colours indicate the prescence of a particular prophage.

As well as identifying the novel BTP1 and BTP5 homologs in UK-ST313, the genome comparison showed that all the 64 UK-ST313 isolates carry the prophages Gifsy-1 and Gifsy-2. Gifsy prophages are present in the majority of *S.* Typhimurium strains (Lemire *et al.*, 2008). A prophage is defined as being defective if no viable phage particles are produced either spontaneously or after induction with a DNA damaging agent (Bobay *et al.*, 2014). It has been reported that both Gifsy-1 and Gifsy-2 are defective in D23580 and the molecular basis of the defective prophages has been determined (Figure 3.3; Owen *et al.*, 2017).

In the Gisfy-2 of D23580, the gene that encodes the O replication protein is a pseudogene, with a 71 bp deletion in the centre of the gene (Figure 3.3A; Owen *et al.*, 2017). The O protein is required for phage replication during lytic cycle (Taylor and Wegrzn, 1995). By using ACT, it was identified that most UK-ST313 isolates shared the same deletion as D23580 in centre of the *o* gene, suggesting that UK-ST313-associated Gifsy-2 prophages are also likely to be inactive. In UK isolate U56, an even larger deletion in the centre of the *o* gene in Gifsy-2 was identified, which would also be likely to inactivate the prophage.

In D23580 Gifsy-1, a single SNP mutation is located in the -10 promoter region of gene *dinI* (Figure 3.3B). It was established experimentally that this single nucleotide change altered the expression of the downstream *gfoA* gene and made Gifsy-1 defective (Owen *et al.*, 2017). The exactly same SNP was identified in all UK-ST313

isolates, suggesting that the Gifsy-1 prophages carried by the UK-ST313 were also defective.

All the UK-ST313 isolates carry prophage ST64B, except for one isolate (U16). ST64B is a lambdoid phage associated with modulation of host immune response during infection (Günster *et al.*, 2017; Yang *et al.*, 2015). The ST64B prophage is defective in the ST19 strain 14028s and the ST313 strain D23580 due to a frame-shift mutation that split the tail gene into two genes *sb21* and *sb22* (Figure 3.3C; Figueroa-Bossi and Bossi, 2004; Owen *et al.*, 2017). The same frame-shift mutation was identified in all the UK-ST313 isolates, which is likely to make the ST64B prophages defective in the UK-ST313 isolates.

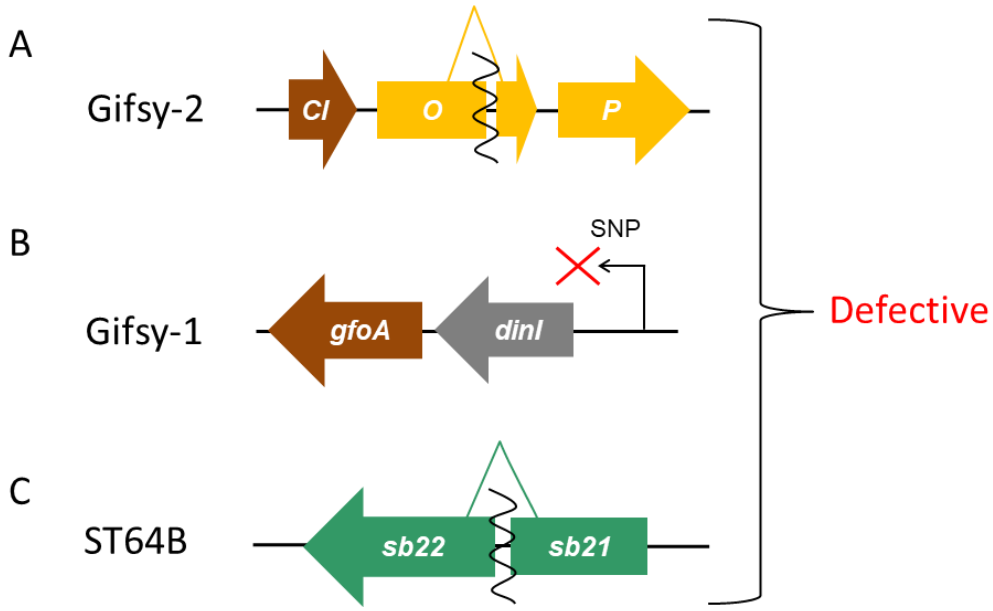


Figure 3.3 The prophage Gifsy-2 (A), Gifsy-1 (B) and ST64B (C) are defective in UK-ST313. Arrows represent a single gene and black arrow in Gifsy-1 represents promoter of *dinI*. The wavy line inside arrows represents mutations inside the gene. The figure is modified from Owen *et al.*, 2017.

Genome alignment between UK-ST313 and African ST313 L2 identified a Fels2-like prophage in 8 of the total 64 UK isolates. The Fels2-like prophage is absent from African ST313 L2 but is present in the ST313 L3 representative strain BKQZM9 (Pulford *et al.*, 2021). This Fels2-like prophage has 99.9% nucleotide identity to the prophage RE2010 that is found in many *Salmonella* serovars including *S.* Enteritidis, one of the major serovars responsible for iNTS in sub-Saharan Africa (Feasey *et al.*, 2016; Pulford *et al.*, 2021).

The similarity in the prophage content between UK-ST313 and L3 would be consistent with UK-ST313 being an intermediate phylogenetic group between L2 and L3 as has been suggested by previous phylogenetic analysis (Ashton *et al.*, 2017).

3.3 BTP1 homologs

3.3.1 Gene content of the four BTP1-like prophages

Following identification of the four homologs of BTP1 prophages, $BTP1^{UK1}$, $BTP1^{UK2}$, $BTP1^{UK3}$ and $BTP1^{UK4}$ during genome alignment, the location of each prophage was identified (Table 3.1). All the UK BTP1-like prophages were located downstream of the chromosomal *proA* gene, sharing the same position as BTP1. The identity of the sequence between the phage *attR* and *attL* sites for each UK prophage and BTP1 at the nucleotide level is shown in Table 3.1.

Table 3.1. Similarity between the BTP1 prophage and its UK homologs.

	Nucleotide Similarity (%)	Prophage location between	Plaque-forming ability*
BTP1		*proA* & *dhaF*	Yes
BTP1^{UK1}	74	*proA* & *dhaF*	Yes
BTP1^{UK2}	53	*proA* & contig ends	No
BTP1^{UK3}	62	*proA* & contig end	Yes
BTP1^{UK4}	91	*proA* & end of *STMMW_04141*	No

*Plaque-forming ability was tested on the D23580 ΔΦ strain (JH3949, Table 2.2)

Recent study showed that the novel ST313 lineage L3 carries a P22-like prophage which is 99.9% identical to BTP1^{UK1} (Pulford *et al.*, 2021). Because *S.* Typhimurium ST313 isolates have also been reported in Brazil (Almedia *et al.*, 2017), BLASTn was used to identify BTP1-like prophages in the Brazilian ST313 strains (accession numbers available in Appendix I). It was discovered that 8 of the 9 Brazilian ST313 isolates carried the BTP1^{UK1} prophage. The results suggest that BTP1^{UK1} is associated with ST313 variants that have been isolated from three continents, namely Europe, America and Africa. The wide geographical spread of the BTP1^{UK1} prophage raised the possibility that all these ST313 clusters share the same ancestor.

To investigate the gene content of BTP1 homologs, totally 293 *S.* Typhimurium strains including ST19 and ST313 strains (Details in Appendix I) were compared for the presence or absence of BTP1 genes. To put these prophages into a

phylogenetic context, gene contents of BTP1 for all the 293 strains were then mapped onto a phylogenetic tree with the assistance of Yan Li (Figure 3.4).

UK-ST313 is a diverse group of *Salmonella* isolates, and the majority of UK-ST313 is associated with an intermediate grouping that lies between ST313 L1 and L2 in Africa (Figure 3.4; Ashton *et al.*, 2017). Both Brazilian ST313 and African ST313 L3 are located within the UK-ST313 grouping on the phylogeny, and all carry $BTP1^{UK1}$ (Figure 3.4). In addition, L3 contains the same RE2010 prophage as some of the UK-ST313 isolates. It has been proposed that an international transmission event was responsible for arisal of ST313 L3 in Malawi (Pulford *et al.*, 2021). Here the discovery that the $BTP1^{UK1}$ prophage is carried by UK-ST313 isolates, Brazilian ST313 isolates and African ST313 L3 isolates offers new evidence to support the international and inter-continental transmission of *S.* Typhimurium ST313.

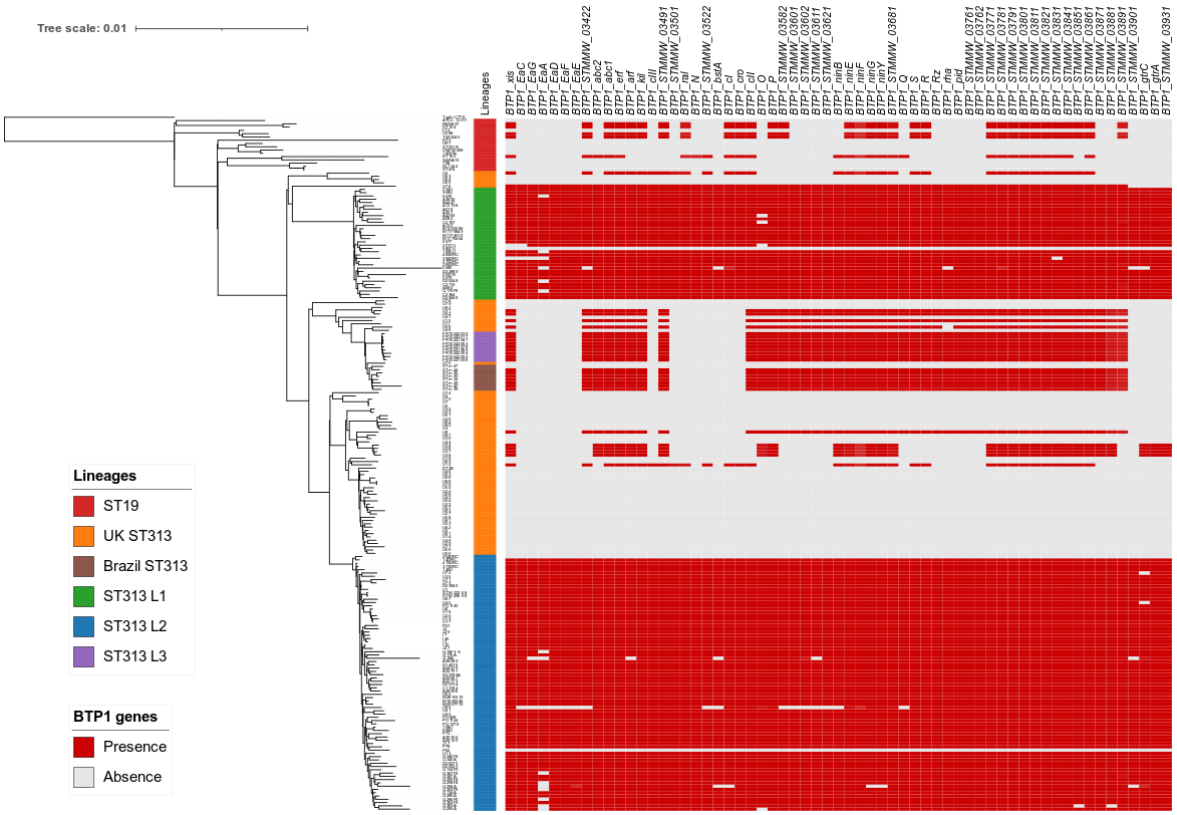


Figure 3.4. The gene complement of BTP1 mapped onto a phylogenetic tree of *S.* Typhimurium ST19 and ST313 from Africa and the UK. The names and genome accession numbers of each of the 293 *S.* Typhimurium strains is given in Appendix I. RAxML-NG v0.9.0 (Kozlov *et al.*, 2019) was used to build a phylogenetic tree with substitution model GRT + G. A heatmap is used to indicate the presence and absence of BTP1 gene in each *S.* Typhimurium isolates. The red colour indicates the presence of individual BTP1 genes, and the grey colour shows gene absence. This figure was constructed with the help of Yan Li.

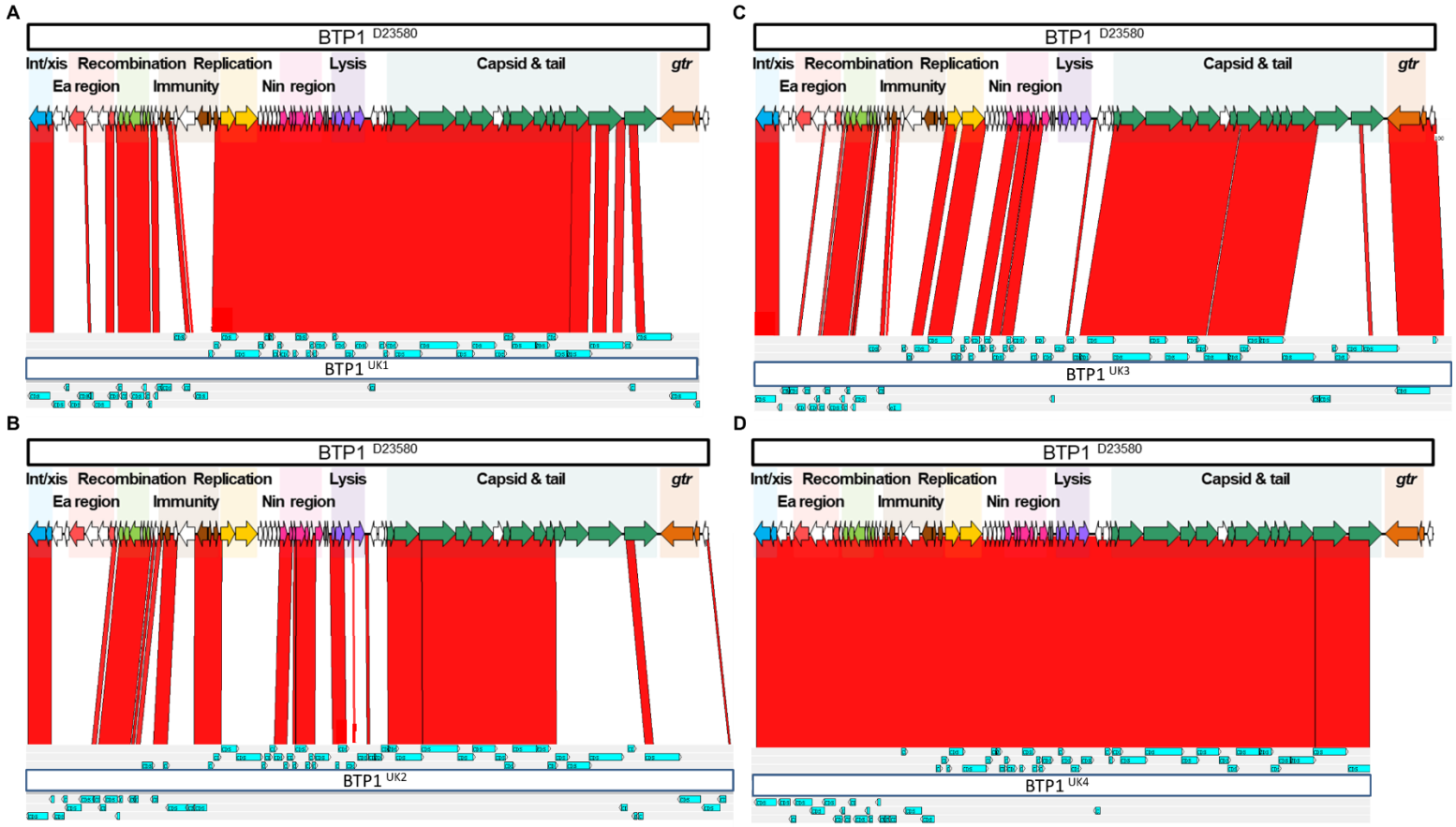


Figure 3.5. The similarities and differences in gene content between BTP1 and the four BTP1-like homologs. Genomic comparison between BTP1 and the four BTP1-like prophages was done with ACT (Carver *et al.*, 2008). The red region indicates nucleotide sequence identity > 90%. Functional regions of BTP1 (Owen *et al.*, 2017) are shown at the top of each panel. Coding sequences within BTP1-like prophages are shown in blue arrows at the bottom.

The ORFs of these homologs were characterized by the RAST server (Section 2.8.4, Details in Appendix II; Aziz *et al.*, 2008). The BTP1 genome has been characterised recently and the various ORFs were classified into 10 functional blocks: *int/xis*, *ea* region, Recombination, Immunity, Replication, *nin* region, Lysis, Capsid & tail and *gtr* region (Owen *et al.*, 2017). Detailed comparisons between BTP1 and BTP1-like homologs were performed using ACT (Artemis Comparison Tool) with 90% identifity of nucleotide sequence as the threshold (Figure 3.5).

BTP1^{UK1} was found in 5 of the 64 UK-ST313 and shares 74% nucleotide identity with BTP1. BTP1^{UK1} has almost identical functional regions, from the Replication region to the Capsid region and mainly differs in the *ea*, *gtr* and Immunity regions (Figure 3.5). Unlike BTP1 which carries the CI repressor, BTP1^{UK1} has the *sieB* gene. The SieB protein is 92% identical at the amino acid level to the same protein in P22. SieB is a superinfection exclusion phage immunity protein (Susskind *et al.*, 1974). BTP1 encodes the BstA defense system which protects D23580 against attack by other lytic phages, which may provide a selective advantage for the survival of ST313 L2 (Owen *et al.*, 2021). It is possible that the SieB protein encoded by BTP1^{UK1} plays a similar role to BstA against other phages. BTP1^{UK1} also carries the *mnt* gene encoding the Mnt repressor which could contribute to maintainance of the lysogen (Vershon *et al.*, 1987).

The BTP1^{UK2} prophage shared 53% sequence identity with BTP1, and had some important similarities and differences which are summarised in Figure 3.5. The

Recombination regions of BTP1^{UK2} and BTP1 are identical. BTP1^{UK2} carries a CI repressor protein that is 100% identical to the BTP1-encoded CI at the amino acid level. In contrast, the *gtr* operon is absent from BTP1^{UK2}.

A broader BLAST-based search for the presence of BTP1^{UK2} amongst hundreds of *Salmonella* genomes (Section 2.8.4) showed that the prophage shares 99.95% nucleotide identity with a prophage carried by a *S.* Typhimurium ST213 isolate from Mexico (Silva *et al.*, 2019) and is 99.86% identical to phage SE1 from *S.* Enteritidis (Llagostera *et al.*, 1986). However, BTP1^{UK2} prophage was defective in UK-ST313 (Table 3.1). The data suggests that the reason that BTP1^{UK2} did not produce viable phage particles was due to the absence of genes corresponding to the capsid region of BTP1, but this remains to be proven experimentally.

BTP1^{UK3} is 62% identical to BTP1 and shares the same Recombination, Replication, Capsid and *gtr* regions as BTP1 (Figure 3.5). The BTP1^{UK3} prophage was able to form phage plaques but had a CI repressor that was genetically distinct to the CI from BTP1. Like BTP1^{UK1}, BTP1^{UK3} carries the superinfection-exclusion *sieB* gene and lacks the *bstA* gene of BTP1. A broader BLAST-based search for the presence of BTP1^{UK3} amongst hundreds of *Salmonella* genomes (Section 2.8.4) identified a prophage that was almost identical, with only a 1-nucleotide-difference. This prophage was carried by the *Salmonella enterica* strain 16A242 from wild birds in South Korea (GenBank: CP020922).

The BTP1^{UK4} was much rarer than the other BTP1-like prophages, and was only identified in a single UK isolate, U76. The BTP1^{UK4} prophage was 100% identical to BTP1 at the nucleotide level, from the very beginning part to the centre of BTP1, the *STMMW_03901* gene, which encodes the tail spike protein (Figure 3.5). A large chromosomal region downstream of BTP1 was absent from the genome of U76, extending from the gene *dhaF* to *STMMW_04141*. The absent region contained the *stb* fimbrial operon, which is involved in intestinal persistence during *Salmonella* infection (Weening, *et al.*, 2005), and the *leuC2* and *leuD2* leucine biosynthesis genes (Yang and Kessler, 1974). However, as the BTP1^{UK4} was defective (Table 3.1), this prophage was not further studied in this thesis.

In summary, a comparison of the 10 functional blocks between BTP1 and all the BTP1-like prophages is shown in Figure 3.6 using a threshold of 50%-70% identify. It is clear that all the BTP1-like prophages share the identical *int/xis* and Recombination regions, and most part of *nin* region.

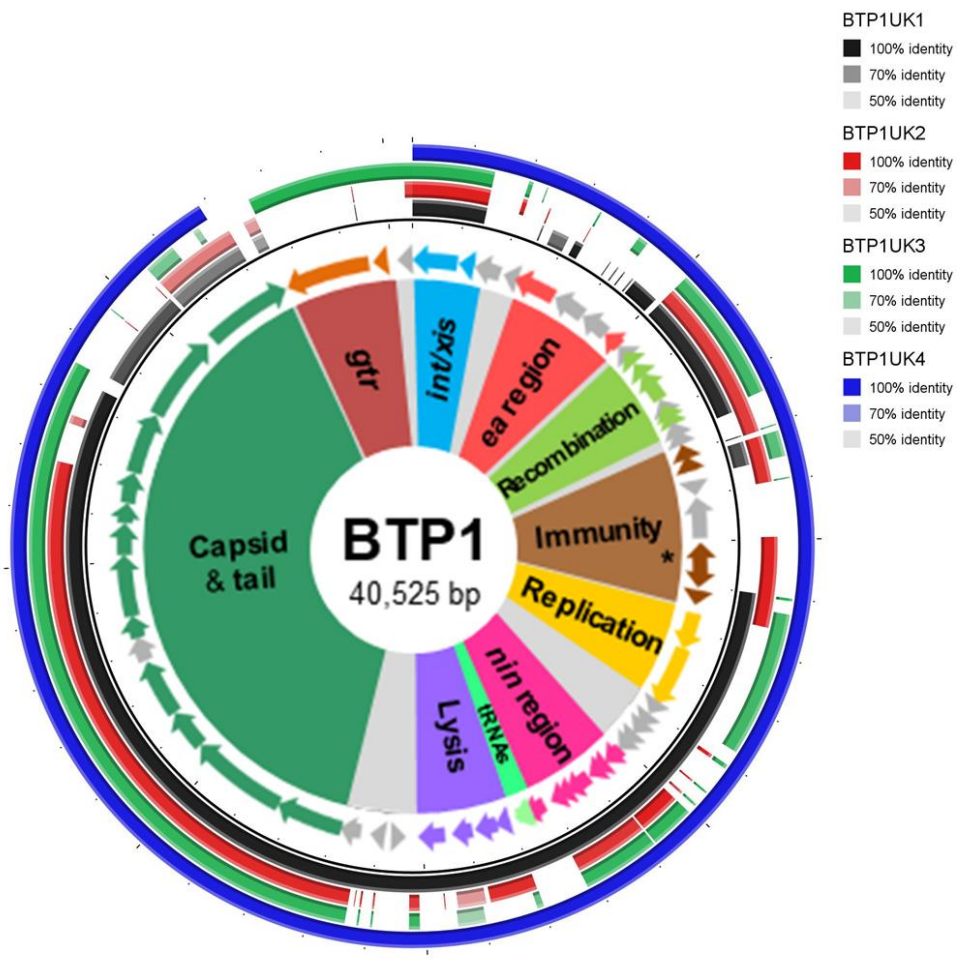


Figure 3.6. Functional genetic architecture of BTP1-like prophages. The diagram represents the similarities and differences between the nucleotide sequences of the BTP1-like prophages and BTP1. The ten functional blocks of the BTP1 genome are shown in different colours and regions of unknown function are indicated in grey (Owen *et al.*, 2017). The CI repressor is marked with an asterisk in the Immunity region. Four UK BTP1 homologs are the outermost rings with different colours. The percentage nucleotide identity is indicated by colour intensity, as described in the key on the right-hand side.

3.3.2 Host range determination of BTP1 and BTP1-like prophages

The *Salmonella* serovars and strains that a phage can lyse determines the host range of the phage (Kutter, 2009). Previously, the host range of phages P22 and BTP1 have been reported for a small number of Africa *S.* Enteritidis and *S.* Typhimurium strains (Rodwell *et al.*, 2021). To determine the extent of the host

range of P22, BTP1 and BTP1-like prophages BTP1^{UK1} and BTP1^{UK3}, experiments were done with 36 *S.* Typhimurium isolates belonging to different sequence types or lineages, plus 39 isolates of different *Salmonella* serovars. Plaque-purified lysates of individual phages were used to infect particular *Salmonella* strains using the method in Section 2.5.1, and the ability to form plaques is reported in Figure 3.7. The results showed that BTP1 and its UK homologs generally had a broad host range.

Specifically, BTP1^{UK1} shared the same host range as P22, while the host range of BTP1^{UK3} resembled that of BTP1. As described in Section 3.1, the BTP1 prophage encodes two distinct phage defense systems. One is associated with the *gtr* locus that is responsible for enzymatic modification of the cell-surface associated O antigen and LPS (Kintz *et al.*, 2015). The second is mediated by BstA (Owen *et al.*, 2021). As African ST313 L1 and L2 already carried the BTP1 prophage, the BTP1-encoded phage defense systems prevented superinfection by BTP1. The data suggests that the resistance of African ST313 L1 and L2 to the BTP1^{UK3} phage was likely to be due to the same reason. BTP1^{UK3} shares the same *gtr* region as BTP1, suggesting that the consequent O antigen modification prevents BTP1^{UK3} plaque formation on ST313 L1 and L2 strains. This remains to be verified experimentally.

Serotyping was originally developed as a phenotypic method for the characterization and classification of *Salmonella* isolates, and was based on the White-Kauffmann-Le Minor scheme (*Salmonella* Subcommittee of the Nomenclature Committee of the

International Society for Microbiology, 1934). The serotype is determined from two protein structures on the surface of *Salmonella*: the O antigen, the outmost part of the cell surface, and the two phases of the flagella H antigen, H1 and H2 (Diep *et al.*, 2019). According to modern serovar nomenclature (Grimont and Weill, 2007), *S.* Typhimurium belongs to serotype Group B and has antigens:

O: 1, 4, 5, 12; H1: i; H2: 1, 2.

Serotypes and the related antigens of each *Salmonella* serovar used for host range determination are shown in Table 3.2. The results from Figure 3.7 showed that the phage P22, BTP1 and BTP1 homologs infected almost all the *Salmonella* strains that had a Group B or D1 serotype. This made biological sense because the Group D1 shares O antigens 1 and 12 with Group B. In contrast, serogroups C1 and E have a totally different O antigen type to *S.* Typhimurium. As detailed in Figure 3.7, the serogroup C and E isolates were resistant to all the phages tested. Taken together, these findings are consistent with the BTP1 and BTP1-like phages sharing the same type of LPS receptor as phage P22 (Lindberg *et al.*, 1970).

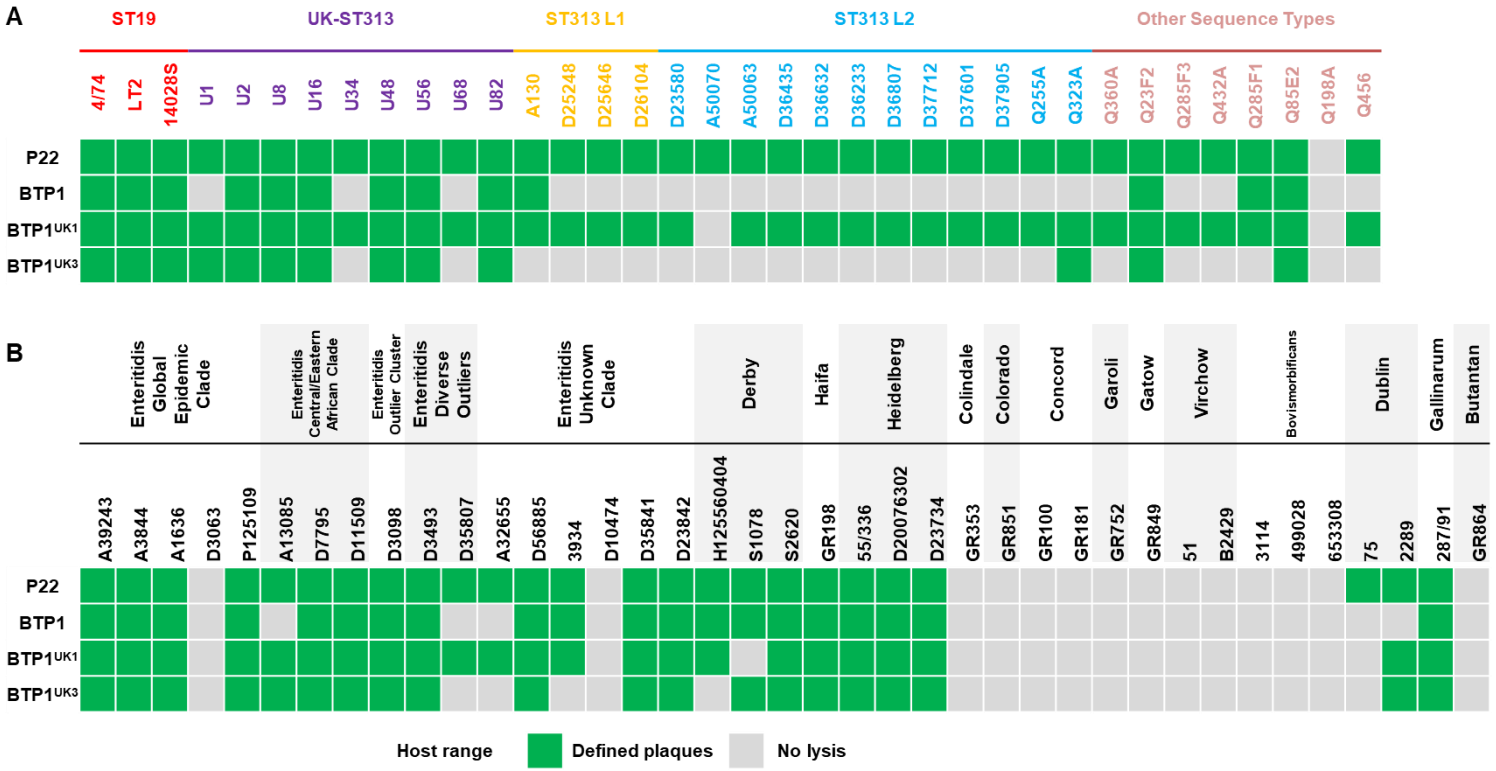


Figure 3.7. Host range of P22, BTP1 and BTP1-like phages. Host range was determined by spot assay on a range of *S.* Typhimurium strains including ST19, ST313 and other Sequence Types (A) and other *Salmonella* serovars (B). The strains were derived from the 10,000 *Salmonella* genomes project (Perez-Sepulveda *et al.*, 2021), and were selected to represent this diversity of the *Salmonella enterica sp.* Subspecies I. Strains from ST313 L1 (yellow) and L2 (blue) carry BTP1 prophage. Strains from *S.* Enteritidis were from five clades associated with epidemics: Global Epidemic, Central/Eastern African, Outlier, Diverse Outlier and Unknown Clade (Feasey *et al.*, 2016). The phage host range was recorded as lysis (defined plaques; green) or no lysis (grey), using three biological replicates.

Table 3.2. Serogroup details for different *Salmonella* serovars used for host range determination

Serovar	Serogroup*	O antigen*	H1 antigen*	H2 antigen*
Typhimurium	B	1,4,[5],12	i	1,2
Enteritidis	D1	1,9,12	g,m	
Derby	B	1,4,[5],12	f,g	[1,2]
Haifa	B	1,4,[5],12	z10	1,2
Heidelberg	B	1,4,[5],12	r	1,2
Colindale	C1	6,7	r	1,7
Colorado	C1	6,7	l,w	1,5
Concord	C1	6,7	l,v	1,2
Garoli	C1	6,7	i	1,6
Gatow	C1	6,7	y	1,7
Virchow	C1	6,7,14	r	1,2
Bovismorbificans	C2-C3	6,8,20	r,[i]	1,5
Dublin	D1	1,[9],12	g,p	\
Gallinarum	D1	1,9,12	\	\
Butantan	E1	3,{10}{15}{15,34}	b	1,5

***Serogroup information from Grimont and Weill, 2007**

3.3.3 Spontaneous induction level of the BTP1-like homolog prophages

S. Typhimurium ST313 from D23580 exhibits a high level of spontaneous induction of BTP1 (approximately 10^9 PFU/mL), compared with the closely related phage P22 (approximately 10^5 PFU/mL) (Owen *et al.*, 2017). Spontaneous prophage induction resulted from the activation of lytic pathways that kill the lysogen and release viable phages (Cortes *et al.*, 2019). This lysis resulted in the lysis of about 0.2% of *S.* Typhimurium ST313 D23580 cells in an overnight culture (Owen *et al.*, 2017).

An experimental approach was used to determine whether the BTP1-like prophages also exhibited high levels of spontaneous induction. The spontaneous induction

levels of all the BTP1-like prophages were measured using supernatants from representative UK-ST313 isolates (Table 3.3), following the method described in Section 2.5.2. As reported in Table 3.1, prophages BTP1^{UK2} and BTP1^{UK4} were unable to form phage plaques, and were classified as defective. BTP1^{UK1} had a spontaneous induction level lower than P22. BTP1^{UK3} showed the same high level of spontaneous induction as BTP1.

Table 3.3. Spontaneous induction level of each prophages. The numbers of phages produced in individual overnight culture was determined as detailed in Section 2.5.2.

Prophage	Host	PFU/mL	Reference
P22	D23580	10^5	Owen et al., 2017
BTP1	D23580	10^9	Owen et al., 2017
BTP1^{UK1}	U8	10^4	This study
BTP1^{UK1}	U15	10^4	This study
BTP1^{UK2}	U5	0	This study
BTP1^{UK3}	U34	10^9	This study
BTP1^{UK3}	U36	10^9	This study
BTP1^{UK3}	U37	10^9	This study
BTP1^{UK3}	U38	10^9	This study
BTP1^{UK4}	U76	0	This study
Lawn: JH3949 (D23580Δϕ)			

3.3.4 Comparative genomic analyses of the BTP1-like prophage CI proteins

There has been speculation about the mechanism responsible for the high level of spontaneous induction of BTP1 since its discovery (Owen *et al.*, 2017). Furthermore, the physiological significance of the high spontaneous induction level of BTP1 has

been unclear. One possible reason for the high levels of spontaneous induction could be amino acid changes in the CI repressor protein that influence activity and modulate the ability to maintain prophage lysogeny. It has been shown previously that overproduction of the phage Lambda CI repressor can prevent the spontaneous induction of prophages in *E. coli* (Czyz *et al.*, 2001).

Accordingly, the protein sequences of the immunity regions of BTP1 and the BTP1 homologs were investigated. In comparison with BTP1, the four BTP1-like prophages had distinct Immunity regions (Figure 3.6). To investigate the molecular basis of the difference between spontaneous induction levels of the BTP1-like prophages, the amino acid sequences of the CI repressor for each prophage were aligned, as described in Section 2.8.4.

Figure 3.8 shows that the CI repressor of BTP1^{UK3} is identical to P22. Because BTP1^{UK3} had the same high level of spontaneous induction as BTP1, and P22 had a much lower level of spontaneous induction, It is now clear that changes in the amino acid sequence of the CI protein are not responsible for the high level of spontaneous induction of BTP1. Consequently, it is possible that another gene carried by BTP1 and BTP1^{UK3} is responsible for the high level of spontaneous induction of these prophages representing an important question to be addressed in future research.

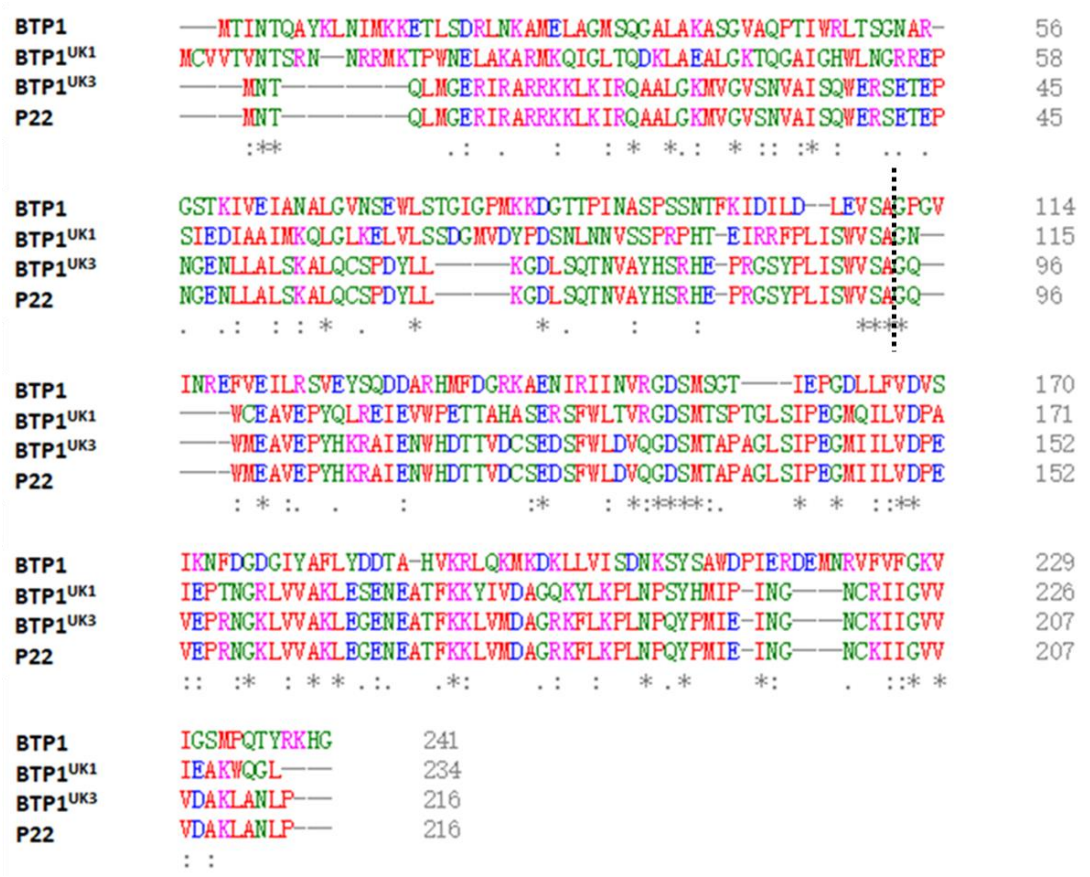


Figure 3.8. Alignment of the amino acid sequence of the CI repressor from four prophages. The entire sequence of the CI protein is shown in five rows, with amino acid numbers indicated on the right-hand side. Amino acids that are conserved in all four proteins are labelled with "*". High levels of amino acid conservation are labelled with ":". Weak similarity is shown by ".". The black dashed line represents the self-cleaveage site for each repressor (Sauer *et al.*, 1982).

3.3.5 Investigation of phenotypic properties associated with lysogeny of BTP1 and UK BTP1-like prophages

It has been established that the RDAR morphotype is a phenotypic property that is responsible for high levels of biofilm formation by *Salmonella* bacteria associated with gastroenteritis (Romling *et al.*, 1998). More recently, it has been discovered that African ST313 and UK-ST313 isolates share key phenotypes that include the lack of RDAR morphotype, swimming motility and catalase activity (Ashton *et al.*, 2017;

Owen *et al.*, 2017). The lack of RDAR phenotype of ST313 L2 is linked to a point mutation in the *bcsG* gene, which encodes cellulose biosynthetic enzyme (Singletary *et al.*, 2016) and to a truncation of 8 amino acids from the C-terminus of the CsgD regulatory protein, which is a modulator of biofilm development (MacKenzie *et al.*, 2019).

However, the RDAR phenotype of UK-ST313 isolates that carried the BTP1^{UK3} prophages has not been determined. To determine whether lysogeny by the BTP1^{UK1} or BTP1^{UK3} prophages modulated the RDAR phenotype, the RDAR phenotype assay was performed. Both strains D23580ΔΦ (prophage-free mutant) and U2 (a UK-ST313 isolate that did not carry a BTP1 or a BTP5-like prophage) were chosen as the genetic background (Figure 3.9). Lysogenic strains were constructed as described in Section 2.5.3.

The BTP1, BTP1^{UK1}, BTP1^{UK3} lysogens shared the same RDAR-negative phenotype, suggesting that all the BTP1-like prophages did not modulate biofilm formation of ST313.

Further research aimed at identifying novel phage-associated phenotypes that distinguish UK-ST313 from African ST313 would be worthwhile.

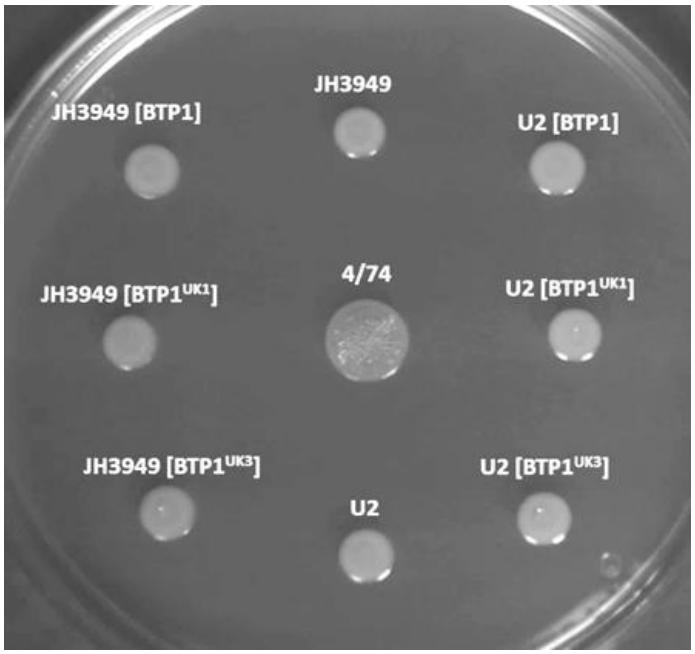


Figure 3.9. RDAR morphotype assay of BTP1 and BTP1-like phage lysogens. The RDAR phenotype was formed after two-day incubation at 25°C (Section 2.4.4). In the centre is the positive control 4/74 which shows a textured RDAR-positive appearance.The other 8 isolates are RDAR negative.

3.4 BTP5 homologs

3.4.1 Gene content of the three BTP5-like prophages

BTP5 is a novel prophage belonging to *Myoviridae* family that has key similarities to phage P2 (Owen *et al.*, 2017). Previous experiments from the Hinton Lab showed that BTP5 could form lysogens, but was unable to form phage plaques on a prophage-free derivative of D23580 (Owen *et al*, 2017). In Section 3.2, three homologs of BTP5 prophages, BTP5^{UK1}, BTP5^{UK2} and BTP5^{UK3} were identified during genome alignment of the ST313 L2 strain D23580 and all UK-ST313 isolates. Similarities between the genome of BTP5 and BTP5-like prophages at nucleotide level and plaque-forming ability are shown in Table 3.4.

Table 3.4. Similarity between the BTP5 prophage and its UK homologs.

	Nucleotide Similarity (%)	Prophage location between	Plaque-forming ability*
BTP5		***mug* & *yqjH***	**No**
BTP5^{UK1}	**74**	***cpxP* & *fieF***	**Yes**
BTP5^{UK2}	**53**	***mug* & *yqjH***	**Yes**
BTP5^{UK3}	**62**	***cpxP* & *fieF***	**Yes**

*Plaque-forming ability was tested on the D23580 ΔΦ strain (JH3949, Table 2.2)

.

The key finding from Table 3.4 was that, unlike the BTP5 phage, the BTP5^{UK1}, BTP5^{UK2} and BTP5^{UK3} phages were all viable, and capable of forming plaques on the phage-free derivative of D23580.

In terms of chromosomal location of the prophages, BTP5^{UK2} was located between the chromosomal gene *mug* and *yqjH*, as was BTP5. The other two prophages BTP5^{UK1} and BTP5^{UK3} were located between the *cpxP* and *fieF* genes, about 1 Mb away from the position of BTP5.

Although the function of BTP5 in ST313 L2 remains unknown (Owen *et al.*, 2017), the ORFs of BTP5 have been classified into 6 functional blocks: *int* region, Immunity, Replication, Capsid, Lysis and Tail region (Owen, 2017). To investigate the gene content of BTP5-like prophages, the RAST server was used to characterise the ORFs of BTP5 homologs (Section 2.8.4, in Appendix II; Aziz *et al.*, 2008). Genomic comparison between BTP5 and BTP5-like prophages was visualised by ACT (Carver *et al.*, 2008), as shown in Figure 3.10.

BTP5^{UK1} is 67% identical to BTP5 with a totally different functional region between *int* and the Capsid region. Though no function could be inferred for most of the different ORFs of BTP5^{UK1}, this prophage carries genes that encode the Cox protein and the C protein. It has been reported that the Cox protein of phage P2 inhibits the formation of C protein, and modulates viral DNA packaging (Saha *et al.*, 1987).

BTP5^{UK2} shares 91% identical nucleotide sequence with BTP5 and the phage location in the chromosome is the same as BTP5. BTP5 has the phage late control genes *gpB* (brown region at the very left in Figure 3.10), with *gbD* downstream. Both of these two genes are involved in prohead assembly of phage (Richardson *et al.*, 1988). The *gpB* gene is absent from BTP5^{UK1} and BTP5^{UK3}, but is identical in BTP5^{UK2}. In addition, BTP5 contains the *orf97* and *tum* genes (black region in Figure 3.10) that are homologous to the genes in the SOS operon of coliphage 186 (Owen *et al.*, 2017; Shearwin *et al.*, 1998); these two genes are found in BTP5^{UK2} but not BTP5^{UK1} and BTP5^{UK3}. The function of the BTP5^{UK2}-specific ORFs is unknown.

BTP5^{UK3} is 66% similar to BTP5 and is 89% similar to BTP5^{UK1} at the nucleotide level. The biggest difference between BTP5^{UK3} and BTP5^{UK1} was in the leading part of Replication region with the addition of 4 unknown genes.

3.4.2 Discussion – BTP5-like prophages

Overall, the UK-ST313 isolates carried three BTP5-like prophage variants. Unlike

BTP5, all three BTP5-homologs were viable phages with plaque-forming ability. Because BTP5^{UK2} has 91% identity with BTP5 at the nucleotide level, further investigation to characterise BTP5^{UK2} could help to explain why BTP5 is unable to form plaques.

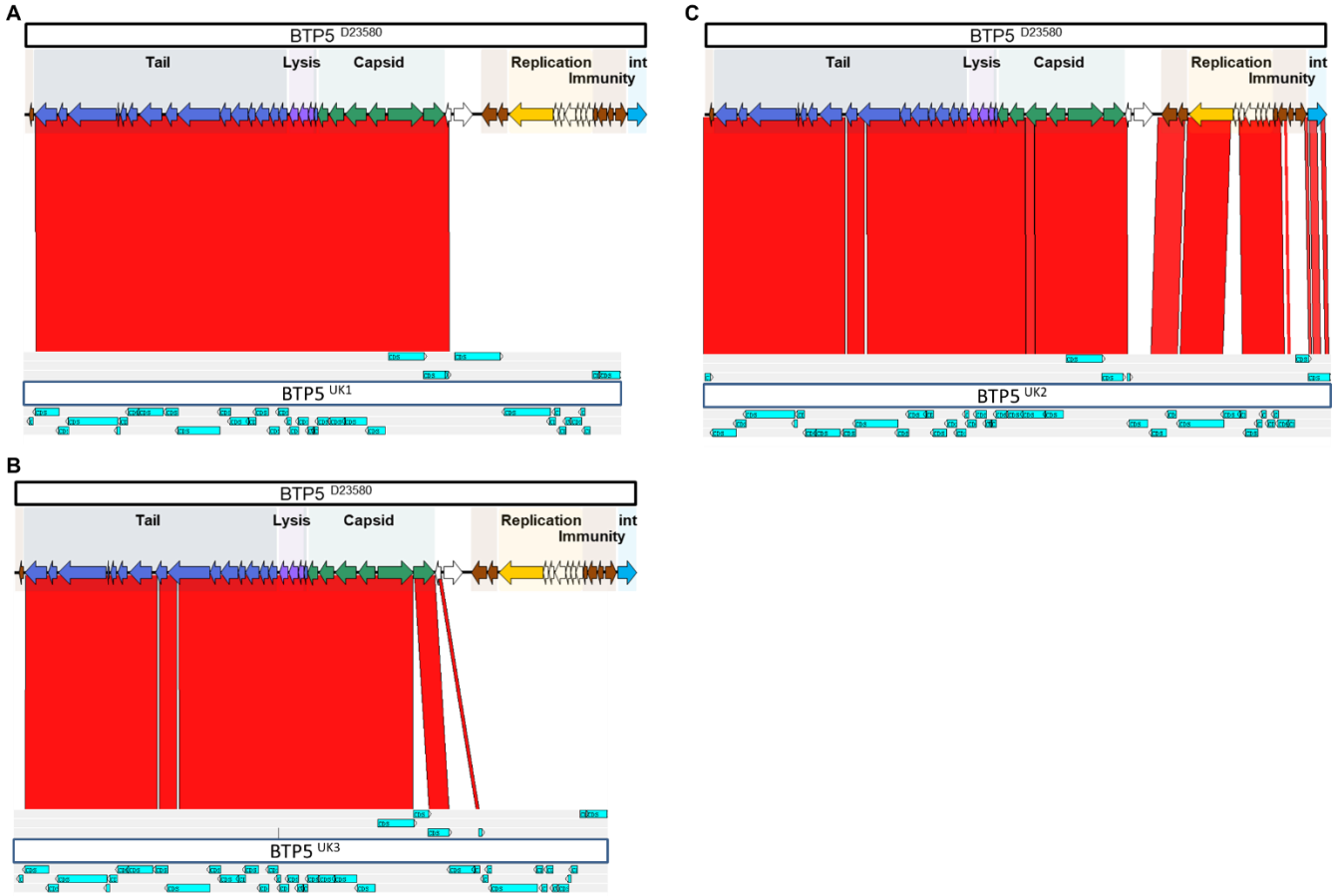


Figure 3.10. The similarities and differences in gene content between BTP5 and the three BTP5-like homologs. Genomic comparison between BTP5 and the three BTP5-like prophages was done with ACT (Carver *et al.*, 2008). The red region indicates nucleotide sequence identity > 90%. Functional regions of BTP5 (Owen *et al.*, 2017) are shown at the top of each panel. Coding sequences within BTP5-like prophages are shown in blue arrows at the bottom.

3.4.3 Host range determination

To investigate the host range of the BTP5-like phages, experiments was done for all the BTP5 homologs using the same *Salmonella* isolates as described in Section 3.3.3. Briefly, plaque-purified lysates of the individual phages were used to infect particular *Salmonella* strains using the method in Section 2.5.1.

The results showed that BTP5-like prophages had a broad host range (Figure 3.11). Specifically, the BTP5^{UK2} and BTP5^{UK3} phages formed plaques on a very wide group of *Salmonella* strains that included all serovars except serotype group C1 (Figure 3.11). The *Salmonella* serovars Colindale, Colorado, Concord, Garoli, Gatow and Virchow belong to serotype group C1 (Table 3.2; Grimont and Weill, 2007). The *S.* Enteritidis strains that were tested represented five distinct clades (Feasey *et al.*, 2016); only the Central East African clade was resistant to all three of the BTP5-like prophages.

All ST313 L2 strains were resistant to BTP5^{UK1} and were susceptible to BTP5^{UK3}. As the only difference betweens BTP5^{UK1} and BTP5^{UK3} are the additional genes in the replication region of BTP5^{UK3}, it is possible that these genes are involved in the O or H antigen modification or phage superinfection immunity. Taken together, the findings reveal that the three BTP5 homologs have a broad host range.

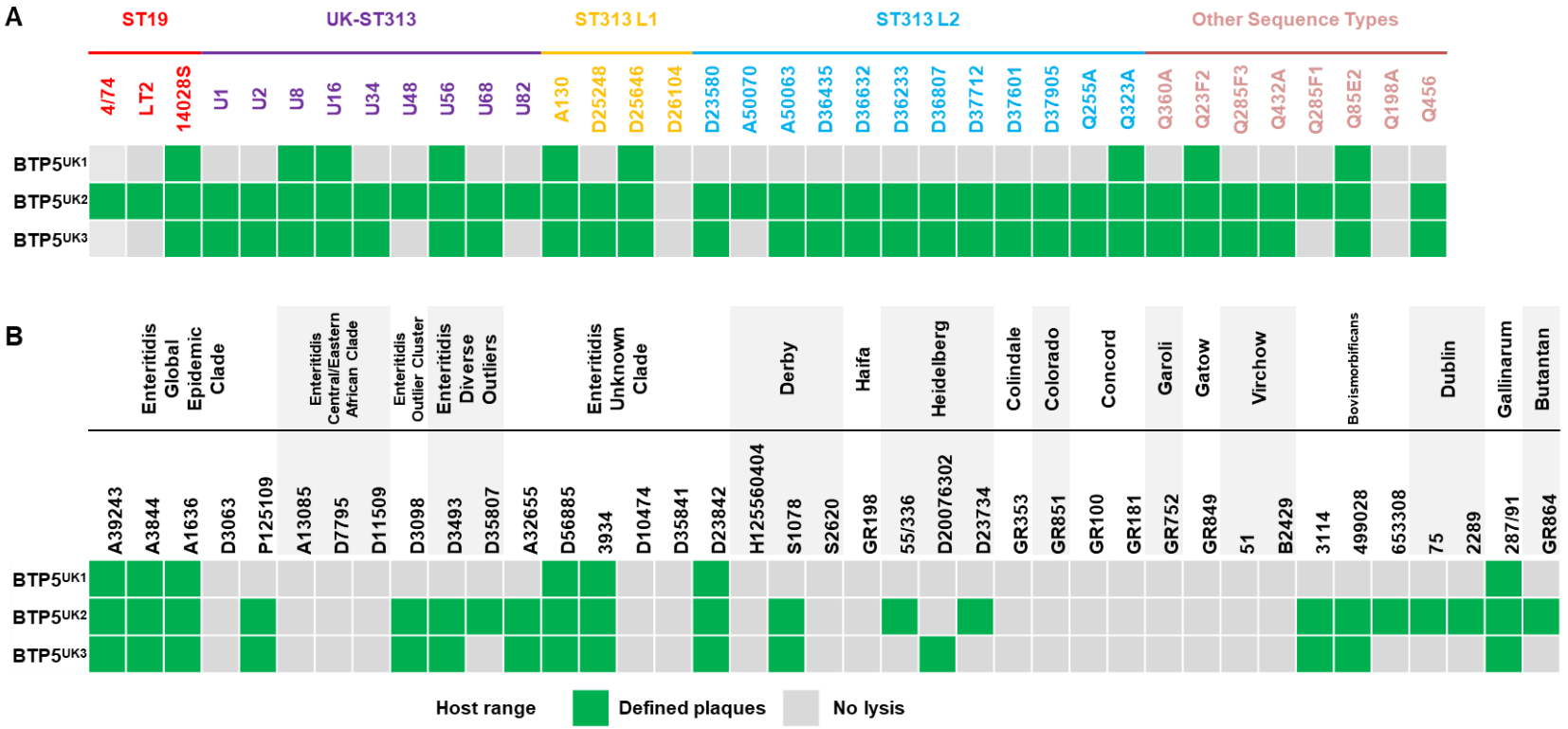


Figure 3.11. Host range of BTP5-like phages. Host range was determined by spot assay on a range of *S.* Typhimurium strains including ST19, ST313 and other Sequence Types (A) and other *Salmonella* serovars (B). The strains were derived from the 10,000 *Salmonella* genomes project (Perez-Sepulveda *et al.*, 2021), and were selected to represent this diversity of the *Salmonella enterica sp.* Subspecies I. Strains from ST313 L1 (yellow) and L2 (blue) carry BTP5 prophage. Strains from *S.* Enteritidis were from five clades associated with epidemics: Global Epidemic, Central/Eastern African, Outlier, Diverse Outlier and Unknown Clade (Feasey *et al.*, 2016). The phage host range was recorded as lysis (defined plaques; green) or no lysis (grey), using three biological replicates.

3.5 Discussion

The prophage component of the accessory genome plays an important role in bacterial fitness and virulence (Adler *et al.*, 2021). ST313 L2 strains carry 3 prophages Gifsy-1, Gifsy-2 and ST64B that have previously been studied in *S.* Typhimurium ST19, and the 2 novel prophages BTP1 and BTP5 (Kingsley *et al.*, 2009; Owen *et al.*, 2017). The *bstA* gene encoded by BTP1 was involved in both virulence and anti-virulence of ST313 (Herrero-Fresno *et al.*, 2014; Spiegelhaier *et al.*, 2020). The function of the BstA protein permits prophages to defend host cells against exogenous phage attack, which could give ST313 L2 bacteria a selective advantage (Owen *et al.*, 2021).

The high rate of spontaneous induction of BTP1 is caused by an unknown mechanism. It was hypothesized that the high level of spontaneous induction of BTP1 might alter the growth dynamics of ST313 during infection, contributing to fitness in the host.

To investigate the role of BTP1 and BTP5 in Africa ST313 L2, similar prophages were searched from a variety of *Salmonella* genomes. ST313 isolates from beyond the African continent were found from both Brazil and the UK (Almeida *et al.*, 2017; Ashton *et al.*, 2017). Some of these isolates belong to a novel ST313 subgroup UK-ST313, which caused gastroenteritis rather than invasive disease (Ashton *et al.*, 2017). The altered pathogenesis of UK-ST313 compared with African ST313 L2,

leads to a question of whether prophages associated with non-African ST313 isolates might confer interesting biological properties. In this chapter, genome alignment between African strain D23580 and 64 UK-ST313 isolates was performed. Totally 4 homologs of BTP1 and 3 homologs of BTP5 were identified.

Although the homologs of BTP1 and BTP5 shared many identical sequence regions with BTP1 and BTP5, most of the UK-specific regions carried ORFs of unknown function (Figure 3.6). However, a number of potentially interesting gene differences were identified. BTP1^{UK1} and BTP1^{UK3} both carry a superinfection exclusion system encoded by the *sieB* gene instead of the BTP1-encoded BstA system, which could contribute to protecting the bacterial host against phage attack (Owen *et al.*, 2021; Susskind *et al.*, 1974). BTP1^{UK1} also carries the *mnt* gene that helps the maintenance of prophage lysogen (Vershon *et al.*, 1987).

BTP1^{UK3} shared the same high level of spontaneous induction as BTP1. The fact that the CI repressor of BTP1^{UK3} was identical to the P22 CI protein at the amino acid level suggested that the molecular basis for the high level of spontaneous induction of BTP1 was not linked to the CI protein.

It has previously been reported that BTP5 was not capable of forming visible phage plaques (Owen *et al.*, 2017). BTP5^{UK2} shared most of the genes that have been characterised in BTP5 and it formed viable plaques, suggesting that one or more

BTP5^{UK2} specific genes are responsible for the ability of BTP5^{UK2} to form viable viral particles that are capable of plaque formation. This could be worthy of experimental investigation in the future.

One limitation of the bioinformatic analyses that have been reported in this chapter is that the annotation was done with the RAST server towards the beginning of the project in 2019 (Aziz *et al.*, 2008). Subsequent iterations of the database may well result in an improved annotation that would contribute more to our understanding of the genome of the UK-ST313-associated BTP1 and BTP5-like prophages.

Future experimental approaches should focus on identifying the key genetic similarities between BTP1^{UK3} and BTP1 that responsible for the high spontaneous induction level of both prophages. Furthermore, a detailed functional genomic investigation of BTP5^{UK2} and BTP5 could identify novel genes that allow BTP5^{UK2} to form viable phage particles capable of plaque formation.

To conclude, BLAST was a useful tool for identifying the prophages component of the accessory genome that shared extensive similarity with BTP1 in the UK-ST313 isolates. An identical prophage of BTP1^{UK1} was found in 8 of the 9 Brazil ST313 isolates and a 99.9% identical prophage was identified in African ST313 L3 (Pulford *et al.*, 2021). As L3 represents a phylogenetic intermediate between L1 and L2 (Pulford *et al.*, 2021), it was hypothesized that UK-ST313 is actually a diverse ST313

grouping that included African L3 and had been transmitted from the UK to both the American and African continents.

As further evidence that the BTP1-like prophages might reveal international transmission of *Salmonella enterica*, a prophage similar to BTP1^{UK2} was found in a *S.* Typhimurium ST213 isolate from Mexico and in many *S.* Enteritidis isolates (Llagostera *et al.*, 1986; Silva *et al.*, 2019).

Taken together, the existence of the BTP1-like prophages supports the hypothesis that there has been significant inter-continental transmission of the diverse *S.* Typhimurium UK-ST313 lineages, a topic that is being investigated independently in the Hinton Lab.

Chapter 4.
Comparison of the phenotypic properties of African *S*. Typhimurium ST313 Lineage 2 and Lineage 2.2

4.1 Introduction

Salmonella Typhimurium ST313 is one of the major causes of iNTS in sub-Saharan Africa (Kingsley *et al.*, 2009). Phylogenetic analysis has identified three lineages L1, L2 and L3 (Okoro, *et al.*, 2012; Pulford, *et al.*, 2021). Recent studies showed that multi-drug resistance (MDR) and accumulated pseudogenes contributed to the success of ST313 in Africa (Pulford et al., 2021). The divergence of L1 and L2 of African ST313 occurred after L2 became resistant to chloramphenicol, which was the first-line treatment for septic infections in Malawi in 2001, and then L2 became the primary cause of iNTS in Africa (Okoro, *et al.*, 2012; Tsai and Coombes, 2021). The antibiotic sensitive ST313 L3 first emerged in 2016. The pseudogenes in ST313 L3 is consistent with adaptation from an intestinal lifestyle to a systemic lifestyle (Pulford *et al.*, 2021). In addition, the azithromycin-resistant ST313 sub-lineage L2.1 was first identified in Democratic Republic of Congo in 2008 (Van Puyvelde *et al.*, 2019).

A new sub-lineage of *S.* Typhimurium ST313, L2.2, has emerged in Malawi from 2003 (Msefula *et al.*, 2012), and became increasingly prevalent between 2004 and 2018 as shown in Figure 4.1. This sub-lineage has been associated with the same bloodstream infections as L2, and had a similar multi-drug resistance profile (B. Kumwenda *Pers. Comm.*).

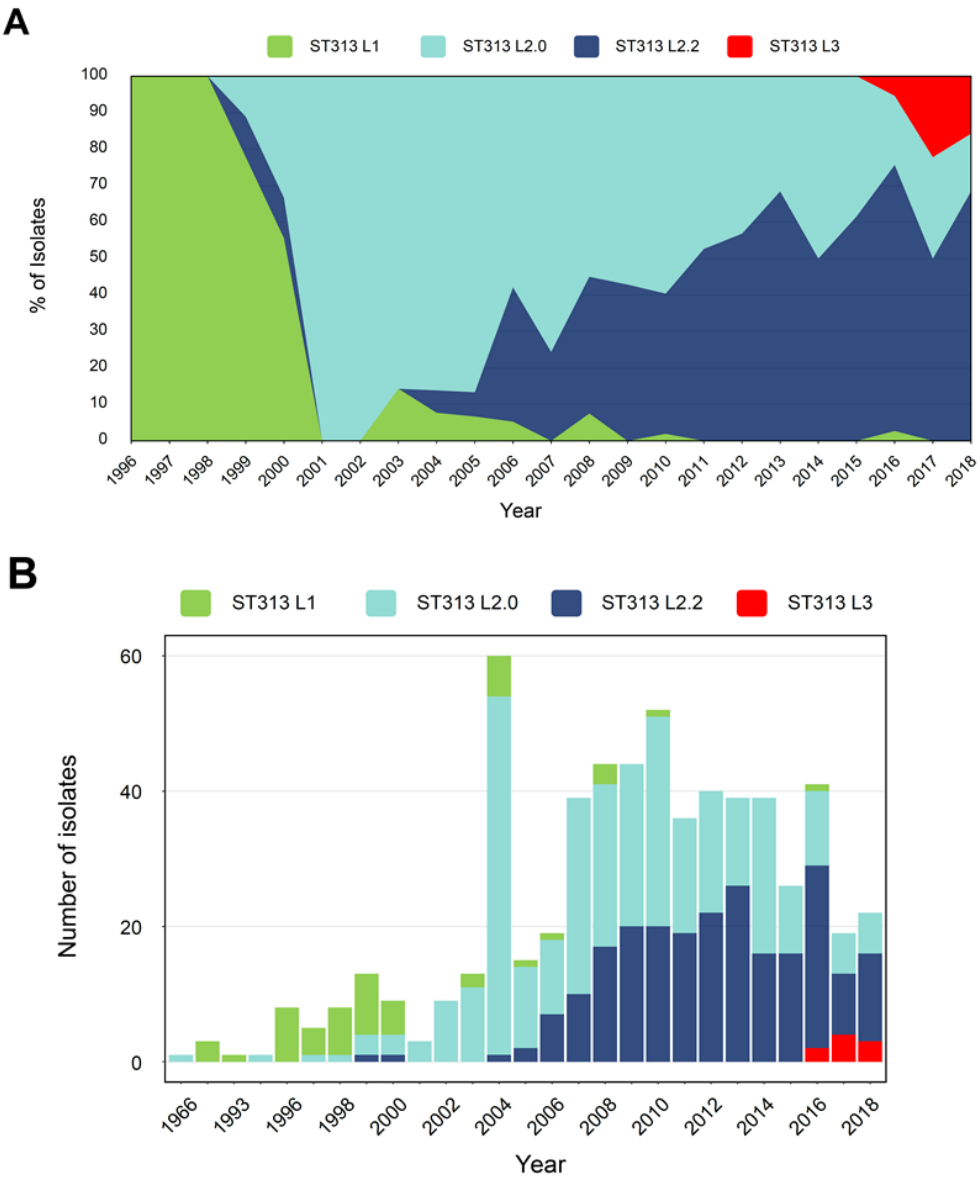


Figure 4.1. Arisal of S. Typhimurium ST313 sublineage 2.2 in Malawi. (A) Evolution of *S.* Typhimurium lineages in Blantyre, Malawi from 1996 to 2018. The genomes from 549 *S.* Typhimurium ST313 isolates from human bllodstream infections at the Queen Elizabeth Central Hospital in Blantyre, Malawi. Proportions are shown per year as follows: L1 is green, L2 is light blue, L2.2 is dark blue and L3 is red. (B) The total number of isolates per year. Genomic data are from Pulford *et al.*, 2021. Analysis was performed and visualised by Yan Li and Alex Predeus.

To investigate the genetic differences between L2 and L2.2, a comparative genomic analysis was conducted between L2 strain D23580 and the representative L2.2 strain D37712. Both D23580 and D37712 carried the same five prophages in the similar position of their chromosomes, but the plasmid complement of the two strains varied. Plasmid pBT1 was absent in D37712 but there was a novel plasmids pBT4, which was 80% identical to plasmid pCol1B9 (B. Kumwenda *Pers. Comm.*). To characterize the transcriptome of D37712, an RNA-seq-based transcriptomic analysis was performed with RNA samples from D37712 grown in 15 *in vitro* conditions relevant to *Salmonella* infection (Kröger *et al.*, 2013) but involved no biological replicates. The data have not yet been published, and so have been made available online as SalComD37712, at http://bioinf.gen.tcd.ie/cgi-bin/salcom_v2.pl?db=salcom_D37712_HL.

Subsequently, a second RNA-seq experiment was done that involved three biological replicates for four of the key growth conditions (ESP, InSPI2, NonSPI2 and NoO2: Table 2.5), to generate a statistically-robust differential gene expression comparison between D37712 and D23580 to determine differential gene expressions in D37712. A summary heat map of the analysis of these data is shown in Figure 4.2.

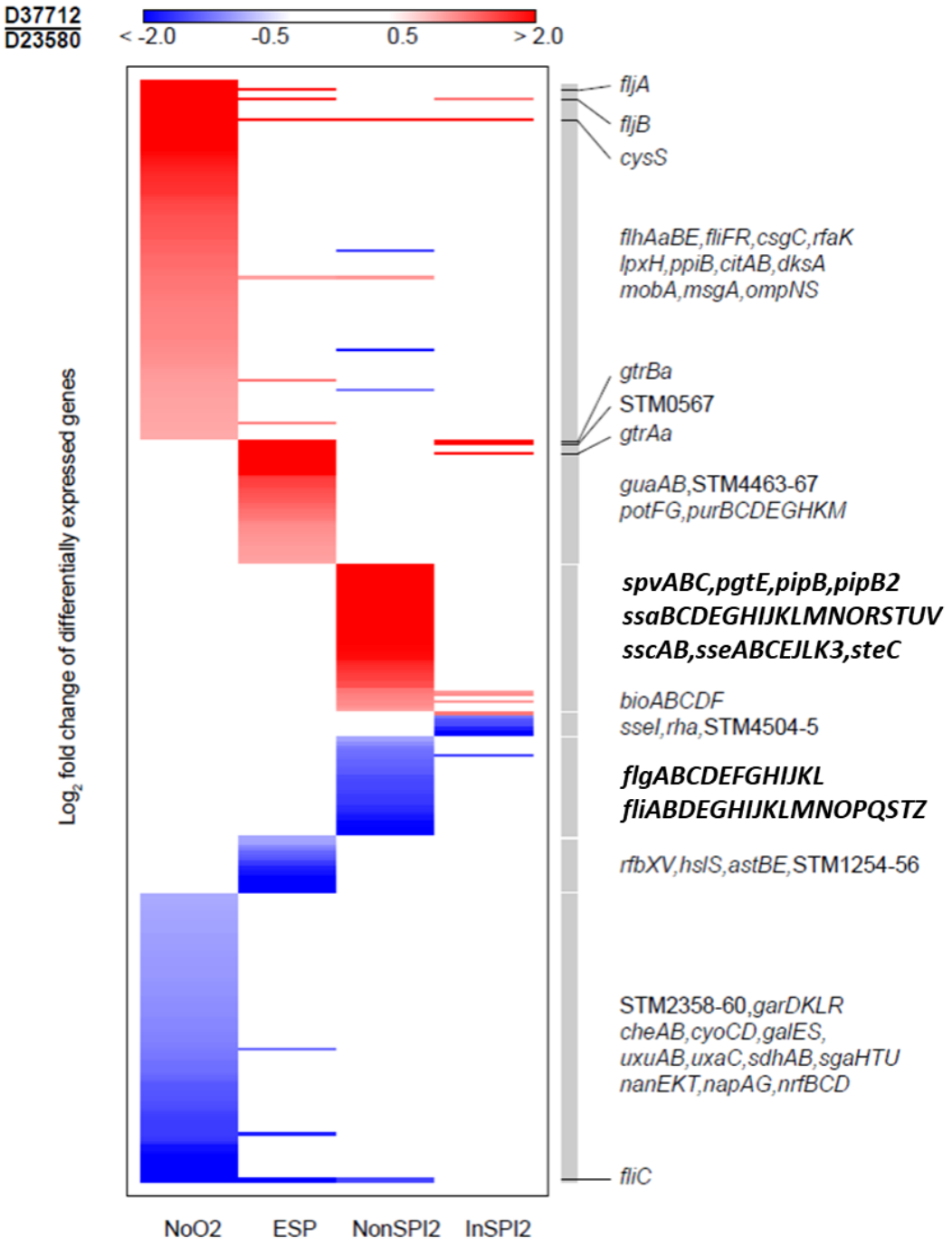


Figure 4.2. Differential gene expression of flagella and SPI2-associated genes in D37712, compared with D23580. Following analysis of RNA-seq data, differentially-expressed genes (FC (Fold Change) > 2-fold, FDR (False Discovery Rate) < 0.05) are shown as a heatmap that compares expression in D37712 and D23580 in four growth conditions (NonSPI2, InSPI2, ESP, NoO2: Table 2.5). Red colour indicates the gene expression level was higher in D37712, while blue colour indicates a lower expression level in D37712. The names of key *S.* Typhimurium genes are shown. The flagella and SPI2 genes are in bold. (A. Predeus *Pres. Comm.*).

Compared to D23580, D37712 showed a lower expression of the flagellar regulon and associated genes, including the *flgA,* the *flgB*, the *fliD* and the *fliF* operons. *S.* Typhimurium produces two forms of flagella filaments made up of either FliC or FljB proteins, which differ slightly in domain D3 (Yamaguchi *et al.*, 2020). The outmost domain D3 is thought to play a role in determining the antigencity of the flagella (Yamaguchi *et al.*, 2020). Both the *fliC* and *fljB* genes were significantly up-regulated in D37712 compared with D23580 during growth at early stationary phase or under anaerobic conditions by between 2-fold and 4-fold. Gene expression heatmaps of *fliC* and *fljB* in D23580 and D37712 are available at the https://tinyurl.com/SalComD37712-fliC-fljB website.

The *cysS* gene was up-regulated in D37712 in all growth conditions. As will be explained in Chapter 5, in the ST313 L2 strain D23580, the chromosomal *cysS* was functionally replaced by a paralogous $cysS^{pBT1}$ gene encoded by plasmid pBT1. Since pBT1 was absent from D37712, the increased expression of *cysS* in D37712 in all conditions made biological sense.

The SPI2 pathogenicity island plays a key role in *Salmonella* infection, and encodes a type III secretion system that is responsible for the translocation of effector proteins into host cells during intracellular infection (Figueria and Holden, 2012). Intriguingly, the expression of SPI2-associated genes was higher than D23580 in NonSPI2 medium, which is a non-inducing growth condition for the SPI2 system (Kröger *et al.*,

2013). It is important to note that the expression level of SPI2-associated genes of D37712 in NonSPI2 media was much lower than that observed during growth under "inducing" conditions in InSPI2 medium, which is low-phosphate and pH 5.8 (Section 2.2). Relevant gene expression heatmaps of SPI2-associated genes are available at https://tinyurl.com/SalComD37712-SPI2 website. It is apparent that SPI2 genes are more highly expressed in D37712 than D23580 in the InSPI2 "inducing" media. However, the intra-macrophage transcriptomic data revealed that SPI2 genes were down-regulated by between 2 to 3-fold during macrophage infection.

The expression of SPI2 genes in the non-inducing NonSPI2 media for strain D37712 is readily apparent from the mapped sequence reads of the RNA-seq experiment for the strain D37712, D23580 and 4/74. A version of JBrowse focused on the SPI2 region is shown in Figure 4.3.

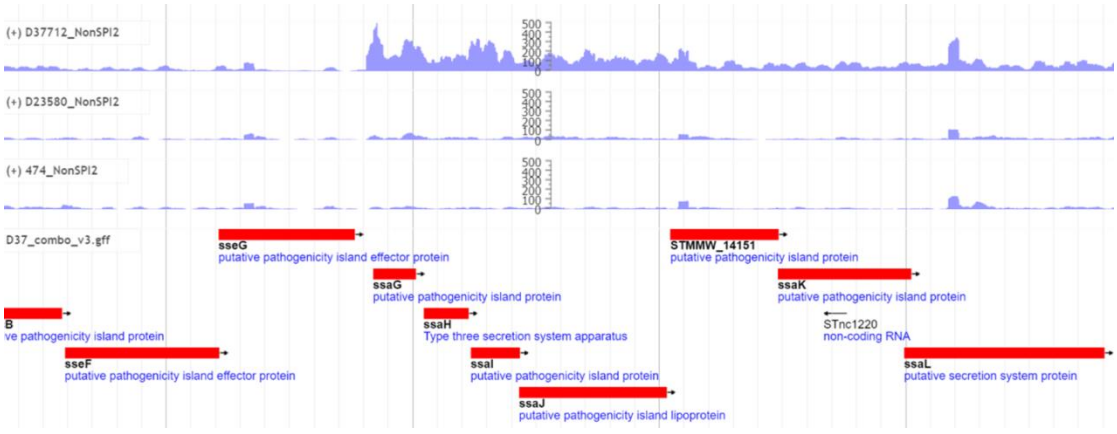


Figure 4.3. RNA-seq data for the SPI2 region of 4/74, D23580 and D37712 in NonSPI2 condition (B. Kumwenda *Pers. Comm.*). The data are available at https://tinyurl.com/SalComD37712-Jbrowse.

To investigate the expression of SPI2 genes in ST313 strain L2.2 D37712, it was important to confirm aberrant expression of the SPI2 system in NonSPI2 media using an independent experimental approach, as described below in Section 4.3.

In this chapter, experimental strategies were developed to investigate certain phenotypic properties of D37712. It was hoped that these experiments might shed light on the reason that ST313 L2.2 has largely replaced ST313 L2 (Figure 4.1A) as the major cause of human bloodstream infections in Malawi in recent years.

4.1.1 Acknowledgement of the contributors to the results in this chapter

I am grateful for the contributions of Ben Kumwenda and other colleagues for their transcriptomic analysis of ST313 L2.2. Specifically, the first 15-condition RNA-seq experiment, which involved no biological replicates, was done by Ben Kumwenda and Carsten Kröger. The second 4-condition RNA-seq experiment involved biological triplicates samples and was done by Ben Kumwenda and Rocío Canals. The RNA-seq data were analysed by Alexander Predeus.

4.2 Identification of phenotypic differences that distinguish D37712 from D23580

The RNA-seq data showed that 364 genes were differentially expressed between D37712 and D23580, in atleast one growth condition (A. Predeus *Pres. Comm.*). A series of experiments were conducted to determine whether the differential

expression of certain genes at the transcriptional level was reflected by an associated phenotypic alteration that distinguished ST313 L2.2 strain D37712 from ST313 L2 strain D23580.

4.2.1 Swimming motility phenotype

Flagellin activates the inflammasome component of the innate immune response, which plays an important role in the host response to *S.* Typhimrium infection (Clare, 2021). Previous work in the Hinton Lab established that the flagellar regulon was expressed at a lower level in *S.* Typhimurium ST313 than it in ST19 strains, including the Class 1 (*flhCD*), Class 2 (e.g. *fliA*, *flgM*) and Class 3 genes (e.g. *fliC*, *fljB*). Phenotypically, the ST313 strain D23580 exhibited a reduction of swimming motility compared with ST19 (Canals *et al.*, 2019b). Although the mechanistic basis of the relatively low level of expression of the flagellar system in ST313 strain D23580 has not been definitively established, there has been speculation that the RflP (YdiV) regulatory protein plays a key role (Canals *et al.*, 2019b). RflP is the EAL-domain protein that is responsible for the bistable expression of flagella, and responds to both cell envelope stress and changes in outer membrane integrity (Spöring *et al.*, 2018).

Because the transcriptomic data suggested that, in NonSPI2 media, the flagellar regulon was expressed at a lower level in D37712 than in D23580 (Section 4.1), swimming motility assays were performed. As well as studying ST313 strains

D23580 and D37712, the ST19 strain 4/74 was also included and this strain has been reported to have a higher level of swimming motility than D23580 (Canals *et al.*, 2019b).

Figure 4.4 shows that D37712 had a significantly decreased level of swimming motility than D23580 and 4/74 in NonSPI2 media. The levels of swimming motility were in the following order: 4/74 > D23580 > D37712. These results were consistent with the RNA-seq data summarised in Figure 4.2.

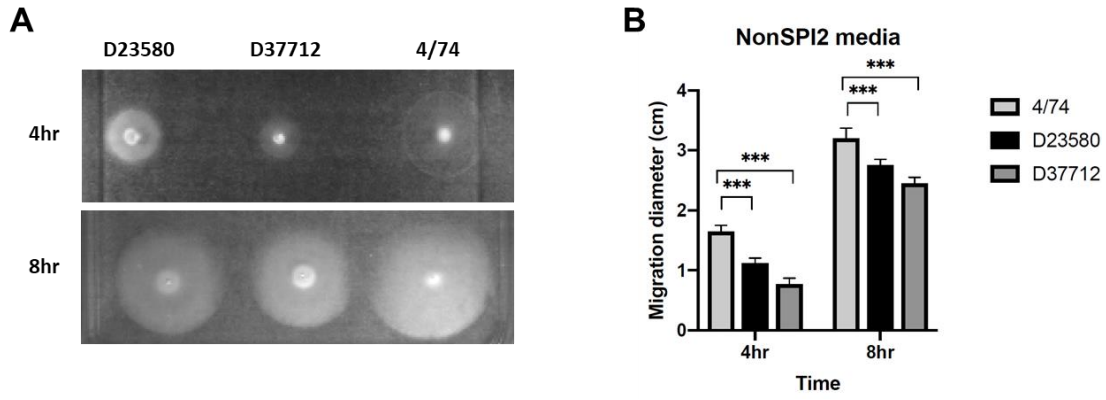


Figure 4.4. Lower swimming motility of D37712 than D23580 in NonSPI2 media. (A) A representative NonSPI2 motility assay plate is shown for strains 4/74, D23580 and D37712 after both 4 and 8 hours incubation (Section 2.4.3). (B) Mean migration diameter of three replicates, n = 3, ***: P<0.001. Error bars show Standard Deviation.

4.2.2 RDAR morphotype

It has been hypothesised that biofilm formation is important for transission of *S.* Typhimurium to mammalian hosts (MacKenzie *et al.*, 2019; White *et al.*, 2006). One of the most well-studied biofilm-associated phenotype of *Salmonella* is the RDAR (red, dry and rough) morphotype that involves the extracellular matrix on the surface of bacterial colonies (Romling *et al.*, 1998). RDAR-positive macrocolonies have

concentric rings and a wrinkled appearance on the surface of Congo red agar plates.

Previous studies determined that a premature stop codon was responsible for *bcsG* (*yhjU*) being an inactive pseudogene in ST313 L2. As BcsG is a cellulose biosynthetic enzyme involved in biofilm formation, the pseudogene *bcsG* resulted in ST313 L2 having a RDAR-negative phenotype; in contrast, ST19 strains 4/74 and SL1344 are RDAR-positive (MacKenzie *et al.*, 2019; Pulford *et al.*, 2021; Singletary *et al.*, 2016).

To determine the RDAR status of ST313 L2.2, the abilities of D23580, D37712 and 4/74 to form biofilms were assessed, using both the RDAR phenotypic assay and a biofilm-inducing 1% tryptone agar plate (MacKenzie *et al.*, 2019; Figure 4.5). D37712 had a biofilm negative phenotype in both assay conditions as did strain D23580. The results are consistent with the comparative genomic analysis between D23580 and D37712, which showed that *bcsG* carried the same pseudogene mutation in the two strains (B. Kumwenda *Pers. Comm.*).

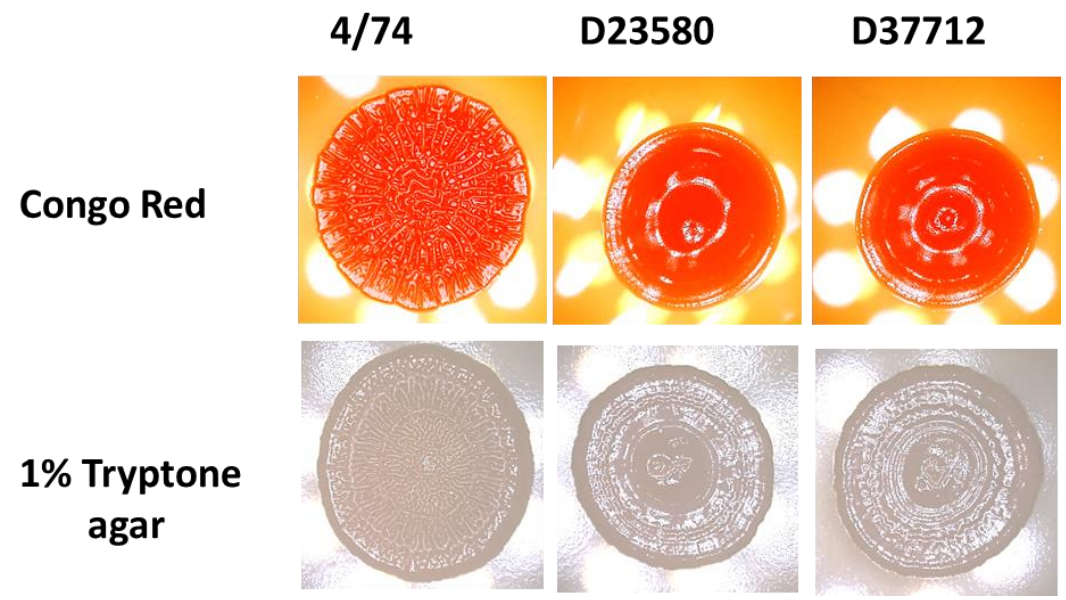


Figure 4.5. Both ST313 L2 and ST313 L2.2 strains have a RDAR-negative phenotype. The top panel shows the RDAR phenotypic test (Section 2.4.4) and the bottom panel shows a complementary experiment that involves the induction of biofilm formation on 1% tryptone agar (MacKenzie *et al.*, 2019). RDAR-positive macrocolonies have concentric rings and a wrinkled appearance (Romling et al., 1998). Strain 4/74 was RDAR-positive in both conditions, whereas D23580 and D37712 were RDAR-negative.

4.3 Expression of SPI2 genes in D37712 in Non-SPI2 inducing condition

As described in Section 4.1, comparative transcriptomic analyses highlighted a key difference between D23580 and D37712: the increased level of expression of SPI2-associated genes of strain D37712 in NonSPI2 media (Figure 4.2; B. Kumwenda *Pers. Comm.*). In *S.* Typhimurium ST19, SPI2 genes are normally induced during growth in InSPI2 media, a minimal medium that contains low levels of phosphate at pH 5.8 (Löber *et al.*, 2006). In contrast, the high level of phosphate and neutral pH (7.4) of NonSPI2 media does not stimulate SPI2 gene expression of ST19 (Kröger *et al.*, 2013) or ST313 strain D23580 (Canals *et al.*, 2019b).

To confirm whether differential expression of SPI2-associated genes was observed in strain D37712 at a single cell level, a transcriptional fusion-based approach was used involving the GFP$^+$ protein, a highly fluorescent version of GFP (Scholz *et al.*, 2000). The SsaG protein, (also known as SctF) is encoded by the SPI2 pathogenicity island, and is the key component of the needle of the SPI2 type III system apparatus (Yu *et al.*, 2018). Previous studies showed that the *ssaG* promoter was induced dramatically in both the intracellular environment and in InSPI2 media (Hautefort *et al.*, 2003; Lim *et al.*, 2006). a genetic strategy was used to insert the *gfp*$^+$ gene into chromosome of D37712 and an antibiotic-sensitive derivative D23580 (strain SSS18; Table 2.2) downstream of the *ssaG* gene promoter (Figure 4.7A).

The expression of GFP$^+$ in the D23580 and D37712 *ssaG-gfp*$^+$ derivatives (SZS008 and SZS032, Table 2.2) was firstly assessed following overnight growth in both InSPI2 and NonSPI2 conditions by fluorescent microscopy (Figure 4.6). Both D23580 and D37712 showed a high level of GFP$^+$ expression in InSPI2 media. In NonSPI2 media, a small proportion of cells of both strains exhibited fluorescence. The fluorescence intensity was lower than that was induced by growth in the InSPI2 media.

Figure 4.6 confirmed that the *ssaG-gfp*$^+$ transcriptional fusion showed the expected fluorescence induction in InSPI2 media versus NonSPI2 media. SPI1 promoters show bistable expression during *in vitro* growth of *S.* Typhimurium (Ackermann *et al.*,

2008), but the SPI2 system has not been reported to share this property. To investigate the possibility of bistablity, the level of *ssaG* expression was quantified at a single cell level.

To quantify the levels of GFP^+ fluorescence, flow cytometry was used to measure expression of *ssaG-gfp*$^+$ in both the D23580 and D37712 genetic background. The average fluorescence level (Figure 4.7B) and the proportion of fluorescent cells in each sample were assessed (Figure 4.7C). D37712 showed a significantly higher level of average GFP^+ expression than D23580 ($P<0.0001$), and a slightly higher percentage of GFP-positive cells ($P<0.001$). The higher proportion of GFP positive cells in D37712 was not sufficient to account for the difference between the mean fluorescence levels of the two strains.

The overall conclusion from Figure 4.6 and 4.7 was that the SPI2-associated *ssaG* promoter was expressed at a significantly higher level in NonSPI2 media in ST313 L2.2 strain D37712 than in ST313 L2 strain D23580.

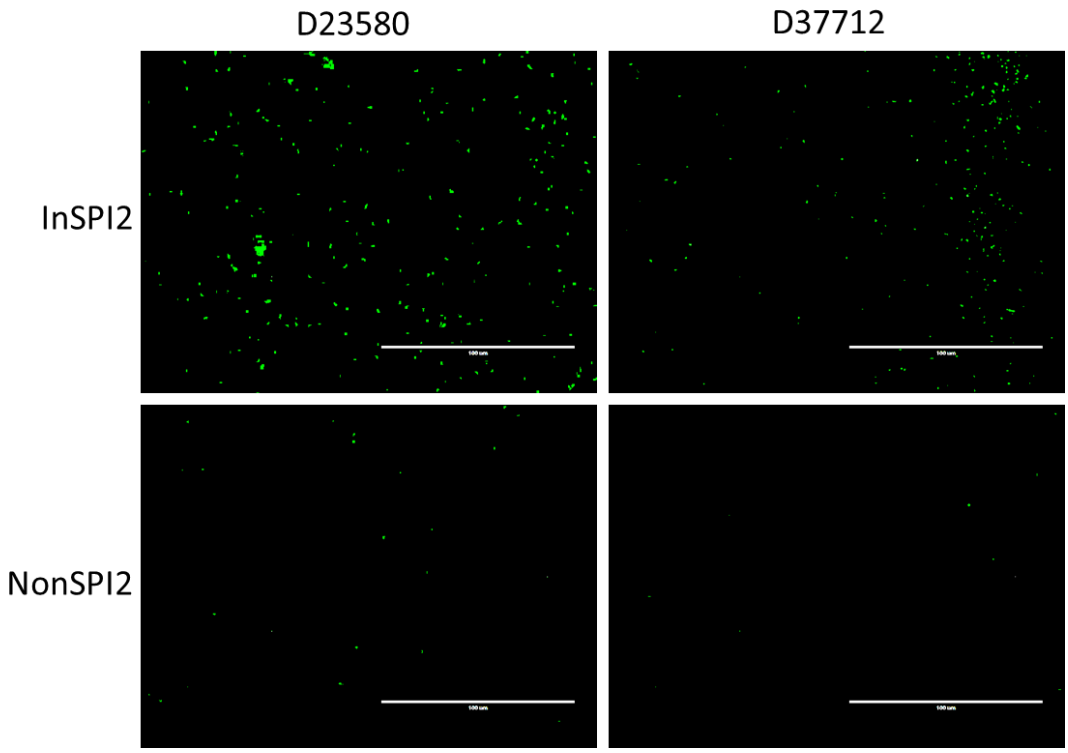


Figure 4.6. Micrograph showing that D23580 and D37712 bacteria carrying the *ssaG-gfp*⁺ transcriptional fusion had a high level of fluorescence in InSPI2 media. Overnight cultures of D23580 and D37712 *ssaG-gfp*⁺ strains (SZS008 and SZS032, Table 2.2) in SPI2 and NonSPI2 medium were diluted to same OD_{600} and then 1 µL was dropped onto the glass slide (Section 2.4.5). Low numbers of individual SPI2-expressing bacteria are visible under NonSPI2 conditions. Image was taken using an EVOS FL cell imaging system (Thermo Fisher) following excitation with blue light. The scale bar (white) is 100 µm (Section 2.4.5).

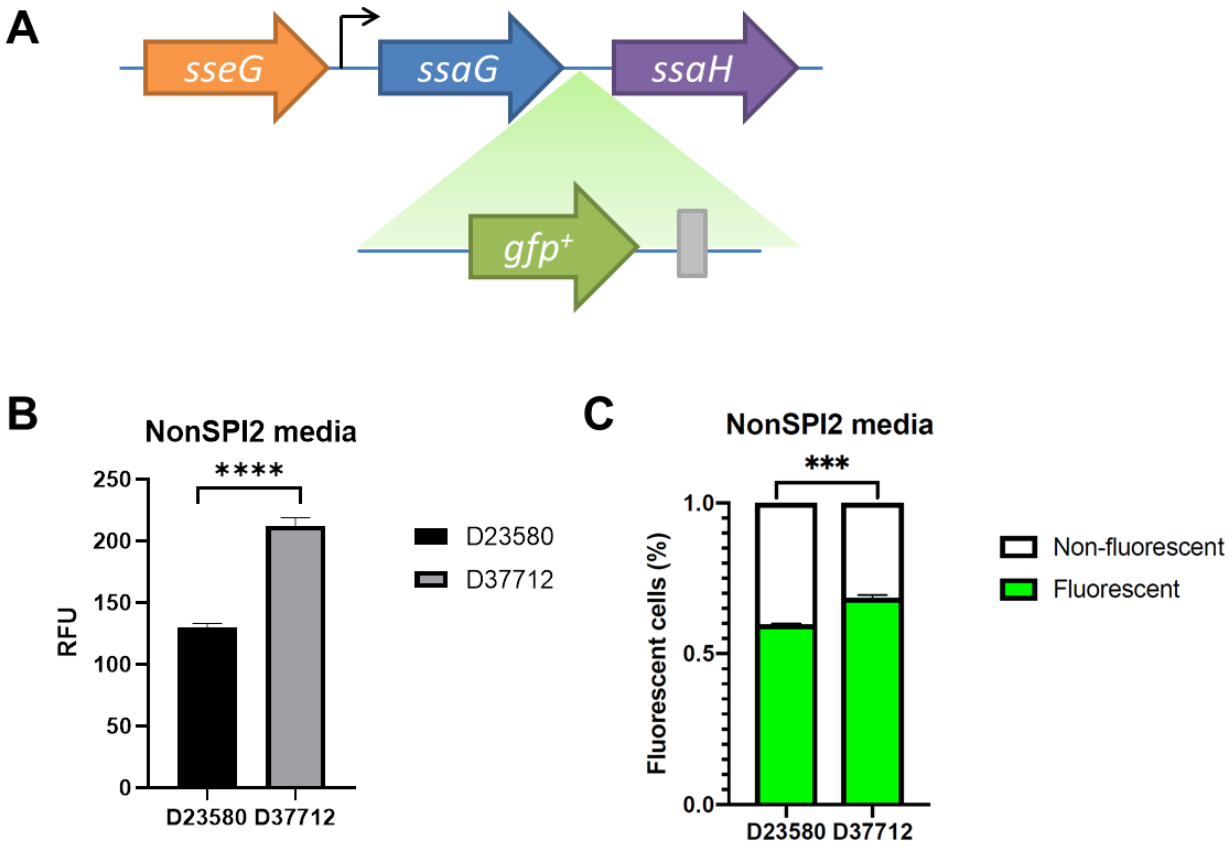
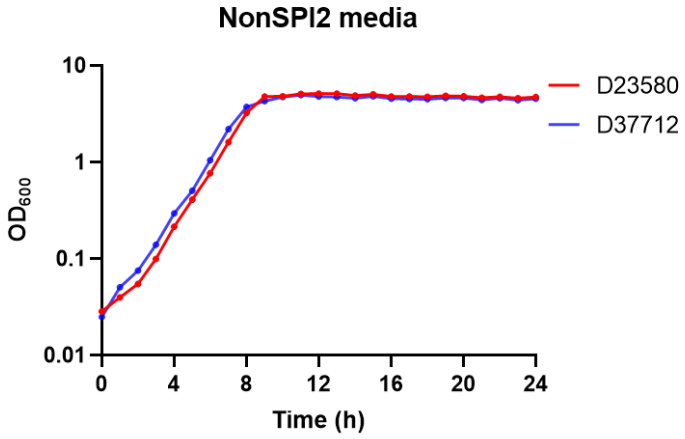


Figure 4.7. The *ssaG* promoter was more highly expressed in D37712 than in D23580 in NonSPI2 media. (A) Genetic context of the GFP$^+$ transcriptional fusion under the control of *ssaG* promoter. Promoter of *ssaG* is shown as a black arrow. Grey bar indicates the 84 nucleotide scar sequence *frt* of pKD4. (B) Mean fluorescence intensity signal of *ssaG*-GFP$^+$ for D23580 (dark grey; SZS008, Table 2.2) and D37712 (grey; SZS032, Table 2.2). RFU = Relative Fluorescence Units. n = 3, ****: P<0.0001 (Section 2.4.6). (C) Percentage of fluorescent-positive (green) and fluorescent-negative cells (white) for *ssaG-gfp*$^+$ expression was determined for each strain by flow cytometry for each sample, n = 3, ***: P<0.001 (Section 2.4.6). Error bars show Standard Deviation.

4.4 Increased fitness of D37712 compared with D23580 in NonSPI2 media

Section 4.3 shows that D37712 had a higher level of expression of SPI2-associated genes in NonSPI2 media than D23580. Previously, it has been reported that increased SPI2 gene expression caused a significant reduction in growth rate of *S.* Typhimurium (Lucchini *et al.*, 2006). Accordingly, the hypothesis was that the higher level of SPI2 gene expression would impose a significant fitness cost upon D37712

during growth in NonSPI2 media. Bacterial fitness is the ability for bacteria to survive and to produce more offspring in a competition environment (Wiser and Lenski, 2005). One simple approach to investigate whether there was a fitness cost to D37712 in NonSPI2 media was to compare its growth rate with D23580 (Santiago *et al.*, 2020). Growth curves for D23580 and D37712 in NonSPI2 media were measured (Figure 4.8). The results showed no significant difference between the growth rates and the final yield of the two strains.



S. Typhimurium Strains	Max Growth Rate	T-test
D23580	0.615	ns (P > 0.05)
D37712	0.608	

Figure 4.8. *S.* Typhimurium strains D37712 and D23580 had similar growth rates in NonSPI2 media. Growth curves of D23580 and D37712 in NonSPI2 medium at 37°C, n = 6 (Section 2.4.2). The table below shows max growth rate of each strains. No significant difference was found between the two strains (t-test: P > 0.05).

An accurate way to assess bacterial fitness defects involves competitive growth assays in liquid media (Hibbing *et al.*, 2010). To investigate the fitness of D37712 in competition with D23580, kanamycin-resistant derivates of D23580 and D37712 were constructed by introducing the *aph* resistance gene into the intergenic region

between *STM4196* and *STM4197* using Lambda Red recombination (Figure 4.9A). In the competition index assay (Section 2.4.7), in each pair, one strain was kanamycin sensitive and the other one was kanamycin resistant. A competitive index of greater than 1 indicates the increased fitness of the kanamycin-resistant strain. All the pairwise combination were studied in LB, InSPI2 and NonSPI2 conditions at OD_{600} = 0.3 (Figure 4.9B). Because antibiotic is one of the factors causing fitness cost (Melnyk *et al.*, 2015). the competition experiments firstly compared kanamycin-resistant and kanamycin-sensitive versions of the same strains (Figure 4.9B). Results showed that the competitive indices of the pairs of strains with the same genetic background were around 1, proving that the insertion of kanamycin resistance gene had no impact on fitness.

Further competition experiments revealed that the fitness of strains D37712 and D23580 was similar in LB media, and identified that D37712 had a selective advantage over D23580 in both InSPI2 and NonSPI2 media. The D37712/D23580 competitive index was 1.79 and in InSPI2 media, and the fitness difference was most apparent in NonSPI2 media, with a D37712/D23580 competitive index of 2.20. The result shows that D37712 has a significant fitness advantage over D23580 during growth in nutrient-limited minimal media. The greatest fitness of strain D37712 was observed under the growth conditions that did not induce a high level of expression of SPI2 genes.

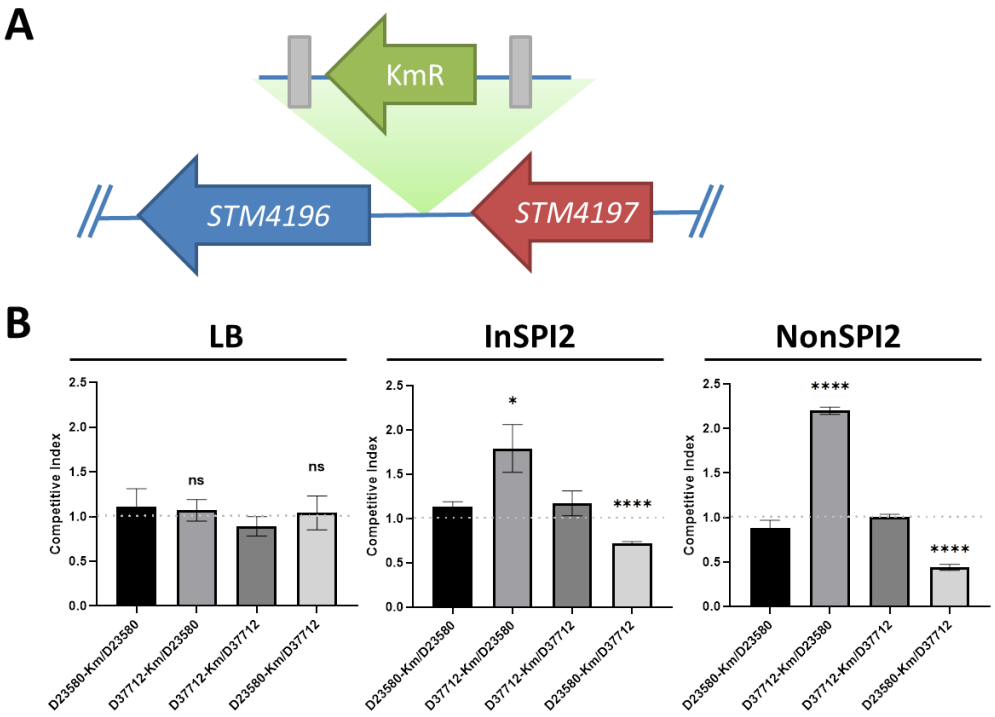
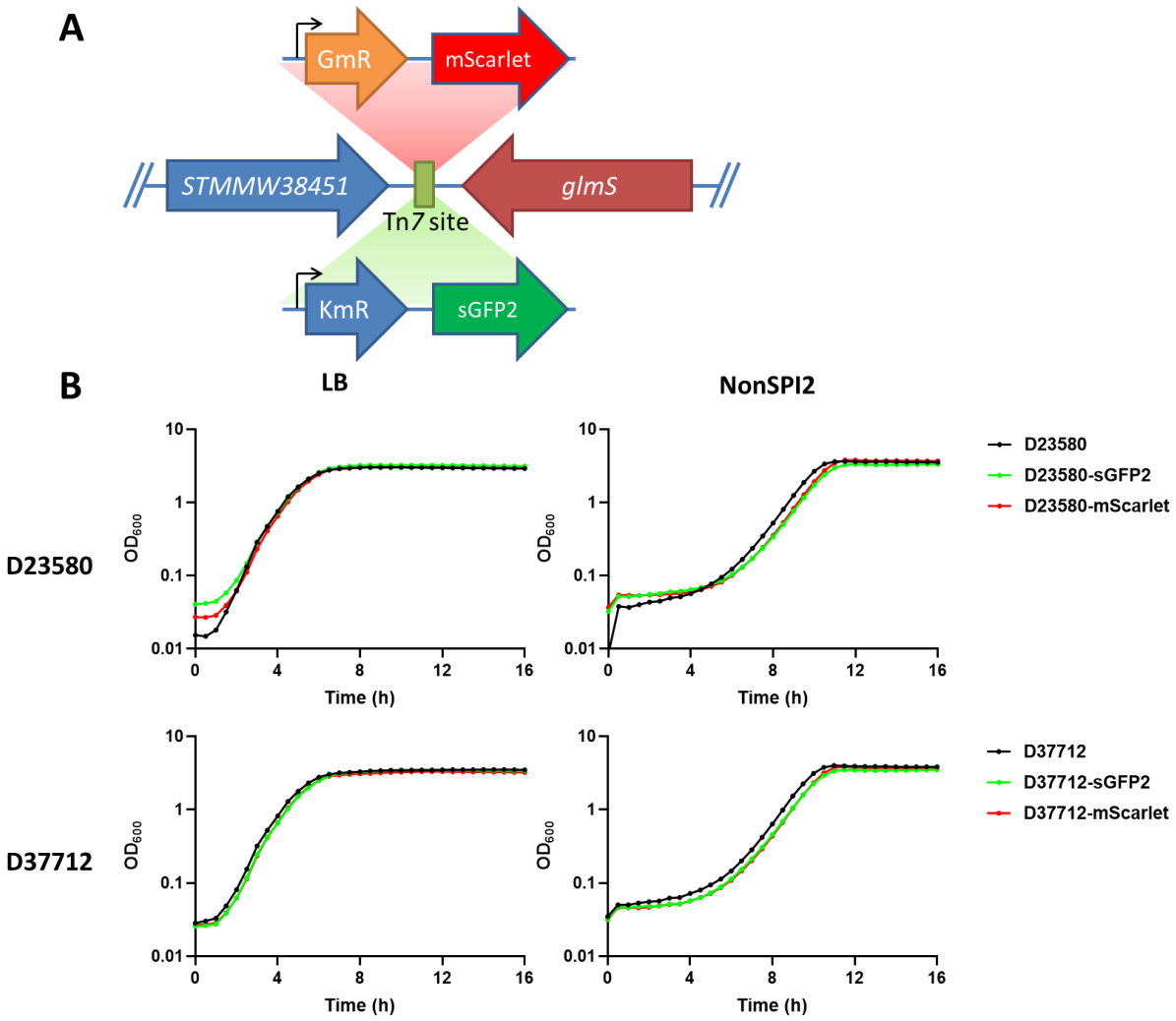


Figure 4.9. D37712 had a higher competitive index than D23580 in both InSPI2 and NonSPI2 conditions. (A) Genetic context of the kanamycin insertion site in the chromosome of D23580 and D37712. Arrows represent a single gene. The grey bar indicates the 84 nucleotide *frt* "scar" sequence of pKD4. (B) Competitive index was determined by CFU counting for each pair-wise combination of antibiotic-resistant versus WT derivatives of D23580 and D37712 following overnight growth in LB, InSPI2 and NonSPI2 media, n = 3. ****: P<0.0001; *: P<0.05; ns: not significant. Error bars show Standard Deviation. A competitive index of 1 indicates the equivalent fitness of two strains, while a number higher than 1 showed that the kanamycin-resistant derivatives had a selective advantage (Section 2.4.7).

To confirm the important finding that D37712 had a fitness advantage over D23580 in NonSPI2 media (Figure 4.9), an independent approach involving fluorescently-tagged *S.* Typhimurium strains was used. An alternate Tn*7*-based genetic approach was used to insert gene cassettes into a different chromosomal region of *S.* Typhimurium that was located >2 Mb from SPI2, based on the ability of Tn*7* to insert into a single highly conserved site located downstream of gene *glmS*. The *glms* gene encodes an essential glucosamine-fructose-6-phosphate

aminotransferase. GmR-mScarlet or Km^R-sGFP2 cassettes were introduced into the chromosomal Tn*7* insertion sites of D23580 and D37712 which are both located between the gene *STMMW38451* and *glmS* (based on annotation of D23580; Figure 4.10A).

To test whether the insertion of the Gm^R-mScarlet or Km^R-sGFP2 cassettes impacted on the fitness of D23580 and D37712, growth curves were first determined in LB or NonSPI2 medium (Figure 4.10B). Though the insertion influenced the fitness of D23580 and D37712 in some of the conditions, the insetion of the Gm^R-mScarlet and Km^R-sGFP2 cassettes had similar effects on D23580 and D37712 (Figure 4.10C), therefore the cassettes were still used for following competition experiments.



S. Typhimurium Strains	Max Growth Rate in LB	T-test	Max Growth Rate in NonSPI2	T-test
D23580	0.212		0.355	
D23580-SGFP2	0.237	ns (P > 0.05)	0.291	ns (P > 0.05)
D23580-mScarlet	0.249		0.308	
D37712	0.211		0.356	
D37712-SGFP2	0.232	ns (P > 0.05)	0.336	ns (P > 0.05)
D37712-mScarlet	0.221		0.339	

Figure 4.10. Chromosomal integration of the two fluorescence gene cassettes had similar impacts on growth rate of D37712 and D23580. (A) Km^R-sGFP2 and Gm^R-mScarlet were inserted into the transposon Tn7 site of D23580 or D37712. Small black arrows represent TSS and coloured directional arrows represent genes. (B) Growth curves of WT and sGFP2 or mScarlet-tagged derivatives of D23580 and D37712 in LB and NonSPI2 medium at 37°C, n = 6 (Section 2.4.2). The table below shows max growth rate of each strain in LB and NonSPI2 medium. No significant difference was found between the two mutations of each strain (t-test: P > 0.05).

The fluorescently-tagged strains were used for competition experiments. Pairwise combinations of strains D23580 and D37712 were grown together in NonSPI2 media and the green and red fluorescence levels were measured using flow cytometry. A quadrant of the flow cytometry data is shown in Figure 4.11A. Both combinations of the two strains showed a higher proportion of D37712 than D23580 bacteria (Figure 4.11B), supporting the conclusion that strain D37712 had a fitness advantage over D23580 in the NonSPI2 environment.

However, Figure 4.11A showed a limitation of the fluorescence-based stragegy, as visualised with the BD FACSAria Flow Cytometer, which involved the presence of diffuse blue dots in both the upper right ($sGFP2^{+}$, $mScarlet^{+}$) and bottom left ($sGFP2^{-}$, $mScarlet^{-}$) regions of the quadrant figure. These blue dots might represent background noise of the flow cytometer. To independently verify the fluorescence-based approach, considering advantage of the fact that the two gene cassettes encoded distinct fluorescent proteins and different antibiotic resistances (Gm^{R}-mScarlet and Km^{R}-sGFP2), the competition experiment was repeated and CFUs were counted on agar plate supplemented with different antibiotics. Both combinations showed that strain D37712 had a competitive index > 1 (Figure 4.11C).

In summary, the three approaches showed that the increased expression of SPI2-associated genes did not have a fitness cost for D37712. In fact, *S.* Typhimurium ST313 L2.2 strain D37712 had a clear competitive advantage over

ST313 L2 strain D23580 that was most apparent in NonSPI2 media.

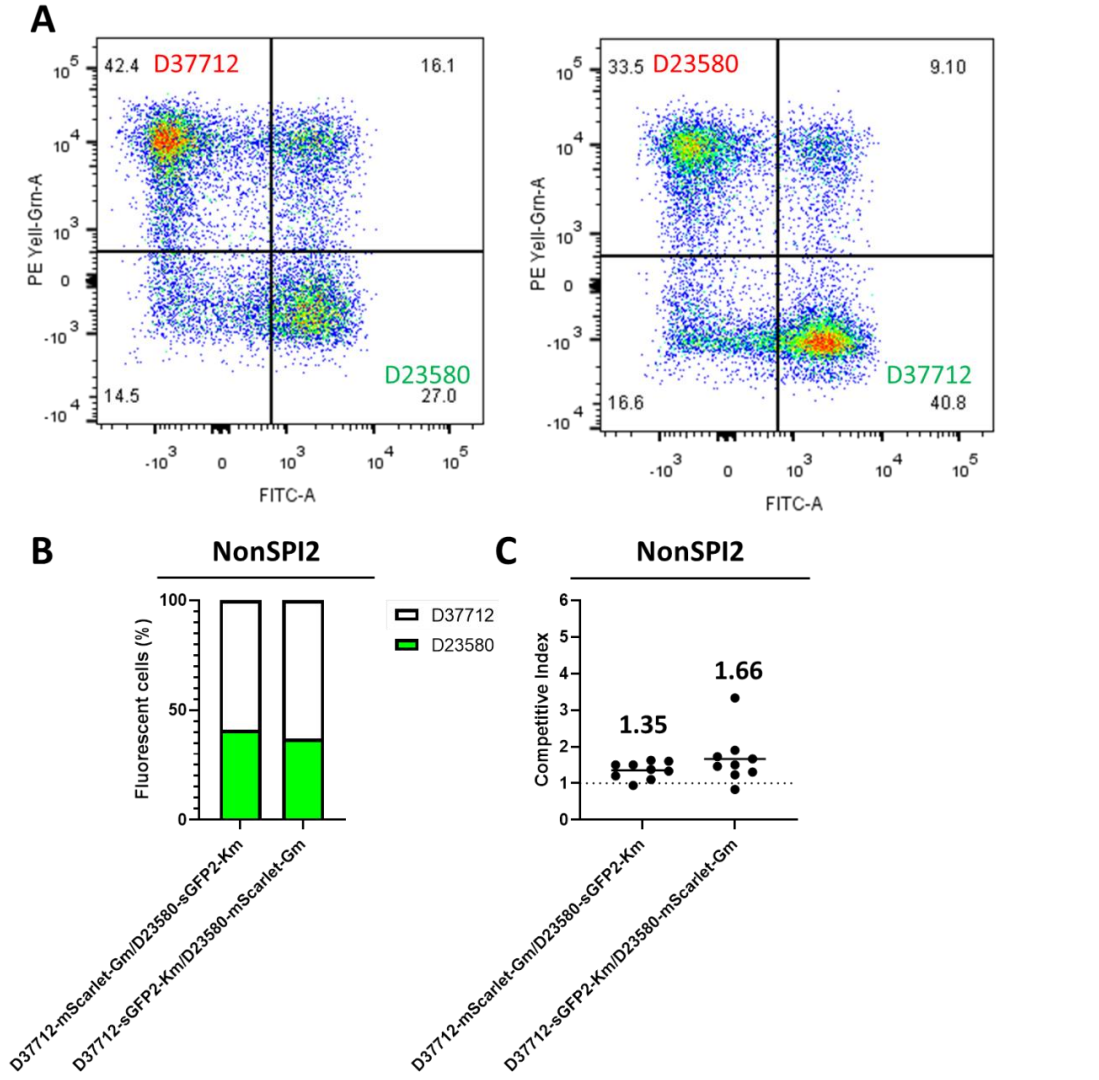


Figure 4.11. Flow cytometric-based competitive index assay showing that D37712 had a selective advantage over D23580 in NonSPI2 conditions. (A) Raw flow cytometric data of pair-wise combinations of fluorescently-tagged derivatives of D23580 and D37712 in a quadrant figure. Sample was taken from a 1:1 mixed culture after incubation in NonSPI2 media to OD_{600} = 0.3. X-axis (labelled FITC) shows the GFP level and the Y-axis (labelled PE Yell-Grn) indicates the mScarlet level. The black number in each region indicates percentage of events in each region. In total, 10,000 events were acquired for each sample (Section 2.4.6). Colour of dots reflects the number bacteria with particular fluorescence characteristics: Red > Green > Blue. (B) Percentage of D37712 (white) and D23580 (Grey) from each experiment shown in Panel A. (C) Competitive index was determined by CFU counting using the antibiotic-resistance phenotype for each pair-wise combination of D23580 and D37712 on LB Km or Gm plates from overnight culture in NonSPI2. The data represents biological replicates with individual data points shown; the group mean is represented by the horizontal line. A competitive index number over 1 reflects increased fitness of D37712 compared with D23580 (Section 2.4.7).

4.5 Discussion

The clonal replacement of *S.* Typhimurium ST313 L2 by the novel sublineage L2.2 in Malawi in recent years was worthy of further investigation. This chapter describes a series of experiments that move beyond comparative genomics to focus on phenotypic properties. Differential gene expression analysis between L2.2 representative strain D37712 and L2 strain D23580 showed key differences in the neutral-pH NoSPI2 growth condition. In comparison with D23580, flagellar genes were down-regulated in D37712 and the SPI2-associated genes were up-regulated (Figure 4.2).

My experiments validated the transcriptomic data by identifying two different phenotypes that clearly distinguished the strains D37712 and D23580.

The functional significance of the reduced level of motility is worthy of consideration. It is possible that reduced flagella expression by ST313 L2.2 could lead to an increased “stealth phenotype” of ST313 during infection. This type of immune system evasion has previously been reported for other entroinvasive pathogens (Toslis *et al.*, 2008) and a flagellar expression could help L2.2 to escape from the host innate immune response (Hayashi *et al.*, 2001; Ilyas *et al.*, 2018).

Previously, it has been reported that *S.* Typhimurium ST313 strain D23580 caused lower levels of inflammasome activation during infection of murine and human

macrophages than *S.* Typhimurium ST19 (Carden *et al.*, 2015), as well as lower levels of apoptosis and production of proinflammatory cytokines (Ramachandran *et al.*, 2015). Future research should compare the level of inflammasome activation by strains D37712 and D23580 during cellular infection.

Since the discovery of SPI2 in 1995 (Hensel *et al.*, 1995), the transcriptional regulation of SPI2 genes has been studied extensively. Previous studies showed that SPI2 genes in *S.* Typhimurium are regulated by a complex system, which is activated by the transcriptional regulators HilD and SlyA (Fass and Groisman, 2009) and silenced via H-NS (Lucchini, *et al.*, 2006).

In this chapter, the flow cytometry data showed that strain D37712 had a higher level of SPI2 gene expression and a higher proportion of SPI2-positive cells compared with D23580 on the NonSPI2 growth condition (Figure 4.7). The results raised the possibility that either the activation or the silencing process of SPI2-associated genes had been disturbed in D37712 during growth in NonSPI2 media.

Coupled with the RNA-seq analysis of D37712 (Section 4.1), the findings raise the posibility that SPI2-associated genes can be activated or de-repressed by unknown regulatory factors in D37712 genome. The SsrAB two-component system activates SPI2 transcription, and down-regulates flagella and swimming motility in *S.* Typhimurium ST19 (Brown *et al.*, 2014). Further experiments to investigate the

expression of other SPI2 regulatory genes such as *ssrA* and *ssrB* at the single cell level would be worthwhile, as the RNA-seq data suggested that *ssrAB* were slightly more highly expressed in NonSPI2 media (available at SalComD37712_SPI2 website). It is conceivable that SsrAB could have opposing influences that are responsible for the differential expression of the flagella and SPI2 regulons in D37712 described in this chapter.

Recent studies have compared the virulence of D37712 and D23580. Both of the two strains were able to survive during human serum-killing (Hammarlöf *et al.*, 2018) and escape from the activation of mucosal invariant T (MAIT) cells (Preciado-Llanes *et al.*, 2020), and they both induced expression of the immunoregulatory cytokine IL-6 and IL-10 (Aulicino *et al.*, 2022). These three studies suggest that the pathogenesis of ST313 L2.2 strain D37712 does not exhibit a lineage-specific virulence phenotype, thus, the aberrant expression of SPI2 genes identified in this chapter does not have a direct impact upon infection biology.

The lack of a lineage-specific virulence phenotype promoted hypothesis that strains D23580 and D37712 differs in terms of fitness. One possibility was that the aberrant expression of the membrane-located SPI2 secretion machinery imposed a significant fitness cost upon D37712 that was not seen for D23580 during growth in NonSPI2 media.

The fitness experiments with three different approaches showed that D37712 had an increased competitive index in NonSPI2 media (Figure 4.10 and 4.11), which disproved the hypothesis above. One explanation for the fitness of D37712 could be the differences in the accessory genome between D23580 and D37712. Plasmid pBT1 is absent from D37712, and the strain carries the novel pBT4 plasmid. The plasmid pBT1 in D23580 plays a physiologically relevant role by encoding an alternate aminoacyl-tRNA synthetase that functionally replaces the chromosomal CysS (Canals *et al.*, 2019a). It is possible that the decreased fitness of D23580 reflects the involvement of a non-canonical cysteinyl-tRNA synthetase. Further, pBT1 also represses transcription of the chromosomal *cysS* gene of D23580. The pBT1-mediated repression of the chromosomal *cysS* gene will be investigated in Chapter 5.

The L2.2-specific plasmid pBT4 shared 80% nucleotide sequence identity with pCol1B9 (B. Kumwenda *Pers. Comm.*) and carried the *impC-umuCD* operon which is responsible for an increased mutation rate when *E. coli* cells are exposed to DNA damage (Sikand et al., 2021). However, no other pBT4-encoded genes have been functionally characterised. Future competition experiments could assess a D37712 derivative that lacked pBT4, or investigate the effect of moving pBT1 into D37712. Both experiments might reveal a role of these plasmids in the modulation of bacterial fitness. Future studies of the fitness difference between ST313 L2 and L2.2 may shed more light on mechanisms that underpin the success and evolution of *S.*

Typhimurium ST313 in Africa.

Chapter 5.
The search for a plasmid-encoded repressor of the chromosomal *cysS* gene in D23580

5.1 Introduction

S. Typhimurium ST313 causes iNTS diseases in humans, while the globally distributed *S.* Typhimurium ST19 causes self-limiting gastroenteritis. An in-depth genomic comparison between strain D23580 (the representative strain for ST313) and strain 4/74 (the representative strain for ST19), showed that 95% of their genomes were identical (Canals *et al.*, 2019b).

A further transcriptomic comparison between D23580 and 4/74 involved RNA isolated from 16 infection-related in vitro conditions (Kröger *et al.*, 2013) and from infected murine macrophage RAW264.7 (Canals *et al.*, 2019b) to investigate the differences in the level of transcription of the shared orthologous genes. The transcriptomic data identified six genes that were down-regulated in D23580 in most growth conditions compared to strain 4/74. The most unexpected one was the *cysS* chromosomal gene (*cysS*^{chr}) which is an essential gene in 4/74 and other *Salmonella* strains (Baba *et al.*, 2006).

It is known that *cysS* is not required for viability of the methanogenic archaea *Methanocaldococcus jannaschii* (Ruan *et al.*, 2004; Stathopoulos *et al.*, 2001). In contrast, the cysteinyl-tRNA synthetase (CysRS) encoded by *cysS*^{chr} is essential for growth of many organisms (Baba *et al.*, 2006; Barquist *et al.*, 2013). However, a transposon-insertion-sequencing (TIS) experiment revealed that the *cysS*^{chr} gene of D23580 was not required for the survival for bacteria in rich medium (Canals *et al.*,

2019a).

The aim of this chapter is to explore why the *cysS* gene is down-regulated in D23580 and to investigate whether there are factors on the accessory genome of D23580 that repress the expression of the chromosomal *cysS*.

5.1.1 Acknowledgement of the contributions to the results in this chapter

Results in Section 5.2 in this chapter were published in an article I co-authored (Canals *et al.*, 2019b). I acknowledge the contributions of my colleague Rocío Canals for generating the RNA-seq data shown in Figure 5.1, and the evolution experiment 1 described in Section 5.5.1.

5.2 The plasmid pBT1-encoded *cysS*^{pBT1} gene is a paralog of the chromosomal *cysS* in D23580

The nucleotide sequence of the *cysS*^{chr} gene in D23580 chromosome is identical to strain 4/74, but the gene is dramatically down-regulated at the transcriptional level. By searching all the annotated genes in the genome of D23580, we identified an additional plasmid pBT1-encoded gene *pBT1-0241* that encodes a cysteinyl-tRNA synthetase (CysRS). According to the comparison on BLASTn, these two genes are 78.5% identical at the nucleotide level and 89.9% identical at the amino acid level. In this thesis the plasmid-encoded gene is termed *cysS*^{pBT1}.

The RNA-seq-based transcriptomic data for the *cysS*^{chr} gene in D23580 and 4/74, and *cysS*^{pBT1} gene in D23580 are shown in Figure 5.1 A. The expression of *cysS*^{chr} in D23580 was dramatically down-regulated, compared with strain 4/74, while the *cysS*^{pBT1} gene was highly expressed in most growth conditions.

Transposon mutagenesis is a powerful approach for assessing gene essentiality in *Salmonella enterica* (Barquist *et al.*, 2013). A TIS experiment that involved strain D23580 identified large numbers of transposon insertion sites on the *cysS*^{chr} gene, but no insertion sites within the *cysS*^{pBT1} gene (Figure 5.1 B), suggesting that *cysS*^{pBT1} is required for survival of D23580 in rich medium (Canals, *et al.*, 2019b).

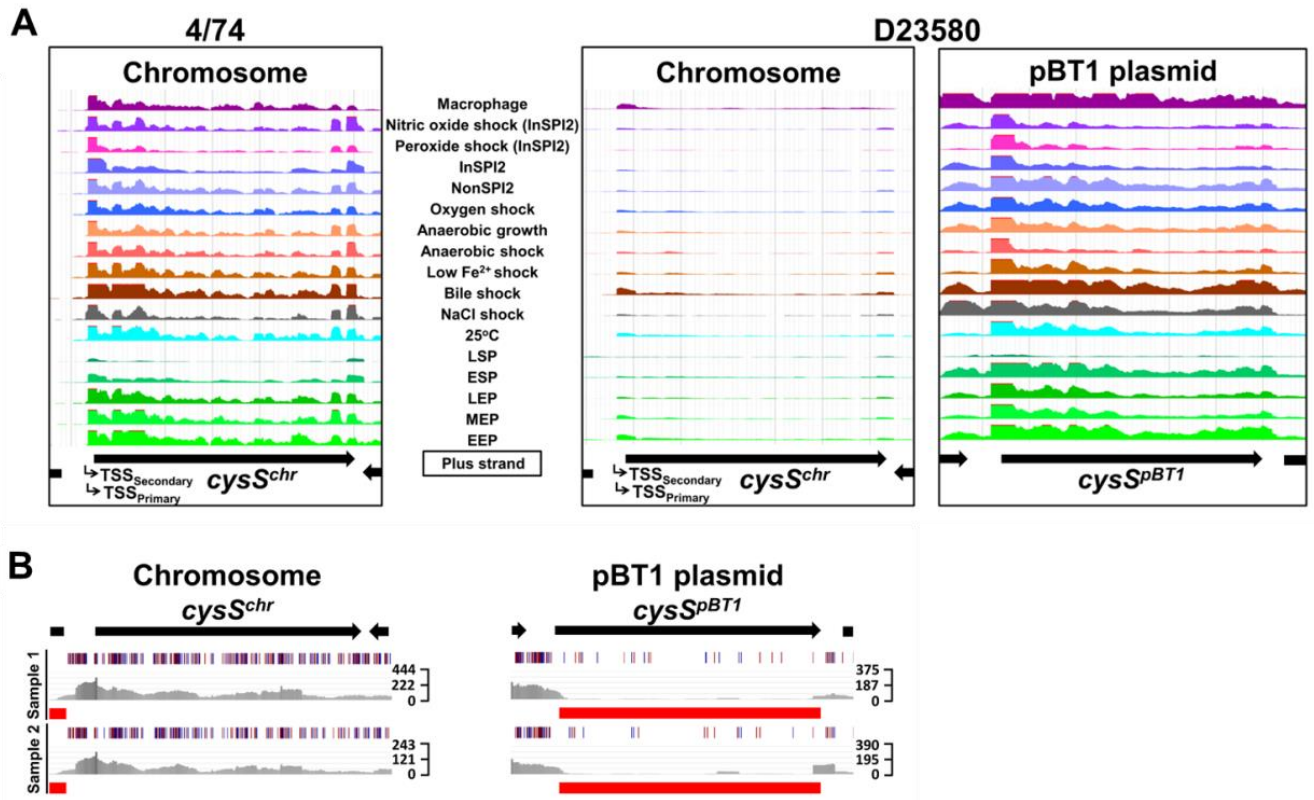


Figure 5.1. RNA-seq-based expression data for the chromosomal *cysS* gene and the plasmid pBT1-encoded *cysS* paralog of in *S.* Typhimurium strains D23580 and 4/74. (A) RNA-seq data for the *cysS* gene of strain 4/74, and both the $cysS^{chr}$ and $cysS^{pBT1}$ genes of D23580 from 16 *in vitro* growth condition, plus macrophage (Canals *et al.,* 2019b; Kröger *et al.*, 2013; Srikumar *et al.,* 2015). The scale of the mapped reads was 1 to 500. Figure from Canals *et al.,* 2019a. (B) Transposon insertion sequencing results for $cysS^{chr}$ and $cysS^{pBT1}$ in strain D23580. The data were from Canals *et al.,* 2019a, visualized in the Dalliance genome viewer and available online at: https://hactar.shef.ac.uk/D23580/. The black arrow represents genes. Below the arrow, the vertical blue and red lines represent transposon insertion sites (red: same direction; blue: opposite direction). The grey histogram show raw data of the mapped sequence reads, and the horizontal red bar indicates that the gene had an insertion index value <0.05 (Barquist *et al.*, 2016; Langridge *et al.*, 2009), and was considered as "required" for growth in rich medium (Canals *et al.,* 2019a).

5.3 Removal of *cysS*^{pBT1} results in poor growth of D23580 in M9 minimal media with glucose

To build on the TIS data that suggested that *cysS*^{pBT1} was required for survival of D23580, individual deletion mutants of *cysS*^{chr} and *cysS*^{pBT1} as well as a plasmid-cured derivative, D23580 ΔpBT1, were constructed (Table 2.2). The growth rates of these three strains were compared with D23580 WT in liquid media and on agar plates, using both the rich medium LB and the M9 minimal medium with 0.4% glucose as a sole carbon source.

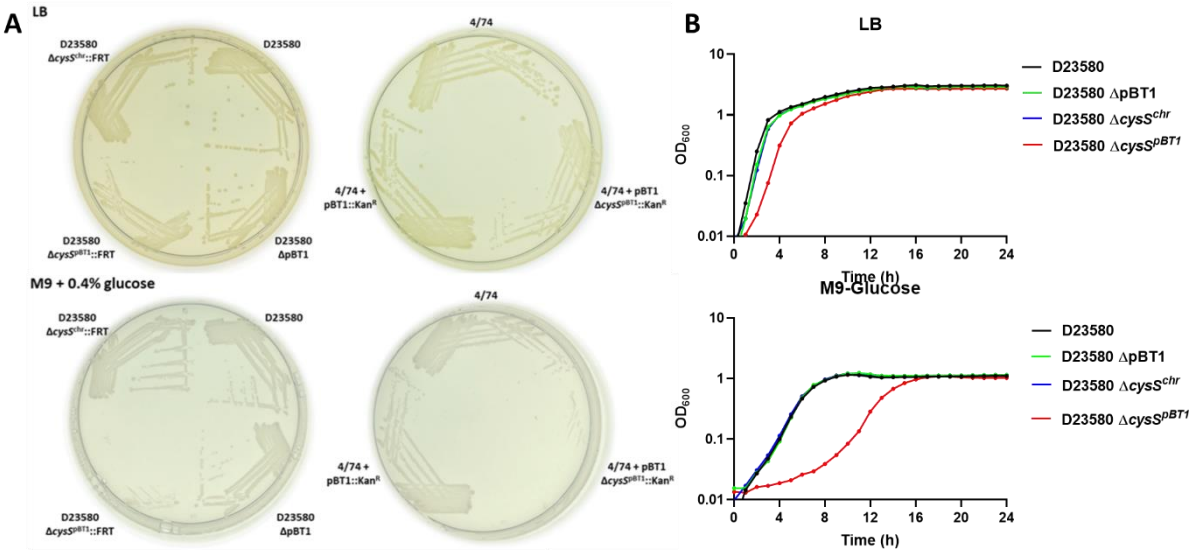
Figure 5.2A shows that all four strains grew well on LB agar plates after overnight incubation. D23580 WT, Δ*cysS*^{chr} mutant and ΔpBT1 grew well on M9 glucose plates, whereas, the deletion mutant of *cysS*^{pBT1} showed poor growth on M9 minimal media. To investigate growth characteristics in more detail, the strains were grown in liquid culture. In LB liquid media, the four strains had a similar exponential growth rates (Figure 5.2B). However, the Δ*cysS*^{pBT1} mutant had a longer lag time that was extended by about 1 hour (Figure 5.2C). In M9 minimal media, the lag time of the Δ*cysS*^{pBT1} mutant was extended by more than 8 hours (Figure 5.2C).

To investigate the influence of pBT1 upon the physiology of another *S.* Typhimurium strain, a heterologous experiment was conducted. Kanamycin-resistant derivatives of both the pBT1 plasmid and the pBT1 Δ*cysS*^{pBT1} plasmid were constructed in D23580, and were conjugated into the ST19 strain 4/74 (Table 2.2). Growths of the

resulting 4.74 derivatives were compared on agar plates. 4/74 WT, 4/74 pBT1*::aph* and 4/74 pBT1 Δ*cysS*^{pBT1}*::aph* grew well on LB plates (Figure 5.2A). 4/74 WT, 4/74 pBT1*::aph* grew well on minimal media, but 4/74 pBT1 Δ*cysS*^{pBT1}*::aph* grew poorly on M9 glucose plates.

These data for strain D23580 suggests that the *cysS*^{pBT1} gene is required for optimal growth, whereas the *cysS*^{chr} gene was dispensable. The fact that 4/74 pBT1 Δ*cysS*^{pBT1}*::aph* grew poorly on minimal media shows that the *cysS*^{chr} of 4/74 had become non-functional. These data are consistent with pBT1 being responsible for transcriptional repression of the *cysS*^{chr} gene.

Because the ΔpBT1 mutant grew at the same rate as the D23580 WT on minimal media (Figure 5.2B), it was deduced that the *cysS*^{chr} was functional in the D23580 ΔpBT1 strain. This finding is consistent with pBT1-encoded genes mediating the down-regulated expression of *cysS*^{chr}.



S. Typhimurium Strains	Lag time (hour)	Max Growth Rate	ANNOVA
	LB medium		
D23580	1.20	0.367	ns (P > 0.05)
D23580ΔpBT1	1.47	0.378	
D23580Δ*cysS*^{chr}	1.54	0.385	
D23580Δ*cysS*^{pBT1}	2.93	0.385	
	M9-glucose medium		
D23580	4.88	0.290	ns (P > 0.05)
D23580ΔpBT1	4.88	0.292	
D23580ΔcysS^{chr}	4.76	0.291	
D23580ΔcysS^{pBT1}	11.48	0.275	

Figure 5.2. The *cysS*^{pBT1} gene is required for growth of *S.* Typhimurium D23580. (A) Growth of D23580 and 4/74 wild type and their mutants on LB (Upper) and M9 glucose (Bottom) plates after overnight growth at 37°C. (B) Growth curves of D23580 WT and 3 mutant derivatives Δ*cysS*^{chr}, Δ*cysS*^{pBT1} and ΔpBT1 (JH4228, JH4229, JH4230, Table 2.2) in LB medium (Upper) and M9 glucose medium (Bottom), n = 8. Error bars show Standard Deviation. The table below shows lag time and max growth rate of each strains. Lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 from each growth curve (Section 2.4.2). ANNOVA was used to check significance for each group of strains (ns: P > 0.05).

An RNA-seq experiment involving the D23580 ΔpBT1 strain confirmed that, *cysS*^{chr} was down-regulation by pBT1 in D23580 (Canals *et al.*, 2019b). Specifically, the transcriptome of D23580 ΔpBT1 showed that *cysS*^{chr} was being transcribed efficiently. The level of expression of *cysS*^{chr} was a TPM value of approximately 90 in most growth conditions, suggesting that factors encoded by the pBT1 plasmid are responsible for down-regulating the expression of *cysS*^{chr} in D23580 (Canals, *et al.*, 2019b). Figure 5.2 suggests that the *cysS*^{pBT1} gene product did not directly influence the expression of *cysS*^{chr} in D23580. If *cysS*^{pBT1} had been the only factor, the Δ*cysS*^{pBT1} mutation would have caused the transcriptional level of *cysS*^{chr} to increase allowing the mutant to grow at the same speed as the wild type or the pBT1-cured mutant without an extended lag phase in minimal media. The opposite results can be seen in Figure 5.2B. These findings suggest that plasmid pBT1-encoded factors other than *cysS*^{pBT1} repressed the expression of *cysS*^{chr}.

5.4 The search for a repressor of *cysS*^{chr} using Tn*5* transposon mutagenesis

To search for pBT1-encoded genes responsible for the down-regulation of *cysS*^{chr} expression, Tn*5* mutagenesis was used to randomly generate single-gene mutations of D23580. The value of Tn*5*-mediated mutagenesis has been described previously (Berg *et al.*, 1983). The poor growth of the Δ*cysS*^{pBT1} mutant in M9 minimal medium, with 0.4% glucose as the sole carbon source, permitted mutants to be identified as large colonies after overnight incubation. The Tn5-delivery suicide plasmid pRL27

(Larsen *et al.*, 2002) was transformed into the strain D23580 Δ*cysS*^{pBT1} following the protocol described in Section 2.7.2. In total, 250,000 transformants were spread onto 10 M9 glucose plates for overnight incubation (Section 2.7.2). A total of 25 large colonies were selected from M9 plates, re-streaked for 2 passages on LB plates followed by confirmation of the fast-growing phenotype on M9 glucose plates. The Tn*5* insertion sites were identified by sequencing the arbitrary PCR products, as described in Section 2.7.2. The locations of 15 transposons were identified from these 25 colonies and the Tn*5* insertion sites are detailed in Table 5.1.

The experiment had been inspired by the hypothesis that transposon insertions into certain pBT1-encoded genes would cause de-repression of the *cysS*^{chr} gene. However, none of the insertion sites were located in the pBT1 plasmid. Instead, they were located in the functionally-uncharacterised chromosomal genes *STMMW-05081* or *STMMW-09631*, or in an intergenic region of the pSLT-BT plasmid (Table 5.1).

The lack of pBT1-associated Tn*5* mutants suggests that the experiment had resulted in loss of pBT1. A PCR-based approach was used to assess the presence of the pBT1 plasmid in a selection of Tn*5* mutants, and discovered that none of the mutants carried the pBT1 plasmid (Figure 5.3). The result suggests that it was the loss of pBT1 that permitted the mutants to grow rapidly in M9 glucose minimal medium.

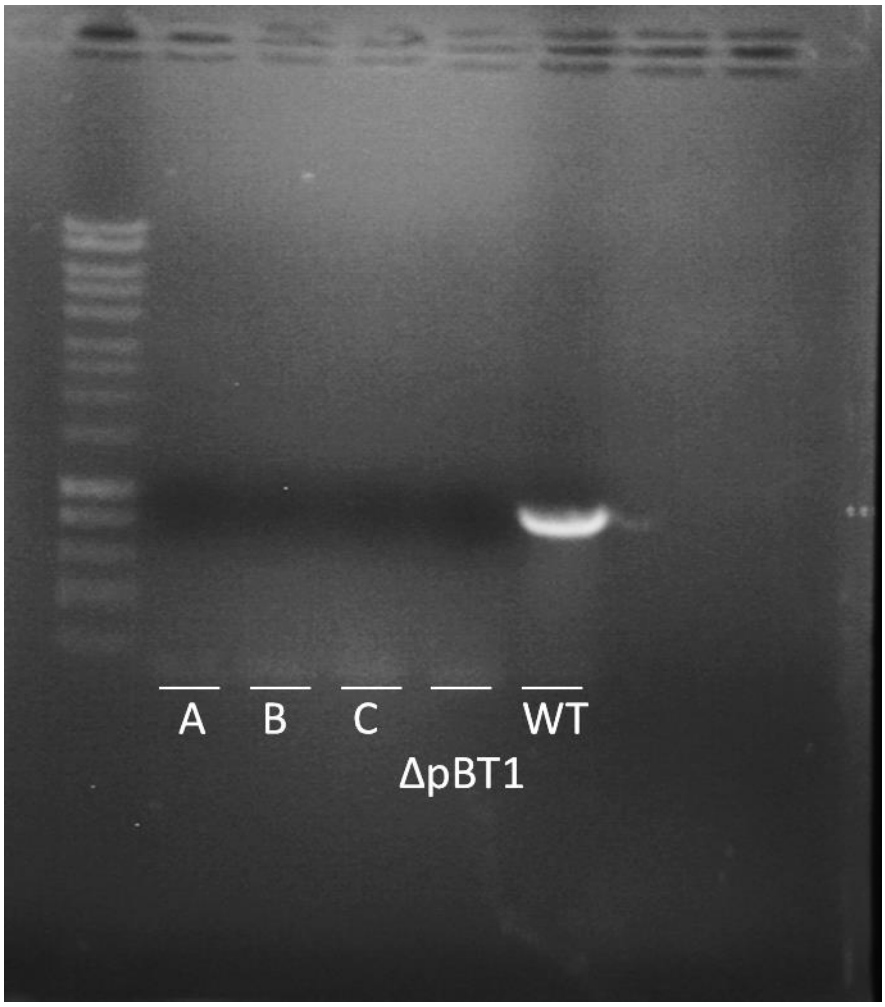


Figure 5.3. Absence of plasmid pBT1 from 3 fast-growing Tn*5* mutants that were selected in M9 glucose medium. A PCR-based approach was used to assess the presence of the pBT1 plasmid. The PCR was performed with primers pBT1-81_ex_fw x pBT1-81_ex_rv to amplify the 800 bp *pBT1-0081* gene from pBT1 (Table 2.3). D23580 WT and ΔpBT1 mutant were used as positive and negative control, respectively. Samples on the photo were: A: CysS-1, insertion site in *STMMW-05081*; B: CysS-6, insertion site after *SLT-BT0571*; C: CysS-10, insertion site in *STMMW-09631*.

Table 5.1. List of Tn*5* mutants from the Tn*5* mutagenesis experiment 1

Strain name	Insertion site	Location	Annotation
CysS-1	*STMMW-05081*	Chromosome	Putative exported protein
CysS-2	*STMMW-05081*	Chromosome	Putative exported protein
CysS-3	*STMMW-05081*	Chromosome	Putative exported protein
CysS-4	*STMMW-05081*	Chromosome	Putative exported protein
CysS-5	*STMMW-05081*	Chromosome	Putative exported protein
CysS-6	*SLT-BT0571* downstream	pSLT-BT	Putative carbonic anhydrase
CysS-7	*SLT-BT0571* downstream	pSLT-BT	Putative carbonic anhydrase
CysS-8	*SLT-BT0571* downstream	pSLT-BT	Putative carbonic anhydrase
CysS-9	*STMMW-05081*	Chromosome	Putative exported protein
CysS-10	*STMMW-09631*	Chromosome	Putative pirin-relative protein
CysS-11	*STMMW-09631*	Chromosome	Putative pirin-relative protein
CysS-12	*STMMW-09631*	Chromosome	Putative pirin-relative protein
CysS-13	*STMMW-09631*	Chromosome	Putative pirin-relative protein
CysS-14	*STMMW-09631*	Chromosome	Putative pirin-relative protein
CysS-15	*STMMW-09631*	Chromosome	Putative pirin-relative protein

A new experimental strategy was designed to isolate Tn*5* insertions in both pBT1 and the chromosome. First, a derivative of D23580 that only carried one plasmid was constructed. Briefly the pBT2, pBT3 and pSLT-BT plasmids were sequentially cured, one by one, using a Cas9-CRISPR plasmid-based method following the protocol described in Section 2.7.3. The sequential plasmid-curing approach was used to generate D23580 ΔpBT2 ΔpBT3 ΔpSLT-BT (SZS010, Table 2.2), which was referred to as "D23580 pBT1^{+}" in the remainder of the thesis.

To assist understanding of the subsequent experiments involving plasmid-cure strains, the construction details have been summarized in Table 5.2.

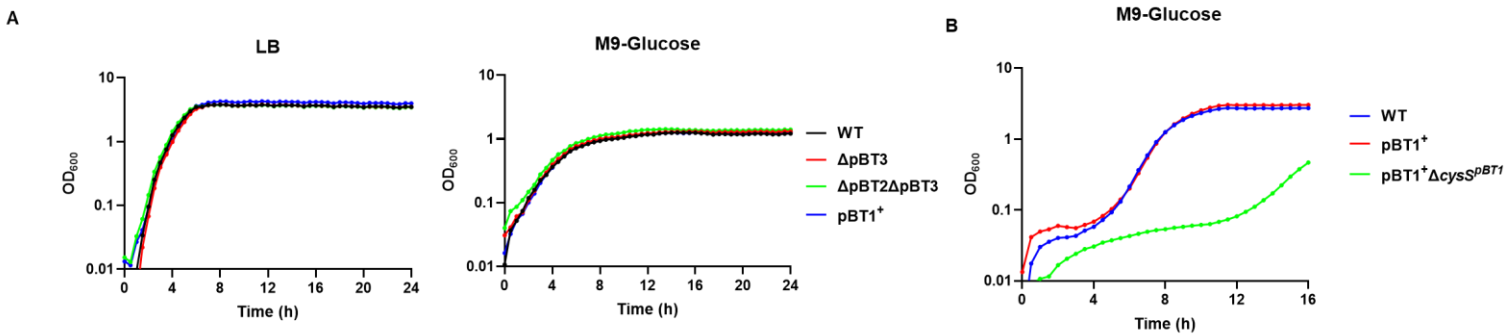
Table 5.2. The plasmid-free derivatives of D23580[a]

Strain identifier	Plasmid content	Description & Construction Strategy
Salmonella Typhimurium D23580[a]		
SNW88	pBT1 pBT2 pSLT-BT	D23580 ΔpBT3
SZS009	pBT1 pSLT-BT	D23580 ΔpBT2, ΔpBT3
SZS010[b]	pBT1	D23580 ΔpBT2, ΔpBT3, ΔpSLT-BT
[a]Details of strain construction are in Section 2.7.3. [b]Strain SZS010 is referred to as D23580 pBT1^{+} in this thesis		

The growth rate of the interim D23580 ΔpBT3 and D23580 ΔpBT2 ΔpBT3 plasmid-cured strains, and the D23580 pBT1^{+}, was compared in LB and M9 glucose media (Figure 5.4 A). No significant difference was found in the exponential growth

rate or the lag time of these plasmid-cured mutants, leading to the inference that plasmid pBT2, pBT3 and pSLT-BT were not responsible for the down-regulated expression of *cysS*^{chr}.

The *cysS*^{pBT1} gene was deleted from D23580 pBT1^{+}, using the pKD3 plasmid to recombine the chloramphenicol resistance gene into the *cysS*^{pBT1} locus of pBT1 to generate strain D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat*. A comparison of growth of D23580 WT, D23580 pBT1^{+} and D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* was performed in M9 glucose media (Figure 5.4 B). Both D23580 WT and D23580 pBT1^{+} had a similar exponential growth rate. In contrast, the D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* mutant grew slower than the wild type, and it showed the same extended lag time that had been observed previously for strain D23580 Δ*cysS*^{pBT1} (Figure 5.2B). Accordingly, the same Tn*5* transposon mutagenesis approach could again be used to select large colonies after overnight incubation at 37°C.



S. Typhimurium Strains	Lag time (hour)	Max Growth Rate	T-test
	LB medium		
D23580	2.595	0.432	ns (P > 0.05)
D23580pBT1+	2.691	0.444	
D23580ΔpBT3	2.787	0.463	
D23580ΔpBT2ΔpBT3	2.402	0.390	
	M9-glucose medium		
D23580	2.691	0.129	ns (P > 0.05)
D23580pBT1+	2.594	0.140	
D23580ΔpBT3	2.450	0.134	
D23580ΔpBT2ΔpBT3	2.210	0.125	

S. Typhimurium Strains	Lag time (hour)	Max Growth Rate	T-test
	M9-glucose medium		
D23580	5.750	0.235	
D23580pBT1+	5.814	0.263	ns (P > 0.05)
D23580pBT1+ Δ*cysS*^{pBT1}	14.206	0.124	**** (P < 0.0001)

Figure 5.4. Similarities in growth rate of D23580 WT and plasmid-free derivatives, and the growth defect associated with deletion of *cysS*^{pBT1}. (A) Growth curves of D23580 WT and its plasmid-free derivative strains ΔpBT3, ΔpBT2 ΔpBT3 and pBT1+ (strains SNW88, SZS009 and SZS010, Table 2.2) in LB and M9 glucose medium, n = 4. (B) Comparison between D23580 WT, pBT1+ and its *cysS*^{pBT1} deletion mutant (strains SZS010 and SZS014, Table 2.2) in M9 glucose medium, n = 8. The table below shows lag time and max growth rate for each panel. Lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 from each growth curve (Section 2.4.2). T-test was performed for each panel to show significance (ns: P > 0.05; ****: P < 0.0001).

The Tn*5* mutagenesis experiment 2 was performed using D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* (SZS014, Table 2.2), following the protocol from Section 2.7.2. In total, 250,000 transformants were spread onto M9 glucose plates. Four large colonies were observed after overnight growth. Eight large colonies were selected from the agar plate after 2-day incubation (4 were large colonies from Day 1 and 4 from Day 2). An example plate of the Tn*5* mutagenesis is shown in Figure 5.5.

The 8 candidate kanamycin-resistant Tn5 insertions were transduced into a clean D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* genetic background to confrim that the fast growth on M9 glucose was directly linked to the Tn*5* transposon insertion. Arbitrary PCR was then used to identify the Tn*5* insertion sites, which generated 7 DNA bands from all the candidates. Five transposon insertion sites were identified by sequencing and are described in Table 5.3. The genetic context of these insertion sites and flanking genes are shown in Figure 5.6, 5.7, 5.8 and 5.9.

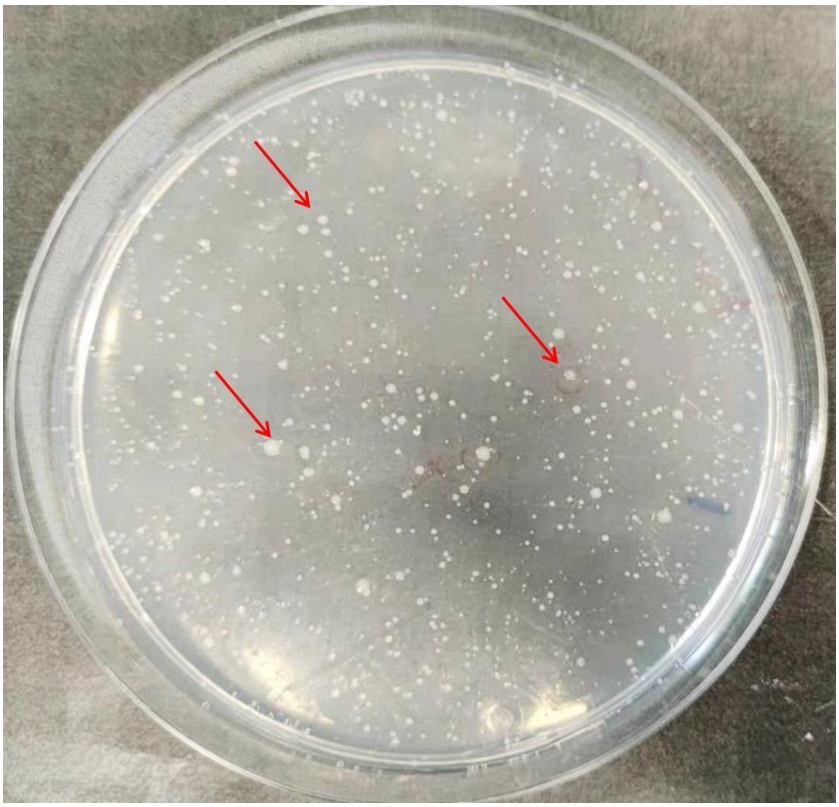


Figure 5.5. Fast-growing Tn*5* mutants appeared as large colonies on an M9 glucose agar plate. A representative agar plate from Tn*5* mutagenesis experiment 2, after two-day incubation at 37°C. Three large colonies are indicated by red arrows.

Table 5.3. List of Tn*5* mutants from the Tn*5* mutagenesis experiment 2

Strain name	Insertion site	Location	Annotation
Tn5-1	*pBT1-0081*	pBT1	Hypothetical protein
Tn5-2	*pBT1-0081*	pBT1	Hypothetical protein
Tn5-3	*traI*	pBT1	Conjugal transfer protein
Tn5-4	Between *ydjA* and *pBT1-0871*	pBT1	ydjA: Putative NAD(P)H nitroreductase *pBT1-0871*: Hypothetical protein
Tn5-5	*STMMW_16671*	chromosome	Hypothetical protein

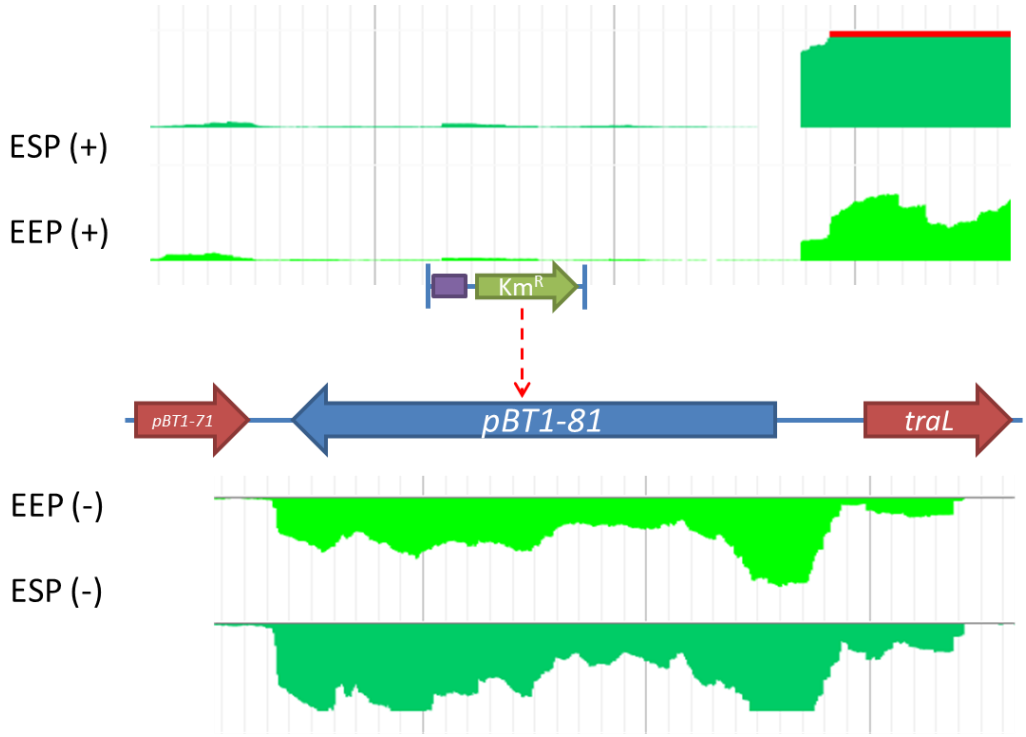


Figure 5.6. Genetic context of Tn*5* insertion site and flanking regions for strains Tn5-1 and Tn5-2. In both the Tn5-1 and Tn5-2 mutants, the insertion site was in the center of the *pBT1-0081* gene in the opposite orientation. The green arrow at the top indicates the kanamycin resistance *aph* gene, and the purple bar shows the origin of replication *ori*R6K from the Tn*5*-delivery plasmid pRL27. Genetic orientation in the plasmid DNA is indicated by red arrows (plus strand) and blue arrows (minus strand). Below, the RNA-seq data for this chromosomal region from the EEP and ESP growth condition are shown (Canals *et al.*, 2019b). The scale of the mapped reads was 1 to 500. The genomic context of *pBT1-0081* and the detailed level of transcription are available at JBrowse from the Hinton Lab at https://tinyurl.com/pBT1-0081.

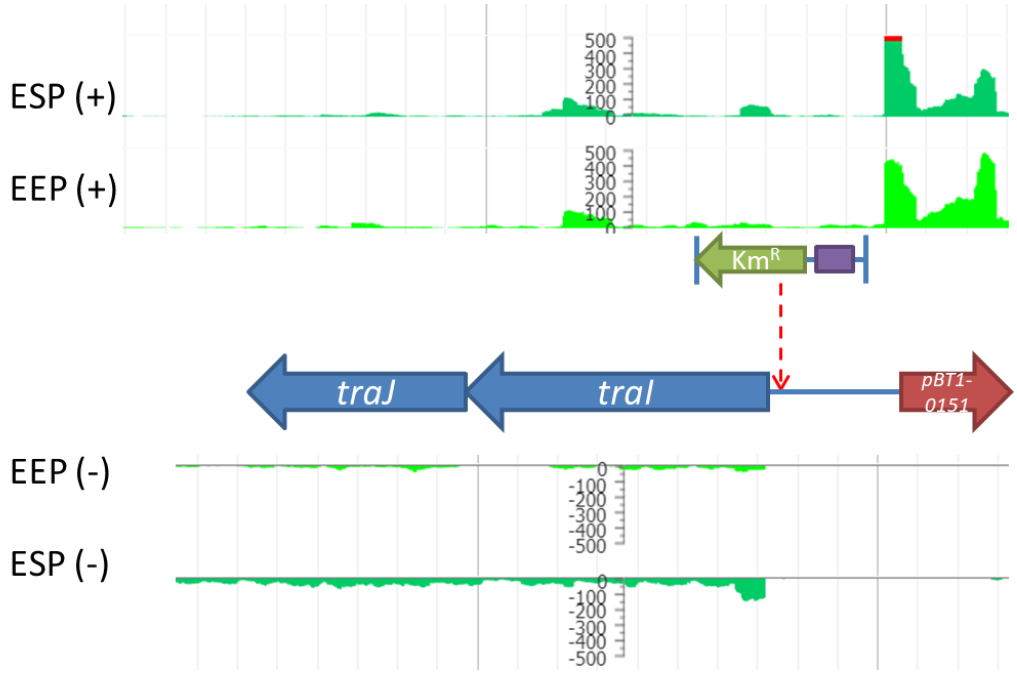


Figure 5.7. Genetic context of Tn*5* insertion site and its flanking regions for strain Tn5-3. In the Tn5-3 mutant, the insertion site was located 3 bp upstream of the start codon of the gene *traI* on the same orientation. The green arrow at the top indicates the kanamycin resistance *aph* gene, and the purple bar shows the origin of replication *ori*R6K from the Tn*5*-delivery plasmid pRL27. Genetic orientation in the plasmid DNA is indicated by red arrows (plus strand) and blue arrows (minus strand). Below, the RNA-seq data for this chromosomal region from the EEP and ESP growth condition are shown (Canals *et al.*, 2019b). The scale of the mapped reads was 1 to 500. The genomic context of *traI* and the detailed level of transcription are available at JBrowse from the Hinton Lab at https://tinyurl.com/pBT1-traI.

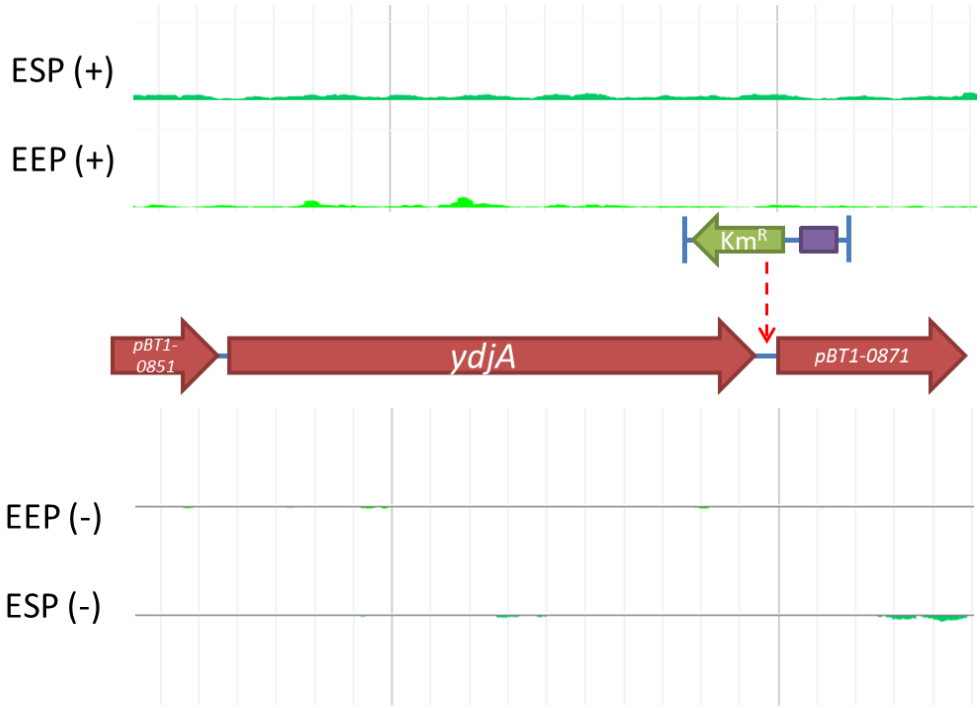


Figure 5.8. Genetic context of Tn*5* insertion site and its flanking regions for strain Tn5-4. In the Tn5-4 mutant, the insertion site was between the gene *ydjA* and *pBT1-0871,* in the opposite orientation. The green arrow at the top indicates the kanamycin resistance *aph* gene, and the purple bar shows the origin of replication *ori*R6K from the Tn*5*-delivery plasmid pRL27. Genetic orientation in the plasmid DNA is indicated by red arrows (plus strand) and blue arrows (minus strand). Below, the RNA-seq data for this chromosomal region from the EEP and ESP growth condition are shown (Canals *et al.*, 2019b). The scale of the mapped reads was 1 to 500. The genomic context of *ydjA* and the detailed level of transcription are available at JBrowse from the Hinton Lab at https://tinyurl.com/pBT1-ydjA.

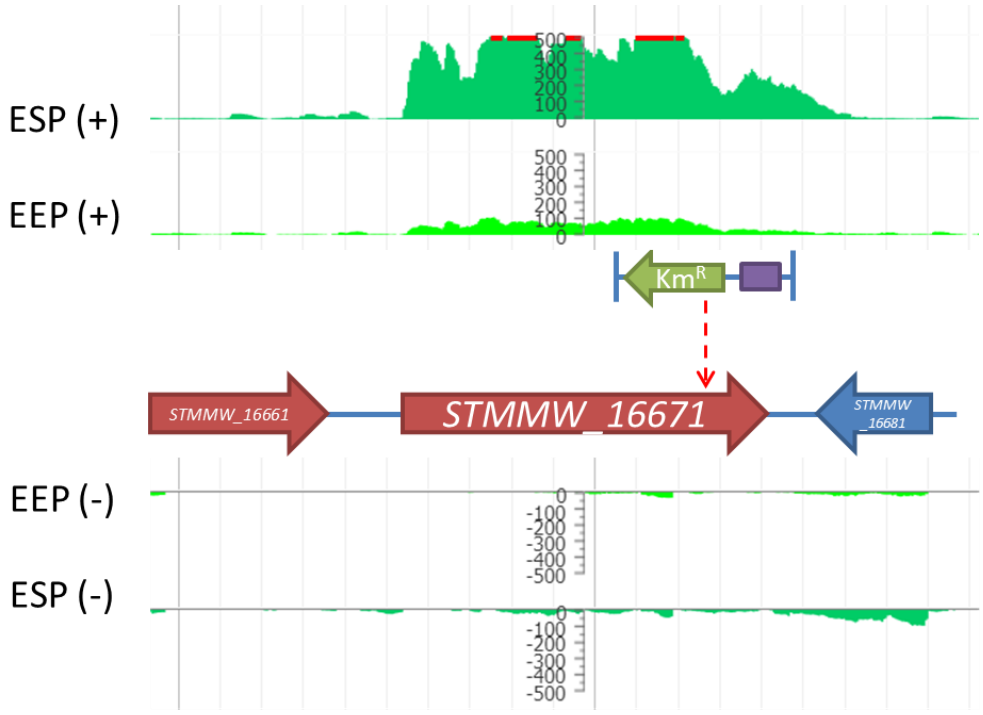


Figure 5.9. Genetic context of Tn*5* insertion site and its flanking regions for strain Tn5-5. In the Tn5-5 mutant, the insertion site was in the gene *STMMW_16671,* in the opposite orientation. The green arrow at the top indicates the kanamycin resistance *aph* gene, and the purple bar shows the origin of replication *ori*R6K from the Tn*5*-delivery plasmid pRL27. Genetic orientation in the plasmid DNA is indicated by red arrows (plus strand) and blue arrows (minus strand). Below, the RNA-seq data for this chromosomal region from the EEP and ESP growth condition are shown (Canals *et al.*, 2019b). The scale of the mapped reads was 1 to 500. The genomic context of *STMMW_16671* and the detailed level of transcription are available at JBrowse JBrowse from the Hinton Lab at https://tinyurl.com/STMMW-16671.

Table 5.3 detailed the five Tn*5* mutants from the Tn*5* mutagenesis experiment 2 that were capable of rapid growth on M9 glucose plates. Two candidates carried the Tn*5* insertion site in the same location of the gene *pBT1-0081* (Figure 5.6). The function of *pBT1-0081* remains unknown and is discussed below.

The insertion site in isolate Tn5-3 was located 3 bp upstream of the start ATG codon of the gene *traI* and downstream of the promoter region of *traI*, which can be

visualized in the green RNA-seq track in Figure 5.7. This insertion site could cause a polar effect on *traI* as it inserted the kanamycin resistance gene *aph* promoter upstream of the natural promoter of *traI* (Hutchison *et al.*, 2019). The *traI* gene encodes the conjugal transfer protein, which binds DNA and is essential for conjugative plasmid transfer (Dostál *et al.*, 2010; McLaughlin *et al.*, 2014). It is possible that this transposon insertion could alter the rate of pBT1 conjugation, but it is not clear how this might influence expression of *cysS* gene or growth in minimal medium.

The Tn*5* insertion in isolate Tn5-4 was located between the *ydjA* gene, which encodes a minimal nitroreductase (Choi *et al.*, 2008), and an unknown gene *pBT1-0871*. Insertion in this position resulted in a longer distance between *pBT1-0871* gene and the promoter, and would likely to repress the expression of *pBT1-0871*. As the mutagenesis experiment was designed to identify plasmid pBT1-encoded genes responsible for the down-regulation of $cysS^{chr}$ in D23580, the chromosomal insertion in strain Tn5-5 was not further investigated in this thesis.

5.5 The search for a repressor for $cysS^{chr}$ using experimental evolution

One limitation of Tn*5*-mediated mutagenesis is that the insertion of a transposon into coding genes generally causes loss of function. Accordingly, a complementary approach involving experimental evolution methodology was used in parallel, using the protocol from Section 2.4.7 to search for other types of mutations that controlled

the regulation of the *cysS*^{chr} gene. As the *cysS*^{pBT1} knock-out mutant grew poorly in M9 minimal medium with 0.4% glucose as its sole carbon source, a small amount of the mutant cells were used to start a bacterial culture to generate spontaneous mutants that were able to grow fast in M9 glucose medium.

5.5.1 Evolution experiment 1

The “liquid” method described in Section 2.4.7 was used to generate spontaneous mutants for D23580 Δ*cysS*^{pBT1}*::frt* and 4/74 pBT1 Δ*cysS*^{pBT1}*::aph* capable of rapid growth on minimal media. Candidate mutants were identified as large colonies on M9 glucose plate (for 4/74, supplemented with kanamycin) and were genome sequenced by MicrobesNG (England). Snippy (https://github.com/tseemann/snippy) was used to identify the mutation site and SNPs on the genome of candidates. The two evolution experiments yielded eight mutants that grew rapidly on M9 glucose media (4 for the D23580 background and 4 for the 4/74 pBT1 background). The detailed description of the mutations is shown in Table 5.4.

Half of the mutations identified from these two evolution experiments were located in the gene *pBT1-0081*. Specifically, the same single nucleotide deletion mutant was found in both D23580 and 4/74 pBT1 backgrounds (strains JHL817 and JHL809). In Section 5.4, the same gene *pBT1-0081* was identified by Tn*5* transposon mutagenesis, raising the possibility that *pBT1-0081* plays an important role in restricting the growth of Δ*cysS*^{pBT1} mutant in minimal medium which I infer was due

to down-regulation of *cysS*^{chr}.

Two different mutations were identified in the *relA* gene (one SNP and one deletion mutation). *relA* encodes the enzyme RelA, a GTP pyrophosphokinase that synthesizes (p)ppGpp, a molecule involved in the stringent response of *E. coli*, *Salmonella* and many other gram-negative bacteria (Fitzsimmons *et al.*,2018; Zhang *et al.*, 2016). The stringent response is trigged by nutrient starvation and controls transcriptional regulation (Potrykus and Cashel, 2008; Srivatsan *et al.*, 2008; Thompson *et al.*, 2006). The process is trigged by the ribosome-associated RelA factor that senses the presence of uncharged tRNA (Agirrezabala *et al.*, 2013; Haseltine *et al.*, 1973). Whether these *relA* mutations influence the expression of the *cysS*^{chr} gene has not been investigated yet due to the time limitations of this project.

Table 5.4. List of spontaneous mutants from the evolution experiment 1

Strain name	Background	Location	Position	Mutation	Type	Gene	Annotation
JHL807	D23580 Δ*cysS*^{pBT1}*::frt*	pBT1	8349	C -> A	SNP	*pBT1-0081*	Hypothetical protein
JHL815	D23580 Δ*cysS*^{pBT1}*::frt*	Chromosome	3095476	ATATCGCCGCCGCC -> A	Del	*relA*	GTP pyrophosphokinase
JHL817	D23580 Δ*cysS*^{pBT1}*::frt*	pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
JHL820	D23580 Δ*cysS*^{pBT1}*::frt*	pBT1	20713	G -> T	SNP	*traY*	Plasmid conjugative transfer integral membrane protein TraY
JHL801	4/74pBT1 Δ*cysS*^{pBT1}*::aph*	pBT1	8047	G -> T	SNP	*pBT1-0081*	Hypothetical protein
JHL809	4/74pBT1 Δ*cysS*^{pBT1}*::aph*	pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
JHL813	4/74pBT1 Δ*cysS*^{pBT1}*::aph*	pBT1	7665	A -> G	SNP	*pBT1-0061*	Hypothetical protein
JHL827	4/74pBT1 Δ*cysS*^{pBT1}*::aph*	Chromosome	3124345	A -> C	SNP	*relA*	GTP pyrophosphokinase

5.5.2 Evolution experiment 2

In the evolution experiment 1 described in Section 5.5.1, spontaneous mutants of D23580 Δ*cysS*^{pBT1}*::frt* and 4/74 pBT1 Δ*cysS*^{pBT1}*::aph* capable of rapid growth on minimal media were isolated from “liquid” method.

A second evolution experiment 2 was performed using D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain (SZS014, Table 2.2), a strain that grew poorly in minimal media (Figure 5.4B). The strain carries pBT1 as its sole plasmid and the plasmid is tagged with chloramphenicol resistance, allowing selection on agar plates with chloramphenicol to maintain the presence of pBT1 in all cells, and to identify mutants only in chromosome or pBT1 plasmid. Evolution experiment 2 involved both the “liquid” and “agar plate” methods from Section 2.4.7 to select mutants capable of rapid growth on minimal media. The “agar plate” method involved two plates, and the resulting cultures from the “liquid” method were spread onto 10 M9 glucose plates.

During the “liquid” method, the growth of each of eight biological replicates was tracked using the Growth Profiler over a 48 hour period (Figure 5.10 A). All the biological replicates started to grow after about 36 hours and had then exhibited a similar exponential growth rate. The extended lag time was greater than observed previously (Figure 5.4B), probably due to the low number of cells used in the inoculum for these evolution experiments.

An example M9 glucose plate that was spread after the 48-hour incubation is shown in Figure 5.10 B. The poor growth of D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain was reflected by a diffuse bacterial lawn, with tiny colonies appearing at the edges of the plate. The large colonies that arose occasionally represented candidate spontaneous mutants capable of rapid growth on M9 glucose media. For every M9 glucose plate obtained from the evolution experiment (2 from “agar plate” method and 8 from “liquid” method), 4 large colonies were carefully collected, minimizing contact with the bacterial lawn underneath. The 40 candidate mutants were re-streaked on M9 glucose plates.

The 10 candidate mutants that grew most rapidly on minimal media were selected for further study. The 10 candidate mutants were screened to check that they did not carry mutations of the *rpoS* gene. Previous evolution experiments performed in the Hinton Lab generated large numbers of *rpoS* mutants (N. Wenner *Pers. Comm.*). The *rpoS* mutants can be identified because they rapidly grow on M9 plate with succinate as the sole carbon source, unlike *S.* Typhimurium WT grew poorly (N. Wenner *Pers. Comm.*).

The 10 candidate mutants were screen on M9 plates supplement with 20 mM Na-succinate, and showed the same slow growth phenotype as the D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain which show that the RpoS protein was fully functional

The 10 candidate mutants were sent to MicrobesNG (England) for whole genome sequencing, and the genomes were analysed to identify nucleotide changes (Section 2.8.2). The candidate mutants isolated from the “liquid” evolution experiments were designated CysS-L1 to CysS-L4. The candidate mutants isolated from the “agar plate” evolution experiments were designated CysS-P1 to CysS-P6. Details of the mutations identified from the 10 isolates are shown in Table 5.5.

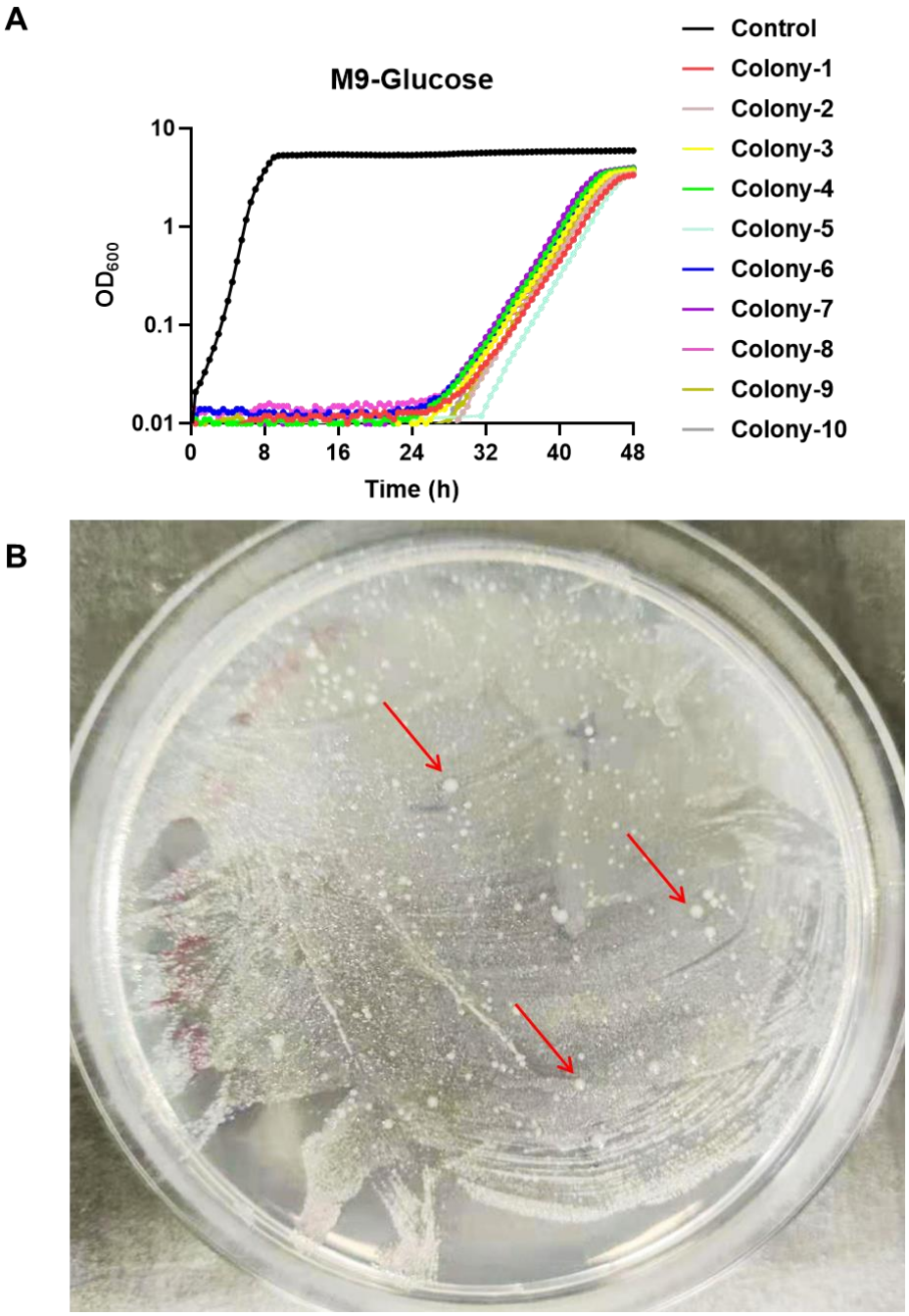


Figure 5.10. Fast-growing spontaneous mutants of D23580 pBT1⁺ Δ*cysS*^{pBT1}*::cat* in M9 glucose media. (A) Following experimental evolution in liquid media (Section 2.4.8), the growth curves of the D23580 pBT1⁺ (Strain SZS010, Table 2.2) and 10 biological replicates of D23580 pBT1⁺ Δ*cysS*^{pBT1}*::cat* (strain SZS014, Table 2.2) were determined in M9 glucose medium using the Growth Profiler (Section 2.4.2; labeled Colony 1 to Colony 10). Growth of the D23580 WT is shown as a black line (labeled as control). The 10 spontaneous mutants are indicated by different colours. (B) Following experimental evolution on solid media (Section 2.4.8), a representative agar plate is shown after two-day incubation at 37°C. Large colonies, representing candidate mutants, are indicated by red arrows.

Table 5.5. List of spontaneous mutants from the evolution experiment 2

Strain name	Background	Location	Position	Mutation	Type	Gene	Annotation
CysS-L1	D23580 ΔpBT2 ΔpBT3 ΔpSLT-BT Δ*cysS*^{pBT1}*::cat*	N/A*	N/A	N/A	N/A	N/A	N/A
CysS-L2		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-L3		pBT1	7664	T -> C	SNP	*pBT1-0061*	Hypothetical protein
CysS-L4		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-P1		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-P2		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-P3		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-P4		Chromosome	637890	G -> T	SNP	P_{cysS}^{chr}	Promoter of *cysS*^{chr}
CysS-P5		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein
CysS-P6		pBT1	7884	AT -> A	Del	*pBT1-0081*	Hypothetical protein

***N/A indicated that no mutant site was identified in the CysS-L1 strain.**

The genome sequences of the 10 candidate mutants had very few nucleotide changes, showing that the experimental evolution approach had not caused significant changes to the *S.* Typhimurium genome. Eight of the ten mutants carried nucleotide changes in pBT1 plasmid-encoded genes, supporting the inference in Section 5.3.

Seven of the mutations identified in evolution experiment 2 involved the *pBT1-0081* gene. Four mutations in the same gene were identified in evolution experiment 1 (Table 5.4), in both the D23580 and 4/74 pBT1 genetic backgrounds, and one mutation in the same gene was identified in the Tn*5* mutagenesis experiment 2 (Figure 5.6). Because mutations in the *pBT1-0081* gene were found so frequently to rescue the slow growth of the D23580 Δ*cysS*^{pBT1} strain in minimal media in three independent experiments, evidence pointed towards *pBT1-0081* being the key pBT1-associate factor that modulated expression of *cysS*^{chr}.

Evolution experiment 2 provided further evidence of the central role of the *cysS*^{chr} transcript by identifying a nucleotide change in the *cysS*^{chr} promoter region carried by the CysS-P4 mutant. The nucleotide change in the context of the *cysS*^{chr} promoter is in Figure 6.3. The data suggests that the mutated promoter caused an increased level of expression of the *cysS*^{chr} transcript, or altered the ability of a regulatory factor to bind to the promoter region, resulting in faster growth of D23580 in minimal media. These possibilities will be addressed in more detail in Chapter 6.

Taken together, the findings from the Tn*5* mutagenesis and evolution experiments are consistent with the pBT1-associated *pBT1-0081* gene product modulating expression of *cysS*^{chr}, possibly by targeting the *cysS*^{chr} promoter region.

5.6 Discussion

D23580 carries 3 novel plasmids pBT1, pBT2 and pBT3 (Kingsley *et al.*, 2009). These plasmids have been widely conserved in *S.* Typhimurium ST313 Lineage 2 (Pulford *et al.*, 2021). The plasmid pBT1 (~84.5 kb) is much larger than the other two plasmids (pBT2: ~2.5 kb; pBT3: ~2 kb), making it the focus for the identification of novel genes contributing to the virulence or fitness of *Salmonella* (Canals *et al.*, 2019b). Plasmid pBT1 is not essential for growth of D23580 as the plasmid could be deleted from D23580 and did not modulate the growth rate of D23580 (Figure 5.2 B). However, previous transposon insertion sequencing data (available at: https://hactar.shef.ac.uk/D23580/) have revealed that pBT1 plasmid contains a large number of genes that are required for growth and fitness of D23580 (Canals *et al.*, 2019a). One example was the plasmid-encoded *cysS*^{pBT1}, which is required for growth of D23580 in minimal media.

Cysteinyl-tRNA synthetase (CysRS) is one of the aminoacyl-tRNA synthetase enzymes that link specific amino acids to their cognate tRNAs and is essential for the translation process in many bacteria (Baba *et al.*, 2006; Newberry *et al.*, 2002). Previous RNA-seq-based transcriptomic analyses data by the Hinton Lab showed

that D23580 expressed dramatically low levels of *cysS*^{chr} and high levels of the *cysS*^{pBT1} transcript (Canals, *et al.*, 2019b; Figure 5.1). A recent biochemical study found that both the enzyme activity and the enzyme stability of the *cysS*^{chr}-encoded Cysteinyl-tRNA synthetase were much lower than the *cysS*^{pBT1}-encoded Cysteinyl-tRNA synthetase (Canals, *et al.*, 2019a). It has been proposed that the inefficiency of CysRS in the translation process of D23580 leads to the accumulation of uncharged tRNAs that lead to induction of the stringent responses upon nutrient starvation (Agirrezabala *et al.*, 2013; Canals, *et al.*, 2019a).

Nutrient stress plays an important role in bacteria viability in the host and thus affects bacterial virulence (Eisenreich *et al.*, 2010). For example, host microbiota restricts the availability of certain nutrients (Fang *et al.*, 2016). The *Salmonella* pathogen is known to face nutrient stress at many stages of infection (Pradhan *et al.*, 2019), suggesting that the *cysS*^{pBT1}-encoded Cysteinyl-tRNA synthetase could contribute to the ability of D23580 and *S.* Typhimurium ST313 Lineage 2 to survive nutrient starvation in an important niche such as the human body.

Although *cysS*^{pBT1} rescues the replacement of *cysS*^{chr}-associated defect of D23580 in minimal media (Section 5.2), the molecular basis was unclear. In this chapter, two Tn*5* mutagenesis experiments and two experimental evolution approaches were used to search for plasmid pBT1-associated factors that down-regulate *cysS*^{chr} expression.

The mutations that were identified most frequently involved the novel *pBT1-0081* gene. The *pBT1-0081* mutants were found in two different types of experiments (Tn*5* mutagenesis and evolution) and two further independent experiment approaches, namely evolution experiment 1 and 2. The findings are consistent with the *pBT1-0081* gene product playing a key role in the down-regulation of $cysS^{chr}$ transcription.

However, it is possible that the Tn*5* insertion in *pBT1-0081* had a polar effect on the transcription of the downstream gene, *traL* (Figure 5.6). The TraL protein is important for DNA transfer during plasmid conjugation (Makino *et al.*, 1998). RNA-seq-based transcriptomic data (Figure 5.6) showed that the *pBT1-0081* gene was co-transcribed with the intergenic region between *traL* and *pBT1-0081* and also the beginning part of the *traL* gene, raising the possibility that regions outside *pBT1-0081* could also influence $cysS^{chr}$. In the next chapter, further investigations were conducted involving the gene *pBT1-0081* and nearby gene. It will also show that plasmid-encoded factors can influence the $cysS^{chr}$ promoter at the transcriptional level.

A mutation was identified just before the start codon of *traI* that rescued the slow growth of D23580 pBT1^{+} Δ$cysS^{pBT1}$*::cat* strain in minimal media (Table 5.3). Similar to *traL*, the *traI* gene is also associated with DNA transfer and plasmid conjugation, raising the possibility that the conjugation process is linked to the response to

nutrient starvation of D23580 in some way. However, a detailed mechanism has yet to be determined.

A SNP mutation was identified in the promoter region of the gene *cysS*^{chr} that permitted rapid growth of the pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain in minimal media (Table 5.5). The data suggests that the mutation increased the activity of the *cysS*^{chr} promoter, to yield sufficient CysRS^{chr} for growth in minimal media. Alternatively, it is possible that the promoter region of *cysS*^{chr} was the target of pBT1 plasmid-encoded repressor. In this case, the mutation could have prevented the binding of a putative plasmid repressor, leading to the production of normal levels of CysRS^{chr}.

In summary, this chapter shows that a plasmid-encoded cysteinyl-tRNA synthetase, together with factors repressing the chromosomal gene, was identified in the novel plasmid pBT1 of *S.* Typhimurium D23580 and was associated with its fitness. A recent comparative genomic analysis identified 79 of a total 10,000 plasmids that carried an alternate plasmid-encoded aminoacyl-tRNA synthetase (Canals *et al.*, 2019a). This highlights the role of the alternate plasmid-associated essential genes as a topic for further study.

The complexity of this novel system involves the combination of an alternate essential gene on a plasmid coupled with plasmid-encoded factors that are apparently capable of silencing chromosomal genes. Future experiments involving

further Tn*5* mutagenesis or evolution experiments could reveal additional components of the regulatory system. It would be interesting to use RNA-seq to compare transcriptomic data between D23580 WT and Δ*cysS*^{pBT1} mutant in rich LB medium and in minimal media with different carbon source to generate further understanding of this novel plasmid-associated system.

Chapter 6.
Investigating effects of plasmid pBT1-encoded genes on the promoter activity of chromosomal *cysS* in D23580

6.1 Introduction

The chromosomal gene *cysS* (*cysS*^{chr}), encoding cysteinyl-tRNA synthetase, is an essential gene in many bacteria (Baba *et al.*, 2006; Barquist *et al.*, 2013). Transcriptomic data showed that the chromosomal *cysS* gene was down-regulated in *S.* Typhimurium ST313 strain D23580 (Section 5.2; Canals, *et al.*, 2019b).

At the same time, one *cysS* paralog, *cysS*^{pBT1}, is encoded and expressed by the novel pBT1 plasmid in D23580 (Canals *et al.*, 2019b). Experiments in the previous chapter confirmed that the expression of D23580 *cysS*^{chr} was repressed by the presence of pBT1. Two independent approaches, Tn*5* mutagenesis and evolutionary experiment, were used to identify the factors that regulate *cysS*^{chr}. The majority of mutations identified in pBT1 were located in the uncharacterised gene *pBT1-0081*. Interestingly, spontaneous mutations involving the same gene were observed from two different independent evolution experiments and from two different background strains, raising the possibility that the gene *pBT1-0081* was responsible for the down-regulation of *cycS*^{chr} expression. A single nucleotide spontaneous mutation was also identified in the promoter region of the chromosomal *cysS* gene that rescued the slow growth phenotype of D23580 Δ*cysS*^{pBT1}, suggesting that pBT1-encoded factors target the promoter of *cysS* to repress the gene expression.

The aim of this chapter is to investigate whether the gene *pBT-0081* and other pBT1-encoded genes are responsible for the down-regulation of the expression of

cysS^{chr}, and whether the promoter region of *cysS*^{chr} is the target site of the regulation.

6.2 Identification of pBT1-encoded genes involved in the control of *cysS*^{chr} expression

To measure the promoter activity of the chromosomal *cysS*, a GFP-based reporter system was used to infer the expression of *cysS*^{chr}. The pZEP08 plasmid was used for the construction of a transcriptional fusion of *cysS*^{chr}*::gfp* (Figure 6.1A; Hautefort *et al.*, 2003). To study the expression of *cysS*^{chr} in different genetic backgrounds, two plasmid-borne *gfp* transcriptional fusions were constructed by cloning different types of promoter of *cysS*^{chr} upstream of the *gfp*$^{+}$ gene of plasmid pZEP08 (Figure 6.1A): the pP_{cysS}-*gfp* plasmid carried the native promoter *cysS*^{chr}, and the pP_{cysS}^{mut}-*gfp* plasmid carried a mutated version of the promoter, which had been identified as a spontaneous mutant in the evolution experiment (Strain CysS-P4 from Table 5.5). The pZEP08 plasmid has a pBR322 backbone, a ColE1 replicon and a medium copy number (Hautefort, *et al.*, 2003). The method for the plasmid construction was described in Section 2.6.6.

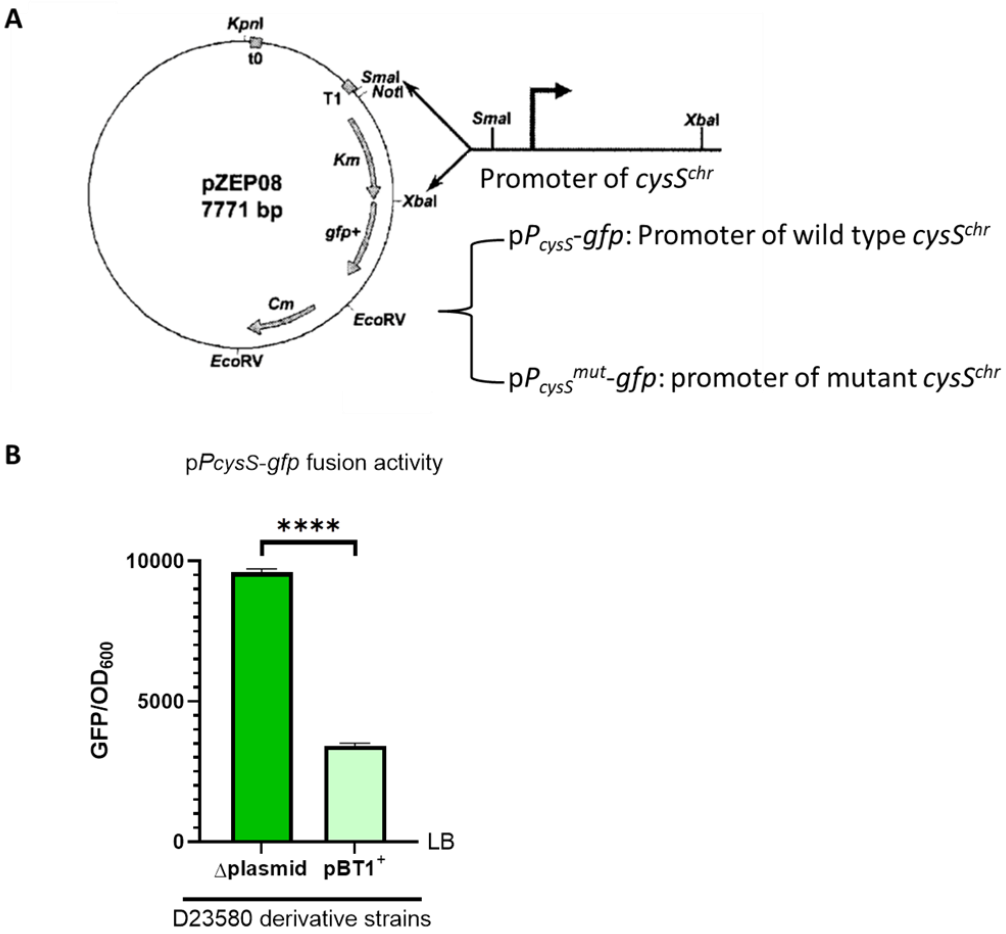
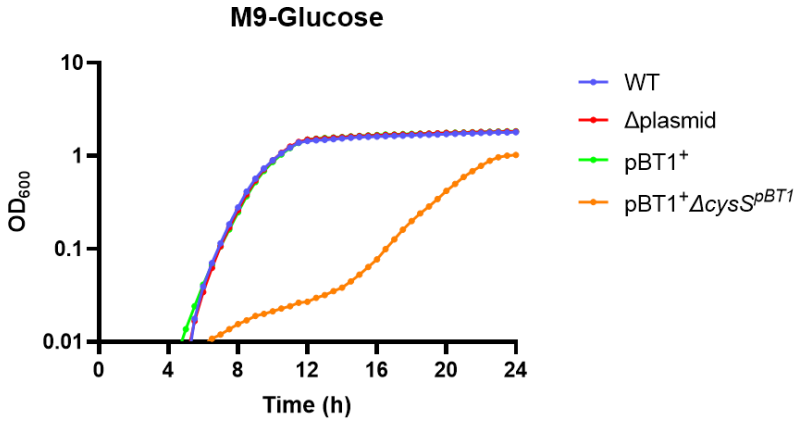


Figure 6.1. Repression of *cysS*^{chr} promoter activity by pBT1. (A) Genetic organisation of the promoter-less plasmid pZEP08 (Hautefort *et al.*, 2003), and the insertion site of the wild type or mutant versions of the *cysS*^{chr} promoter. (B) Comparison of GFP level between D23580 pBT1^{+} and D23580 ∆plasmid (strains SZS010 and SZS011, Table 2.2) with pP_{cysS}-*gfp* after growth in LB media to OD_{600} = 2 (Section 2.6.6), n = 3, ****: P<0.0001. Error bars show Standard Deviation.

To assess the impact of different versions of plasmid pBT1 upon the growth of D23580 in M9 glucose medium, a plasmid-free derivative of D23580 was first constructed (SZS011, Table 2.2) following the protocol described in Section 2.7.3. The growth of D23580 ∆plasmid strain in M9 glucose medium was compared with D23580 WT and the D23580 pBT1^{+} and D23580 pBT1^{+} ∆*cysS*^{pBT1}*::cat* strains (Figure 6.2). Strain D23580 pBT1^{+} ∆*cysS*^{pBT1}*::cat* had an extended lag time in M9

glucose media, as described previously in Chapter 5. Figure 6.2 confirmed that the growth rates of D23580 WT and Δplasmid mutant were similar in M9 glucose medium.



S. Typhimurium Strains	Max Growth Rate	T-test
	M9-glucose medium	
D23580	0.345	
D23580Δplasmid	0.344	ns (P > 0.05)
D23580pBT1+	0.346	ns (P > 0.05)
D23580pBT1+ Δ*cysS*^{pBT1}	0.173	**** (P < 0.0001)

Figure 6.2. Similarities in growth rate of D23580 WT and plasmid-free derivatives, and the growth defect associated with deletion of *cysS*^{pBT1}. Growth curve of D23580 WT, D23580 pBT1+ and D23580 Δplasmid (strains SZS010, SZS011, Table 2.2), and D23580 pBT1+ mutant carrying the pBT1-associated Δ*cysS* mutation (strain SZS014, Table 2.2) in M9 glucose medium at 37°C, n = 6 (Section 2.4.2). The table below shows max growth rate of each strain. T-test was performed to show significance (Ns : P > 0.05; ****: P < 0.0001).

To study the transcription of *cysS*^{chr}, the p*P*$_{cysS}$-*gfp* reporter fusion was introduced into selected strains by transformation, using the method described in Section 2.6.6. As the fusion p*P*$_{cysS}$-*gfp* was transferred into several genetic backgrounds, the abbreviations and genotypes of the key genetic backgrounds used in this chapter are

described in Table 6.1 for clarity.

Table 6.1. Strains abbreviations used in Chapter 6*

Abbreviations	Plasmid content	Strain identifier	Genotype & Construction Strategy
Salmonella Typhimurium D23580			
pBT1^{+}	pBT1	SZS010	D23580 ΔpBT2, ΔpBT3, ΔpSLT-BT The three plasmids were cured using a CRISPR Cas9-based strategy*
∆plasmid	Plasmid-free	SZS011	D23580 ΔpBT1 ΔpBT2 ΔpBT3 ΔpSLT-BT The four plasmids were cured using a CRISPR Cas9-based strategy*
pBT1^{+} *ΔcysS^{pBT1}*	pBT1	SZS014	D23580 ΔpBT2, ΔpBT3, ΔpSLT-BT, *ΔcysS^{pBT1}* The Δ*cysS^{pBT1}* deletion mutant of SZS010
*Details of strain construction are in Section 2.7.3.			

To assess expression of the *cysS^{chr}* promoter, the level of GFP fluorescence of p*P$_{cysS}$-gfp* was compared in the D23580 pBT1^{+} and D23580 ∆plasmid genetic background in LB media (Figure 6.1B) following the protocol from Section 2.6.6. The GFP fluorescence intensity was measured at OD_{600} = 2 in LB as it was the start point of stationary phase according to previous growth curves. Under this condition, the presence of pBT1 caused a 2.8-fold reduction in the GFP levels. The results were consistent with the hypothesis that the regulation of *cysS^{chr}* by pBT1 plasmid was targeting *cysS^{chr}* transcription, and also demonstrated the validity of this p*P$_{cysS}$-gfp* plasmid approach to screen for novel pBT1-encoded factors that might control *cysS^{chr}* expression. As the results in Figure 6.1B already showed the substantial difference between D23580 pBT1^{+} and the D23580 ∆plasmid in LB media, most of

the comparisons between GFP levels in this chapter were conducted using LB media.

6.3 A single nucleotide mutation in the *cysS*^{chr} promoter region promotes transcription in the presence of pBT1.

To investigate the functional impact of the spontaneous mutation at position -22 in the *cysS*^{chr} promoter region (CysS-P4, Table 5.5) on *cysS*^{chr} transcription, the two reporter fusions pP_{cysS}-*gfp* and pP_{cysS}^{mut}-*gfp* were transformed into the D23580 pBT1^{+} and D23580 ∆plasmid strains. In the D23580 ∆plasmid strain, the mutated pP_{cysS}^{mut}-*gfp* fusion had a significantly higher GFP fluorescence (1.2-fold; $P < 0.01$) than the wild type GFP fusion (Figure 6.3), suggesting that the single nucleotide mutation in *cysS*^{chr} promoter region slightly increased *cysS*^{chr} promoter activity. This difference was more accentuated in the pBT1^{+} strain (1.7-fold; $P < 0.001$), indicating that the promoter mutation counteracted the pBT1-mediated down-regulation of *cysS*^{chr} transcription.

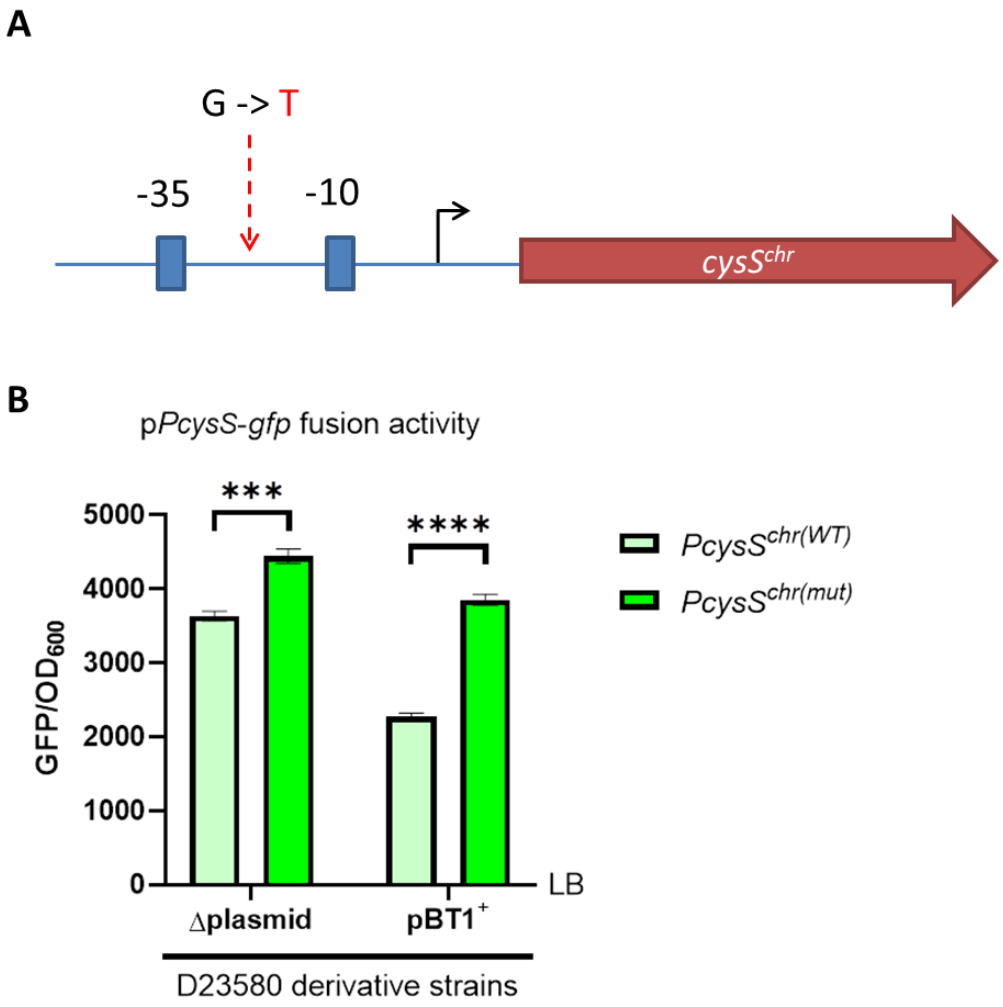


Figure 6.3. Impact of the *cysS*^{chr} single nucleotide mutation on the promoter activity of *cysS*^{chr} in both D23580 Δplasmid and pBT1^{+} strains. (A) Location of mutation in promoter region of *cysS*^{chr} at nucleotide number 637,890 of the D23580 chromosome. The red arrow shows the *cysS*^{chr} gene. The bent black arrow indicates the transcriptional start site (TSS) and blue boxes are -10 and -35 positions (B) Comparison of GFP level between D23580 pBT1^{+} and Δplasmid (strains SZS010 and SZS011, Table 2.2) transformed with pP_{cysS}-*gfp* or pP_{cysS}^{mut}-*gfp* after growth in LB media to OD_{600} = 2 (Section 2.6.6), n = 3, ***: P<0.01, ****: P<0.0001. Error bars show Standard Deviation.

6.4 The *pBT1-0081* gene product influences *cysS*^{chr} promoter activity

Spontaneous mutations in the *pBT1-0081* gene (named *pBT1-0081*^{mut}) were identified frequently in Chapter 5. Similar mutations were identified from two

independent evolution experiments performed by different researchers in three genetic backgrounds: D23580 Δ*cysS*^{pBT1} (Table 5.4), D23580 pBT1^{+} Δ*cysS*^{pBT1} (Table 5.5) and 4/74 pBT1 Δ*cysS*^{pBT1} (Table 5.4). The mutation in *pBT1-0081*^{mut} was a single adenine deletion within the poly (A) region toward 3' end of the coding sequence, shortly before the stop codon TAG (Figure 6.4A). The mutation would increase the protein size by 5 amino acids, which may have structural and functional consequences.

The p*P*$_{cysS}$-*gfp* reporter fusion was used to examine the impact of *pBT1-0081*^{mut} on the modulation of *cysS*^{chr} expression. The D23580 background strains used for evolution experiment were resistant to chloramphenicol (Table 5.4 and 5.5), which is the same antibiotic used to select candidates from transformation of the plasmid p*P*$_{cysS}$-*gfp*. Therefore, the *pBT1-0081*^{mut} mutation was first studied in a chloramphenicol-sensitive 4/74 background (JHL809, Table 5.4). The *pBT1-0081*^{mut} mutant together with the 4/74 pBT1 Δ*cysS*^{pBT1} genetic background strain was transformed with the plasmid p*P*$_{cysS}$-*gfp*, and GFP levels was measured. The mutant with *pBT1-0081*^{mut} showed a significant but modest increase of *cysS*^{chr} promoter activity (Figure 6.4B).

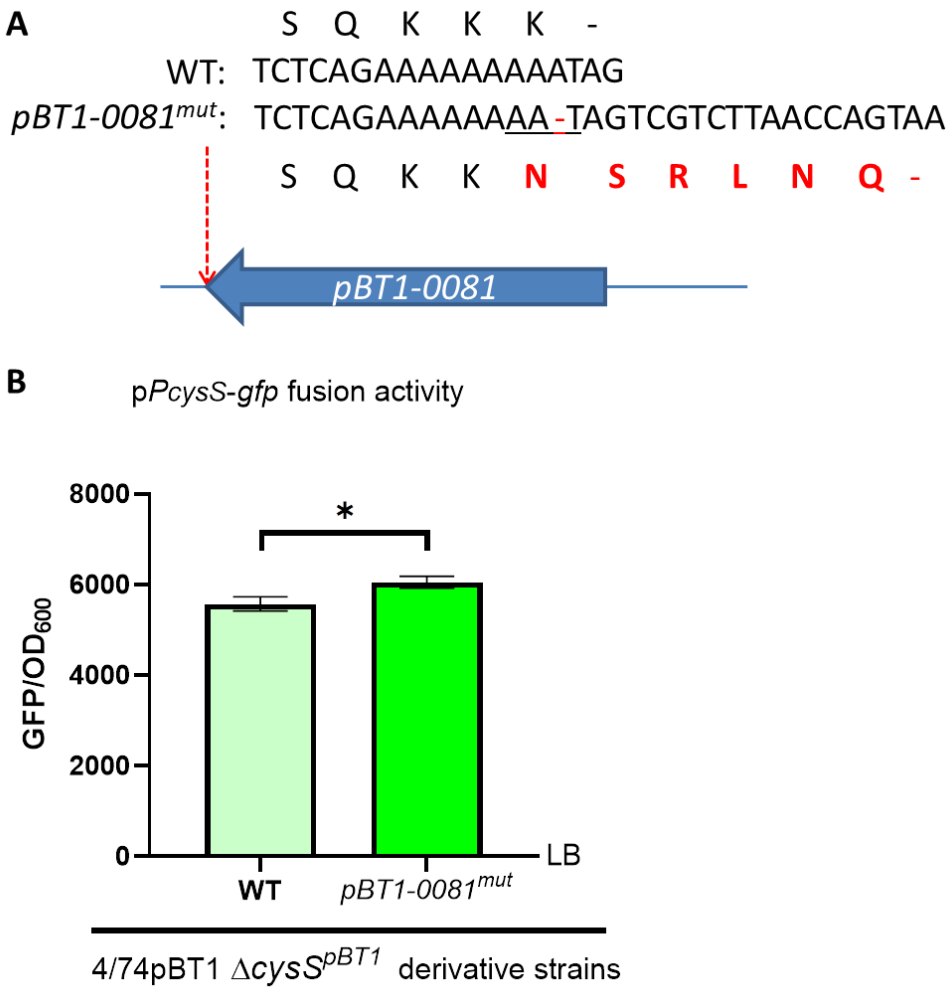


Figure 6.4. Impact of the *pBT1-0081* single nucleotide mutation on promoter activity in 4/74 pBT1 Δ*cysS*^{pBT1}. (A) Location of mutation in *pBT1-0081* at nucleotide number 7,884 of the pBT1 plasmid. Mutation site and the altered nucleotides and amino acids are shown in red. (B) Comparison of GFP levels between 4/74 pBT1 Δ*cysS*^{pBT} WT and its spontaneous *pBT1-0081*^{mut} mutant transformed with pP_{cysS}-*gfp* (strains JH4302, SZS003, Table 2.2) after growth in LB media to OD_{600} = 2 (Section 2.6.6), n = 3, *: P<0.05. Error bars show Standard Deviation.

To learn more about the role of *pBT1-0081* and *cysS*^{pBT1} on the modulation of *cysS*^{chr} expression, the pP_{cysS}-*gfp* reporter plasmid was transformed into both the *cysS*^{pBT1} and *pBT1-0081* deletion mutants of D23580 pBT1^{+}. Comparison of GFP level between D23580 pBT1^{+} and the two deletion mutants showed that the promoter activity modestly but significantly increased in the absence of *cysS*^{pBT1} (P < 0.01),

and decreased slightly in the absence of *pBT1-0081* (P < 0.01; Figure 6.5), indicating that *pBT1-0081* and *cysS*^{pBT1} had opposing effects upon expression of *cysS*^{chr}.

However, in comparison with the 2.8-fold difference on the expression of *cysS*^{chr} between D23580 pBT1^{+} and ∆plasmid (Figure 6.1), the difference between D23580 pBT1^{+} and the two deletion mutants were modest, suggesting that both *pBT1-0081* and *cysS*^{pBT1} may only represent one part of regulatory landscape. The data suggests that the pBT1-encoded down-regulation of *cysS*^{chr} expression is a complex system that involves multiple pBT1 gene products.

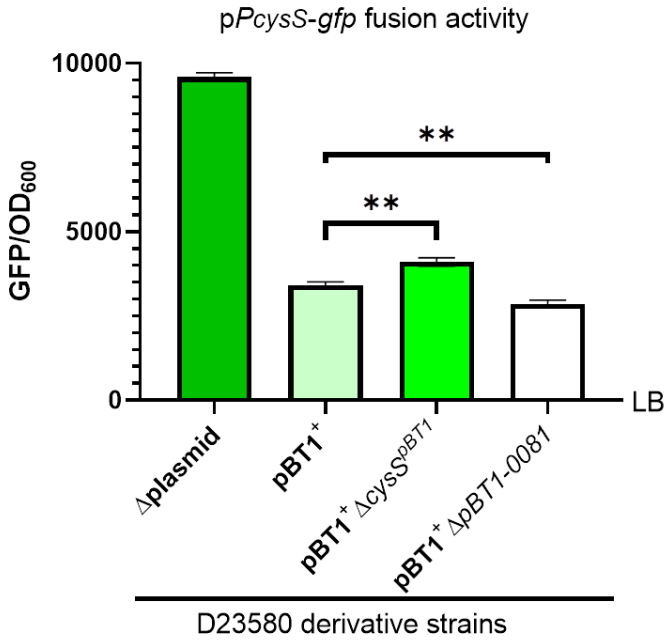
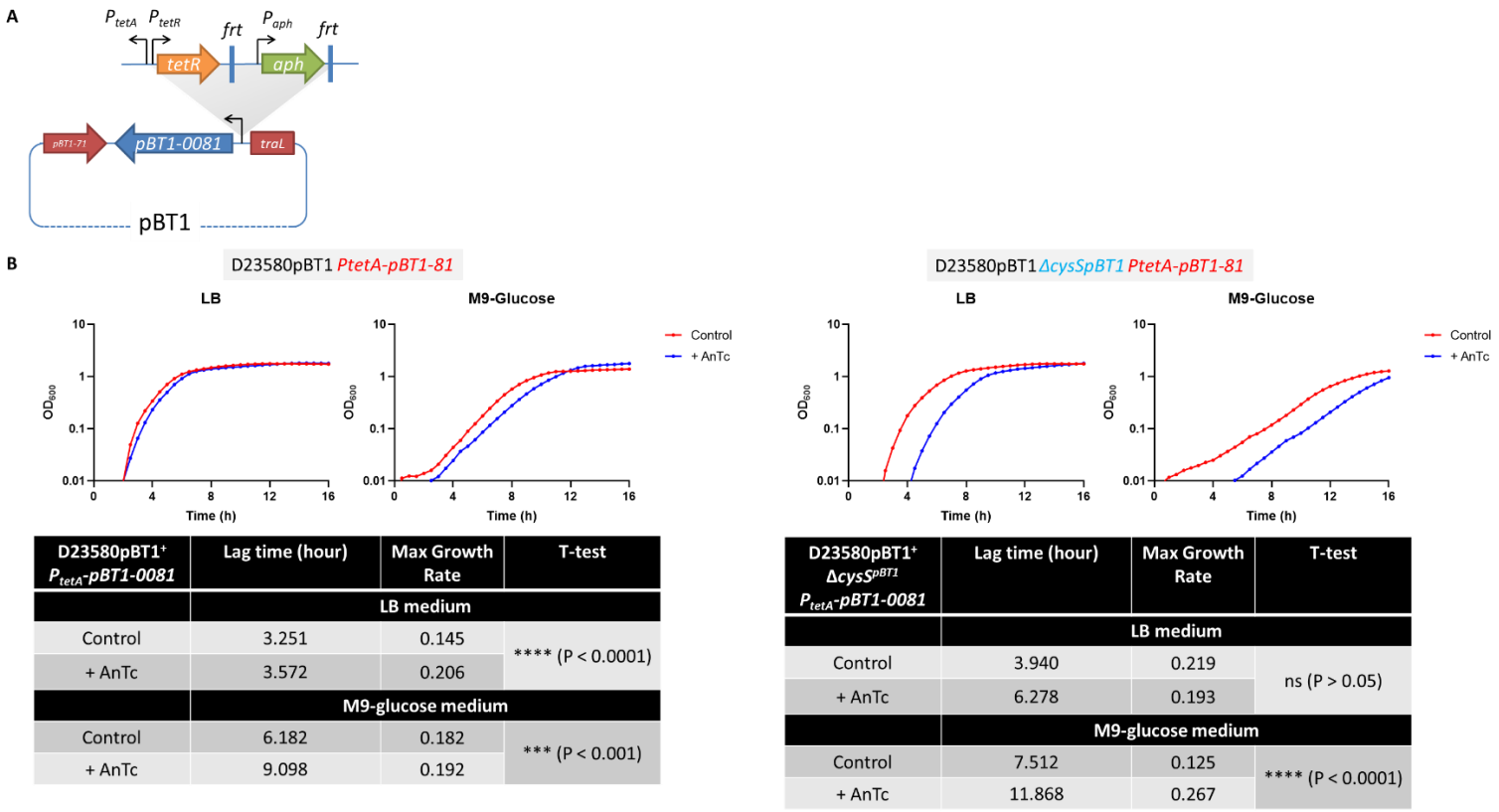


Figure 6.5. Down-regulation of *cysS*^{chr} promoter activity by the pBT1 plasmid is slightly influenced by the *pBT1-0081* and *cysS*^{pBT1} genes. Comparison of GFP levels between D23580 ∆plasmid, D23580 pBT1^{+} and its *cysS*^{pBT1} and *pBT1-0081* deletion mutants (strains SZS010, SZS011, SZS014 and SZS019, Table 2.2) transformed with p*P*$_{cysS}$-*gfp*, after growth in LB media to OD_{600} = 2 (Section 2.6.6), n = 3, **: P<0.01. Error bars show Standard Deviation.

To understand more about the role of *pBT1-0081* on the expression of $cysS^{chr}$ promoter, an inducible *pBT1-0081* was constructed using the *tetR*-P_{tetA} module (Schulte *et al.*, 2019). The detailed genetic context of the resulting *tetR*-P_{tetA}-*pBT1-0081* is shown in Figure 6.6A.

In the *tetR*-P_{tetA} module, expression of the P_{tetA} promoter is triggered by the anhydrotetracycline (AnTc) inducer, and *tetR* encodes the tetracycline repressor that represses the promoter P_{tetA} in the absence of AnTc. Thus, the addition of 500 nM AnTc to the growth medium triggers the expression of *pBT1-0081*, allowing the direct effect of the pBT1-0081 protein on $cysS^{chr}$ regulation to be determined.

The *tetR*-P_{tetA}-*pBT1-0081* module was introduced into D23580 pBT1^{+} and its $cysS^{pBT1}$ deletion mutant. Growth rate was compared in LB and M9 glucose medium with and without 500 nM AnTc for both of the two strains (Figure 6.6B and Figure 6.6C). Following induction of *pBT1-0081*, the two strains showed a longer lag time in both LB and M9 glucose medium; the lag time was increased by more in M9 glucose than in LB.



D23580pBT1⁺ P_{tetA}-*pBT1-0081*	Lag time (hour)	Max Growth Rate	T-test
	LB medium		
Control	3.251	0.145	**** (P < 0.0001)
+ AnTc	3.572	0.206	
	M9-glucose medium		
Control	6.182	0.182	*** (P < 0.001)
+ AnTc	9.098	0.192	

D23580pBT1⁺ Δ*cysS*^{pBT1} P_{tetA}-*pBT1-0081*	Lag time (hour)	Max Growth Rate	T-test
	LB medium		
Control	3.940	0.219	ns (P > 0.05)
+ AnTc	6.278	0.193	
	M9-glucose medium		
Control	7.512	0.125	**** (P < 0.0001)
+ AnTc	11.868	0.267	

Figure 6.6. Induction of *pBT1-0081* expression increased the lag time of D23580. (A) Genetic context of the inducible *pBT1-0081* constructed by integration of the *tetR-P*$_{tetA}$ module. (B) Growth curves of D23580 pBT1⁺ *tetR-P*$_{tetA}$*-pBT1-0081* and D23580 pBT1⁺ Δ*cysS*^{pBT1}*::cat tetR-P*$_{tetA}$*-pBT1-0081* (strains SZS020 and SZS021, Table 2.2) in LB and M9 glucose medium with and without 500 nM AnTc, n = 6. The table below shows lag time and max growth rate for each panel. Lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 from each growth curve (Section 2.4.2). T-test was performed to show significance (ns: P > 0.05; ***: P < 0.001; ****: P < 0.0001).

The pP_{cysS}-*gfp* fusion and the *tetR*-P_{tetA}-*pBT1-0081* module were introduced together into D23580 pBT1^{+} and D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat*. GFP levels were assessed for each of the strains with and without the addition of 500 nM AnTc in both LB and M9-glucose conditions (Figure 6.7). The GFP intensity was measure at OD_{600} = 2 in LB, while in M9-glucose it was measured at OD_{600} = 1 as it was the time point for early stationary phase. Induction of *pBT1-0081* caused a significant increase of GFP level in D23580 pBT1^{+} (Figure 6.7A), demonstrating that *pBT1-0081* induced *cysS*^{chr} promoter activity. This finding was consistent with the decreased GFP level caused by the deletion of *pBT1-0081* (Figure 6.5), proving that *pBT1-0081* plays a key role in the regulation of *cysS*^{chr}.

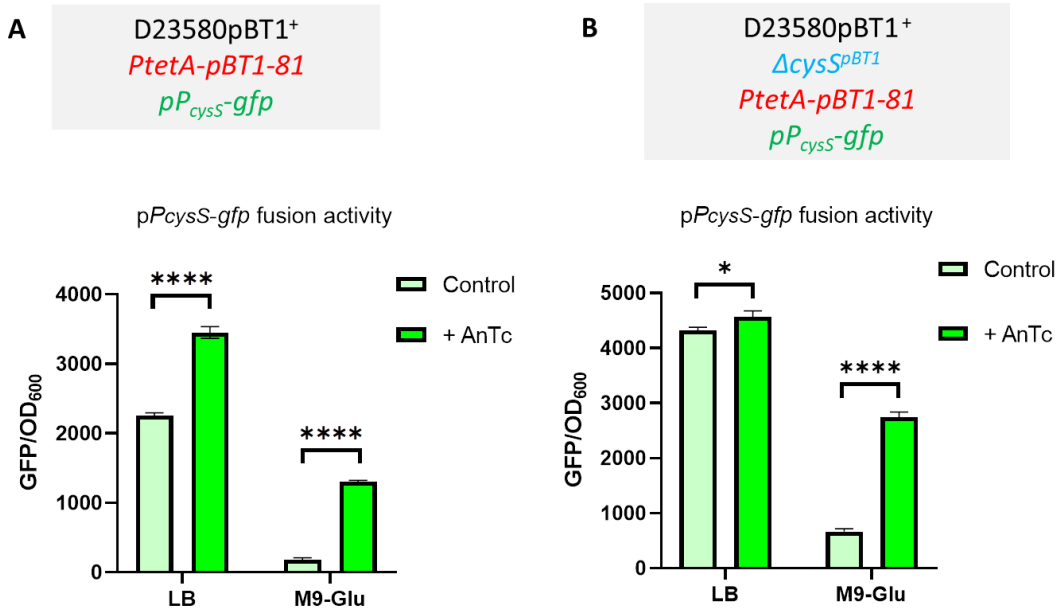


Figure 6.7. Induction of *pBT1-0081* expression up-regulated *cysS*^{chr} promoter activity. (A) Comparison of GFP levels for D23580 pBT1^{+} *tetR-P*$_{tetA}$*-pBT1-0081* transformed with p*P*$_{cysS}$*-gfp* (strain SZS028, Table 2.2) with and without 500 nM AnTc in LB media at OD_{600} = 2 (Section 2.6.6), and in M9 glucose media at OD_{600} = 1, n = 3, ****: P<0.0001. Error bars show Standard Deviation.(B) Comparison of GFP levels for D23580 pBT1^{+} Δ*cysS*^{pBT1} *tetR-P*$_{tetA}$*-pBT1-0081* transformed with p*P*$_{cysS}$*-gfp* (strain SZS029, Table 2.2), with and without 500 nM AnTc in LB media at OD_{600} = 2 (Section 2.6.6), and in M9 glucose media at OD_{600} = 1, n = 3, *: P<0.05, ****: P<0.0001. Error bars show Standard Deviation.

6.5 Induction of *traL* expression infulences *cysS*^{chr} promoter activity

In Section 6.4, the data suggests that multiple pBT1 gene products are involved in the down-regulation of *cysS*^{chr} expression and found one of the genetic determinants is the pBT1-0081 protein. To identify other genetic determinants that control *cysS*^{chr} expression, the nearby genes of *pBT1-0081* were first studied. According to the RNA-seq-based transcriptomic data of D23580 (Figure 5.6; Canals *et al.*, 2019b), the transcription start site of *pBT1-0081* is in the *traL* gene. In addition, the Tn*5* insertion in *pBT1-0081* identified in Section 5.4, which rescued the slow growth of

D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* in minimal media, was in the same orientation as *traL* and may have a polar effect on the transcription of *traL*. The data suggests that the expression of *traL* might be associated with the growth of D23580 in minimal media and affect the transcription of *cysS*^{chr}.

Following the same method to construct an inducible *pBT1-0081*, a *tetR-P*$_{tetA}$*-traL* module was constructed into the promoter region of *traL* of D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat*. The reporter p*P*$_{cysS}$*-gfp* was transformed into the strain and the level of GFP fluorescence was assessed, with and without AnTc in LB media at early stationary phase (Figure 6.8B). Induction of *traL* expression did not cause a statistically significant increase in GFP fluorescence level.

The measurement was repeated using overnight cultures in LB and M9 glucose medium (Figure 6.8C) to investigate the effect of *traL* at late stationary phase. The results showed that induction of *traL* expression caused a significant increase of promoter activity of *cysS*^{chr} ($P < 0.0001$).

In addition the impact of *traL* on growth rate and lag time of D23580 in M9 glucose minimal medium was studied in the D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* genetic background (Figure 6.9). It was apparent that increased expression of *traL* caused a decreased growth rate and an increased lag time in M9 glucose media.

In conclusion, the *traL* gene product regulated expression of the *cysS*^{chr} promoter in a growth phase-dependent manner that was only seen at late stationary phase, and not at early stationary phase.

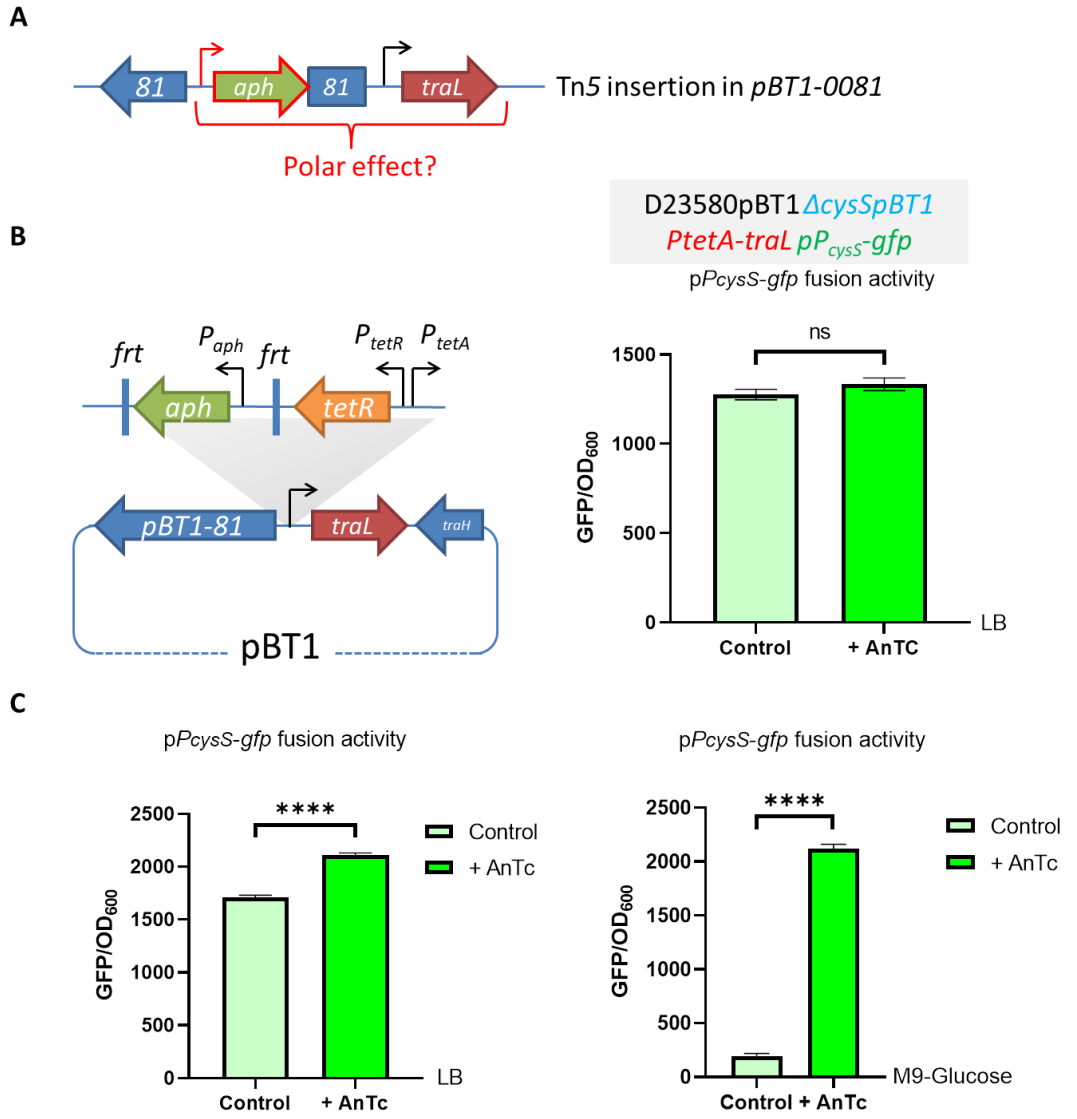
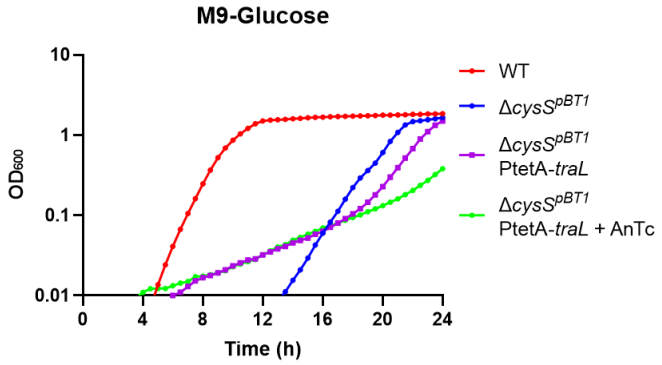


Figure 6.8. Up-regulation of *cysS*^{chr} promoter activity by induction of *traL* expression in late stationary phase. (A) Potential polar effect of Tn*5* insertion in *pBT1-0081* (Table 5.3) upon *traL* transcription. The inserted kanamycin resistance gene and its associated promoter are in red. (B) Genetic context of inducible *traL* by *tetR-P*$_{tetA}$ module, and comparison of GFP levels for D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat tetR-P*$_{tetA}$*-traL* (strain SZS031, Table 2.2) in LB media with and without 500 nM AnTc at OD_{600} = 2 (Section 2.6.6), n = 3, ns: not significant. Error bars show Standard Deviation. (C) Comparison of GFP levels of the same strain as Panel B in LB and M9 glucose medium after overnight growth with and without 500 nM AnTc (Section 2.6.6), n = 3, ****: P<0.0001. Error bars show Standard Deviation.



D23580pBT1⁺ derivative strains	Lag time (hour)	Max Growth Rate	T-test
	M9-glucose medium		
WT	7.568	0.243	
Δ*cysS*^{pBT1}	18.234	0.278	
Δ*cysS*^{pBT1}PtetA-*traL*	19.676	0.261	ns (P > 0.05)
Δ*cysS*^{pBT1}PtetA-*traL* + AnTc	21.574	0.137	**** (P < 0.0001)

Figure 6.9. Lag time is increased by induction of *traL* in a D23580 pBT1⁺ Δ*cysS*^{pBT1}*::cat* background. (A) Growth curves of D23580 pBT1⁺, D23580 pBT1⁺ Δ*cysS*^{pBT1}*::cat* and its *tetR-P*$_{tetA}$*-traL* derivative (strains SZS010, SZS014, SZS030, Table 2.2) in M9 glucose medium, with and without 500 nM AnTc, n = 6. The table below shows lag time and max growth rate for each panel. Lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 from each growth curve (Section 2.4.2). T-test was performed between D23580 pBT1⁺ Δ*cysS*^{pBT1} (SZS014) and the two P_{tetA} mutants to show significance (ns: P > 0.05; P < 0.0001).

6.6 Assigning function to genes around *pBT1-0081*

My previous data suggested that multiple genetic determinants were involved in the complex pBT1-driven repression of *cysS*^{chr} transcription (Section 6.4). In Chapter 5, two independent evolution experiments revealed two spontaneous mutations in the *pBT1-0061* gene that were linked to rapid growth on minimal media (JHL801, Table 5.4 and CysS-L3, Table 5.5). The hypothesis was that additional genetic determinants that control *cysS*^{chr} expression were located between *pBT1-0061* and

traL genes in pBT1.

To understand the function of *pBT1-0061* on the expression of *cysS*^{chr}, the *pBT1-0061* spontaneous mutant (named *pBT1-0061*^{mut}) was studied in a chloramphenicol-sensitive 4/74 pBT1 Δ*cysS*^{pBT1} background (JHL813, Table 5.4). The 4/74 genetic background was used because it has an antibiotic susceptible phenotype (Section 6.4). The *pBT1-0061*^{mut} mutation was an adenine-guanine SNP in the centre of a novel gene *pBT1-0061* (Figure 6.10A). The gene *pBT1-0061* is only 123 bp and is in the same orientation as *pBT1-0081*. This non-synonymous SNP resulted in an amino acid change from isoleucine (I) into arginine (G), which might affect the function of the *pBT1-0061* gene product. Moreover, the RNA-seq data showed some transcripts started from the middle of *pBT1-0061* (Figure 6.10A). The mutation was located between the -10 and -35 motifs of the *pBT1-0051* promoter (Figure 6.10A), suggesting the mutation in *pBT1-0061* could influence the promoter of *pBT1-0051*.

The strain 4/74 pBT1 Δ*cysS*^{pBT1} and its *pBT1-0061*^{mut} mutant derivative (JHL813, Table 5.4) were transformed with the pP_{cysS}-*gfp* plasmid carrying the *cysS*^{chr}*::gfp* transcriptional fusion. In comparison to 4/74 pBT1 Δ*cysS*^{pBT1}, the promoter activity in the mutant carrying *pBT1-0061*^{mut} decreased significantly (Figure 6.10B), suggesting that *pBT1-0061* was involved in the regulation of promoter activity of *cysS*^{chr}.

The facts that the genes upstream (*traL*) and downstream (*pBT1-0061*) of

pBT1-0081 were both associated with *cysS*^{chr} regulation, raised the possibility that the entire *pBT1-0081*-associated region was involved in the regulation of *cysS*^{chr} transcription and the poor growth in M9 glucose medium

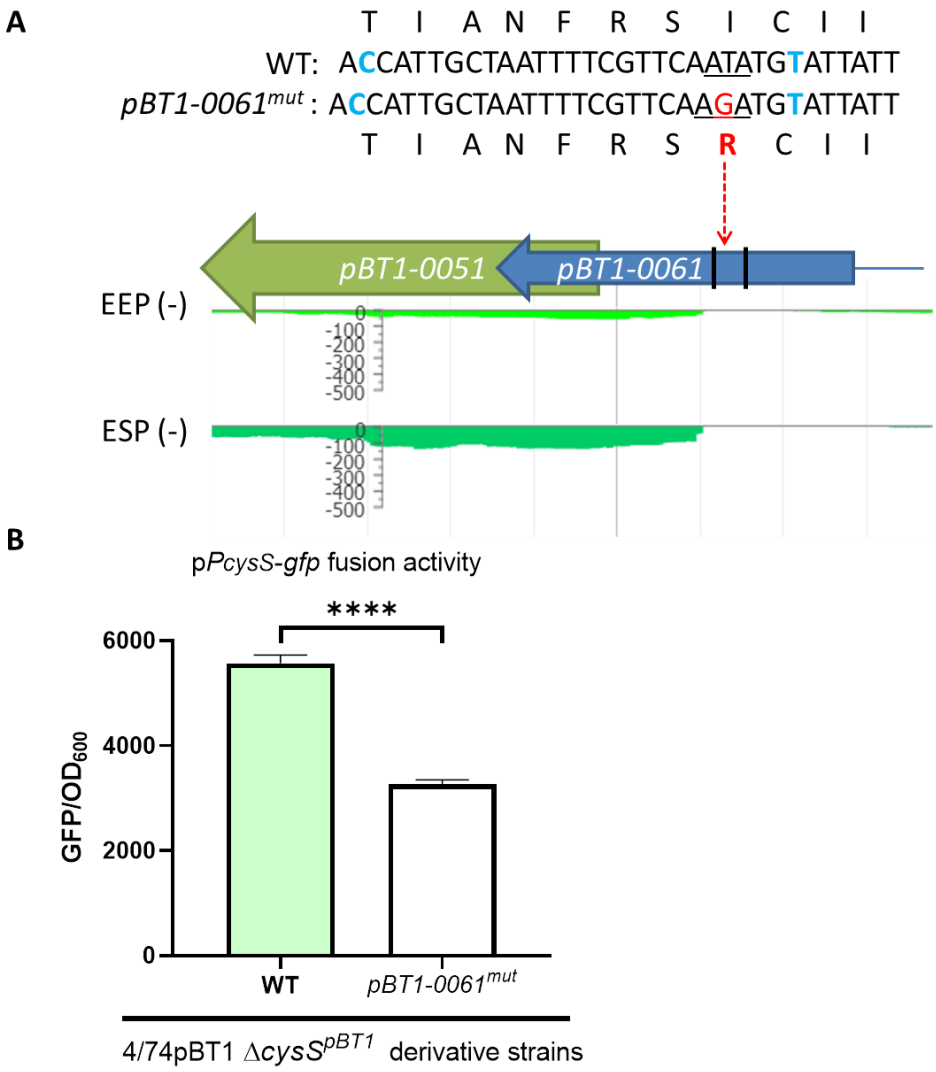


Figure 6.10. Impact of the *pBT1-0061* single nucleotide mutation on *cysS*^{chr} promoter activity in 4/74 pBT1 Δ*cysS*^{pBT1}. (A) Mutation site and the altered nucleotides and amino acids are shown in red, and the -10 and -35 position of the promoter region of *pBT1-0051* are in blue. The black lines in panel B inside *pBT1-0061* indicate the position of the -10 and -35 promoter regions of *pBT1-0051*. Below, the RNA-seq data for this chromosomal region from the EEP and ESP growth condition are shown (Canals *et al.*, 2019b). The scale of the mapped reads was 1 to 500. The genomic context of *pBT1-0061* and the detailed level of transcription are available at JBrowse from the Hinton Lab at https://tinyurl.com/pBT1-0081. (B) Comparison of GFP levels between 4/74 pBT1 Δ*cysS*^{pBT} WT and its *pBT1-0061*^{mut} mutant transformed with p*P*$_{cysS}$-*gfp* (strains JH4302 and SZS002, Table 2.2) after growth in LB media to OD_{600} = 2 (Section 2.6.6), n = 3, ****: P<0.0001. Error bars show Standard Deviation.

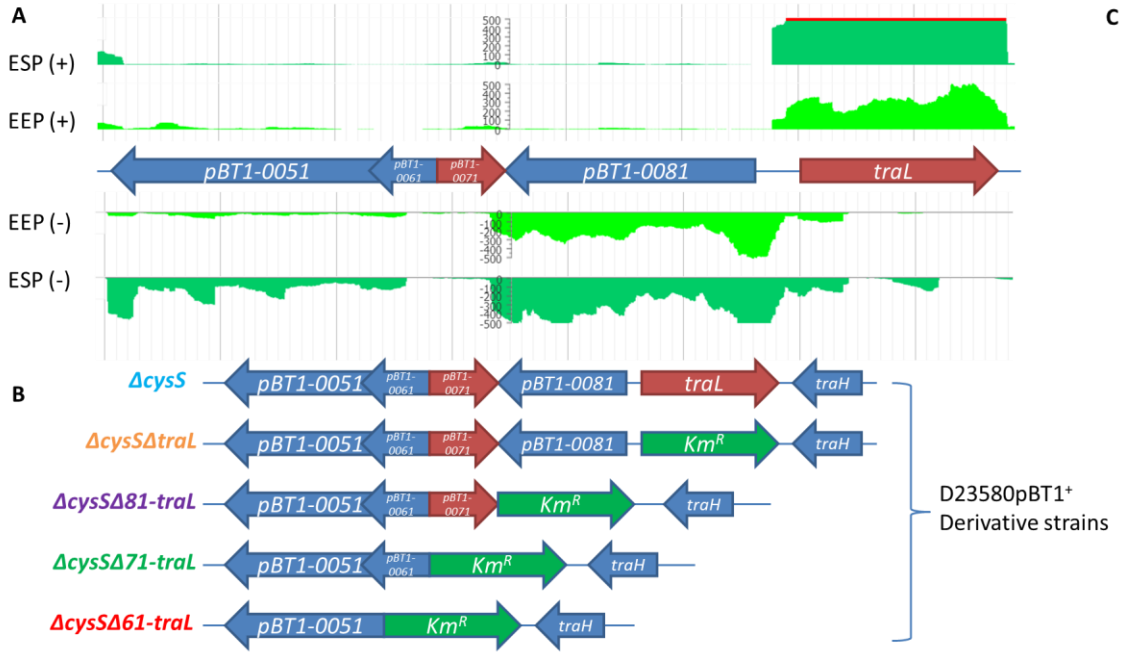
To study the whole genetic region around *pBT1-0081*, genes were sequentially deleted from *traL* to *pBT1-0051*, using D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* as the genetic background via Lambda Red recombination. Although the functions of these genes were unknown, the RNA-seq data (Canals *et al*., 2019b) showed that they were all transcribed during EEP and ESP (Figure 6.11A), suggesting that all these transcripts could have a biological role during these growth stages.

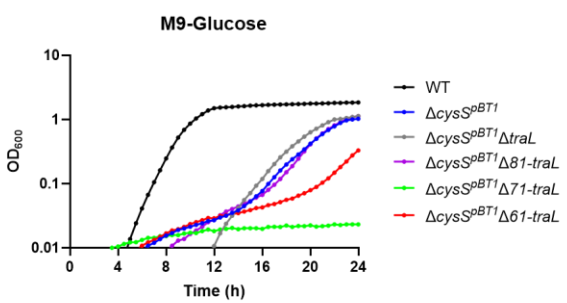
The genetic context of the five mutant derivatives in the *pBT1-0081* region is shown in Figure 6.11B. Despite several attempts, it was not possible to delete the entire *pBT1-0051* to *traL* region of pBT1 using Lambda red recombination. According to data from the transposon-insertion-sequencing (TIS) experiment (Canals *et al.*, 2019a), the *pBT1-0051* gene is not “required” for the survival of D23580 in rich medium. Thus it remains possible that pBT1-0051 is one of the factors in the regulation of *cysS*^{chr} expression.

The growth curves of all the deletion mutants are presented in Figure 6.11C. The D23580 pBT1^{+} Δ*cysS*^{pBT1} Δ*traL* derivative had a shorter lag time than wild type, which was consistent with the results observed in Section 6.5 where the induced *traL* expression lead to a lower growth rate (Figure 6.9). The double deletion mutant of *traL* and *pBT1-0081* showed similar a growth rate to D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat*. The mutant where *pBT1-0071* to *traL* mutant showed an extremely long lag time, suggesting that the *pBT1-0071* gene is important during growth in M9 glucose and is

involved in the *cysS*^{chr} regulation. When the whole region between *pBT1-0061* and *traL* was deleted, the lag time decreased but remained longer than observed for D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat*, suggesting a negative correlation between *pBT1-0061* and *cysS*^{chr} expression.

In summary, all the genes in *pBT1-0051* to *traL* region may have a role in the regulation of *cysS*^{chr} expression. Further experiments involving the construction of deletion mutants of single gene from *traL* to *pBT1-0051* should confirm the regulation caused by each of these genes, and could reveal the mechanisms behind the entire regions.





D23580pBT1+ derivative strains	Lag time (hour)	Max Growth Rate	T-test
	M9-glucose medium		
WT	7.568	0.345	
Δ*cysS*^{pBT1}	18.138	0.173	
Δ*cysS*^{pBT1}Δ*traL*	17.010	0.134	**** (P < 0.0001)
Δ*cysS*^{pBT1}Δ81-*traL*	18.474	0.141	** (P < 0.01)
Δ*cysS*^{pBT1}Δ71-*traL*	ND	0.001	**** (P < 0.0001)
Δ*cysS*^{pBT1}Δ61-*traL*	22.654	0.171	ns (P > 0.05)

Figure 6.11. Impact of sequential gene deletions in the *pBT1-0051 – traL region* on the lag time of D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* in M9 glucose media. (A) Genetic map of the *pBT1-0051 – traL* region. Gene orientation is indicated by red arrows (plus strand) and blue arrows (minus strand). The associated RNA-seq data from the EEP and ESP growth conditions from Canals *et al.,* 2019b are shown. The scale of the mapped reads was 1 to 500, and the detailed data are available at JBrowse from the Hinton Lab at https://tinyurl.com/pBT1-0081. (B) Genetic context of the the *pBT1-0051 – traL* region, showing the five deletion mutants. Gene orientation is indicated by red arrows (plus strand) and blue arrows (minus strand), and green arrows indicate the kanamycin resistance gene that was inserted by Lambda Red recombination. (C) Growth curves of wild type D23580 pBT1^{+} (SZS010, Table 2.2) and each deletion strain in M9 glucose medium. n = 6. The table below shows lag time and max growth rate for each panel. Lag time was assessed by calculating the mean time to reach OD_{600} = 0.2 from each growth curve (Section 2.4.2). T-test was performed between D23580 pBT1^{+} Δ*cysS*^{pBT1} (SZS014) and other mutants to show significance (ns: P > 0.05; **: P < 0.01; ****: P < 0.0001).

6.7 Comparative genomic analysis of the *cysS*^{pBT1} gene

To understand whether there were further examples of plasmid-encoded cysteinyl-tRNA synthetases, the orthologs of *cysS*^{pBT1} were searched using BLASTn using the standard NCBI database (accessed in January 2022). The top 99 sequences from BLASTn identified the *cysS*^{pBT1} gene in 8 *Salmonella* strains other than D23580. The remaining homologues of *cysS*^{pBT1} were found in chromosome of the *Citrobacter* genus, which had a *cysS*^{pBT1} gene with approximately 82% identity at nucleotide level. It has been reported that *Citrobacter koseri* causes invasive diseases (Khashe and Janda, 1996), with some similarities to iNTS disease caused by *Salmonella* ST313. However, *Citrobacter koseri* does not carry a pBT1-like plasmid.

Additional BLASTn searches were done with the *pBT1-0081* gene using the same settings (Section 2.8.2). In total, 9 orthologs of *pBT1-0081* were discovered, which were all plasmid-associated. Among them, 8 of the orthologs were *Salmonella*-associated and the other was from a plasmid pB72C, carried by *Citrobacter* species B72.

The eight *Salmonella* isolates that carried pBT1-like plasmids belonged to the subspecies *enterica*, *salamae* and *diarizonae*, which are generally associated with reptiles such as venomous snakes (Pulford *et al.*, 2019). None of the pBT1-carrying *Salmonella* isolates belonged to the *S.* Enteritidis, *S.* Typhi or *S.* Typhimurium

serovars, which are the major *Salmonella* pathovariants that infect humans. The data suggests that the pBT1 plasmid was transferred from a reptile-associated *Salmonella* serovar to *S.* Typhimurium ST313, possibly in the African setting.

Complete sequences of the 9 plasmids together with the chromosome of *Citrobacter* B72 were aligned with pBT1 sequence to make a Circos plot (Yu *et al.*, 2017; Figure 6.12A). All the 10 sequences and their accession number are listed in Appendix IV. It is evident that the *Citrobacter* plasmid pB72C carries the majority of pBT1 genes while the $cysS^{pBT1}$ gene is located on the chromosome of *Citrobacter* B72. All the *Salmonella* pBT1-like plasmids carried orthologues of both the $cysS^{pBT1}$ and *pBT1-0081* genes, which had 80% nucleotide identity with the pBT1-encoded genes. A maximum likelihood tree (Figure 6.12B) was built from pBT1 and the nine pBT1-like plasmids. The tree showed that the *Citrobacter* plasmid was more closely related to the Most Recent Common Ancestor (MRCA) of pBT1 than the *Salmonella* plasmids.

The results raised the possibility that the $cysS^{pBT1}$ gene moved from the *Citrobacter* chromosome onto a plasmid by recombination, followed by horizontal gene transfer to *Salmonella* via conjugation. One limitation of the study was that the BLAST search was conducted in 2021. Clearly, more *Citrobacter* sequence comparisons will be needed to be done to add weight to this hypothesis.

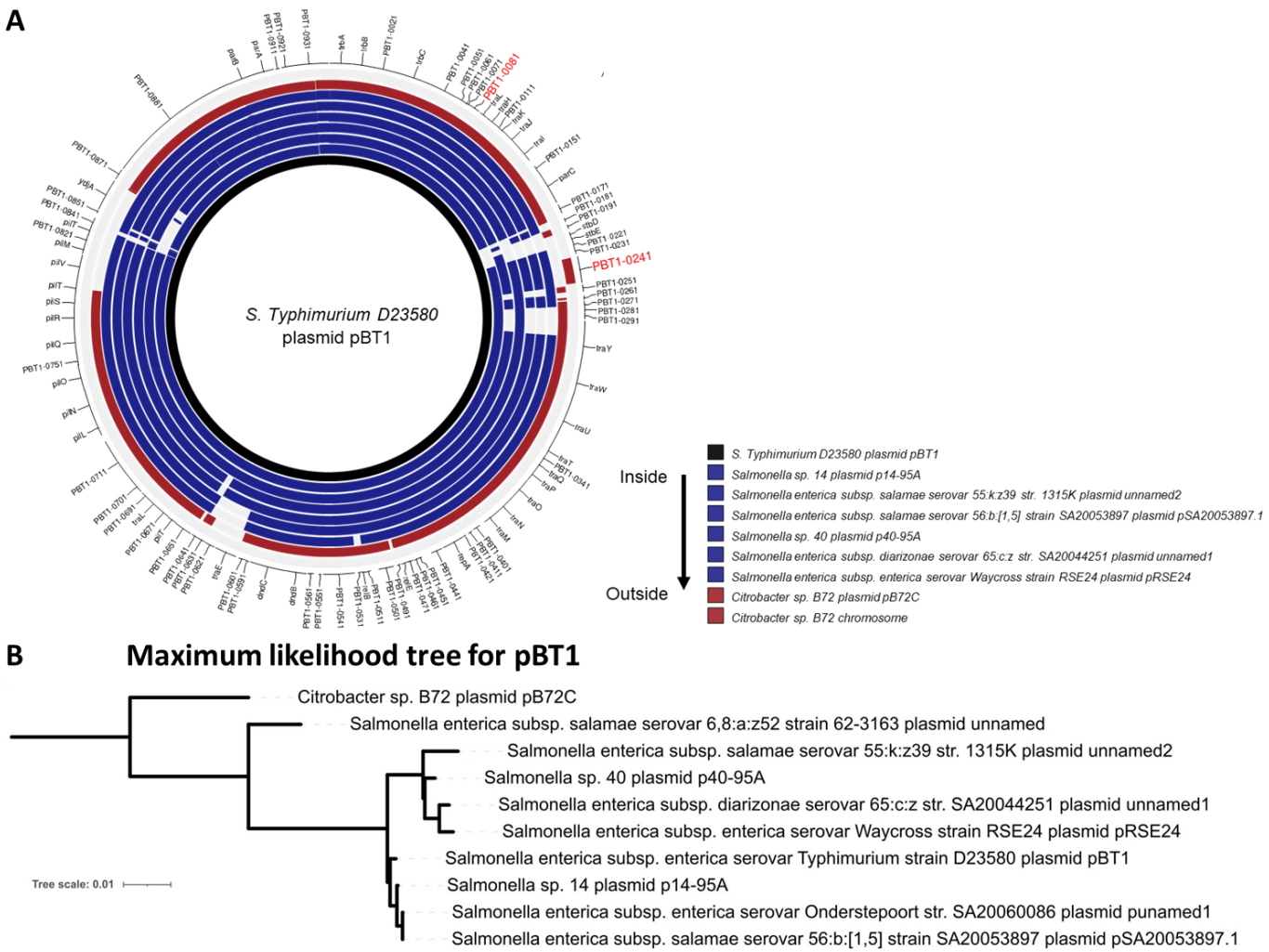


Figure 6.12. Phylogenetic tree for pBT1 and nine pBT1-like plasmids. (A) Circos plot for the alignment of pBT1 and the 9 pBT1-like plasmids plus the chromosome of *Citrobacter* B72 following the method described in Section 2.8.2. The cut-off for the presence/absence of a gene was 80% nucleotide identity. The two key genes *pBT1-0081* and *cysS*^{pBT1} (*pBT1-0241*) are highlighted in red. (B) Maximum likelihood tree of pBT1 and the 9 pBT1-like plasmids. The tree was built using FastTree with the GTR + CAT model (Price *et al.*, 2010; Section 2.8.2). The scale bar shows substitution per site. This comparative genomic analysis was conducted with the assistance of Yan Li.

6.8 Discussion

Expression of the chromosomal *cysS* gene of *S.* Typhimurium ST313 D23580 is repressed by the novel plasmid pBT1, and functionally replaced by the plasmid-encoded $cysS^{pBT1}$ paralog (Canals *et al.*, 2019b). It has recently been shown that plasmid-encoded aminoacyl-tRNA synthetase is associated with a diverse range of Gram-negative and Gram-positive bacteria (Canals *et al.*, 2019a). Study of the mechanisms behind these systems and the biology of plasmid-encoded essential genes can shed a light on a neglected aspect of bacterial evolution. In Chapter 5, experiments that involved two independent approaches were designed to find factors that regulate the expression of $cysS^{chr}$. The most frequent finding from the experiments was a mutation in a novel gene *pBT1-0081* on the plasmid pBT1. The two experiments also identified a single nucleotide change in the promoter region of $cysS^{chr}$ that rescued the slow growth of the D23580 pBT1^{+} $\Delta cysS^{pBT1}$*::cat* strain on minimal media.

By using the transcriptional reporter GFP fusion, it was proved that the single nucleotide mutation in the promoter region of $cysS^{chr}$ lead to an increased promoter activity of $cysS^{chr}$ (Section 6.3). As discussed in Chapter 5, the promoter region is likely to be the target of a pBT1-encoded repressor. The hypothesis was that the mutation modifies the repressor-binding site to prevent the binding of the putative plasmid repressor and maintain a normal level of $CysS^{chr}$ production. This hypothesis awaits experimental verification.

My experiments confirmed that the most frequent *pBT1-0081* spontaneous mutation modulated *cysS*^{chr} transcription and rescued the rapid growth defect of the D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain in minimal media. Following experiments involving the *pBT1-0081* deletion mutant, the transcriptional reporter GFP fusion and the *tetR*-P_{tetA} module, it was shown that over-expression of *pBT1-0081* effected the promoter activity of *cysS*^{chr} (Section 6.4).

However, compared with the dramatic change of *cysS*^{chr} expression caused by the removal of pBT1, the change caused by *pBT1-0081* was modest (Figure 6.5), suggesting that the regulatory pathway for *cysS*^{chr} is a complex system.

Further experiments showed that several gene products encoded within the entire region between *pBT1-0051* and *traL* modulated both expression of *cysS*^{chr} and growth in M9 glucose medium (Section 6.5 and 6.6).

A model representing regulatory links between the gene products encoded by the *pBT1-0051* and *traL* region and the expression of *cysS*^{chr} is shown in Figure 6.13. The data showed that all the genes located between *pBT1-0051* and *traL* were associated with the growth of the D23580 pBT1^{+} Δ*cysS*^{pBT1}*::cat* strain in minimal media.

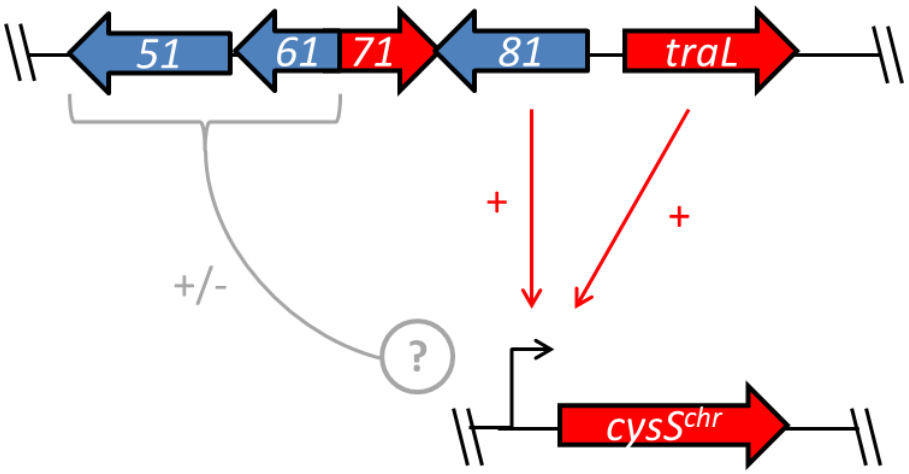


Figure 6.13. Model showing potential regulatory roles of the plasmid-associated *pBT1-0051, pBT1-0061, pBT1-0071, pBT1-0081* and *traL*-encoded proteins upon the promoter activity of the chromosomal *cysS*^{chr} gene. Gene orientation is indicated by red arrows (plus strand) and blue arrows (minus strand). The black arrow indicates the TSS of *cysS*^{chr}. The red line indicates positive-regulation, and the grey line indicates regulation through a hypothetical or unknown mechanism.

However, the regulatory model discovered in this chapter was not consistent with the phenotypes of rapid growth in minimal medium identified in Chapter 5, where *pBT1-0081* mutants from Tn*5* transposon mutagenesis and from experimental evolution both grew rapidly. In addition, the *tetR-P*$_{tetA}$*-pBT1-0081* strain grew faster in minimal medium when *pBT1-0081* was not induced (Figure 6.6). One possible reason was that the *cysS*^{chr}*::gfp* transcriptional fusion used in this chapter had some limitations.

One limitation is that because the pZEP08 plasmid was medium-copy pBR322-derived plasmid (Hautefort *et al.*, 2003), it is possible that plasmid copy number could vary between strains with concomitant impact upon the level of GFP fluorescence. Although the first experiment with the *cysS*^{chr}*::gfp* fusion proved the

validity of the plasmid-encoded GFP system (Section 6.2), in future the findings should be verified by introducing the transcriptional GFP fusion into the chromosome of D23580. Another reason is that the GFP fluorescence instensity was measured at the early stationary phase, and GFP is a stable protein that accumulates over time. Consequently, a strain that grows slowly may have a higher GFP intensity. Further experiments should involve an earlier time point, or the use of non-GFP reporter system. In summary, the relationship between the *cysS*^{chr} expression and growth in minimal media remains to be clarified in further studies.

Another possible limitation in this chapter was that, following Lambda Red recombination, most strains used in this chapter retained the antibiotic cassette to permit selection of the pBT1 plasmid and prevent plasmid loss during growth in stressful conditions. The continued presence of the antibiotic cassette could cause a polar effect on the downstream gene and so have the potential to cause a phenotypic effect.

As genes encoded within the entire region between *pBT1-0051* and *traL* were associated with both the regulation of *cysS*^{chr} and the long lag time in M9 glucose medium, the function of each single gene should be assessed in the future to test the hypothesis is that *pBT1-0081* and *pBT1-0051* are two key genes in the regulatory system. According to the RNA-seq-based transcriptomic data of this genetic region (Figure 6.11A), the annotation of *pBT1-0061* as a distinct gene might be an

annotation error. It remains possible that the mutation of pBT1-0061 affected the expression of *pBT1-0051*, as discussed in Figure 6.10. In addition, the function of the entire genetic region could be tested in an *E. coli* background to determine whether the regulatory mechanism functions in a heterologous host.

The comparative genomic results showed that both the *cysS*^{pBT1} and *pBT1-0081* genes were plasmid-associated in 8 *Salmonella* strains that colonise the gastrointestinal tracts of reptiles. In addition, the *pBT1-0051* to *traL* locus was found in a plasmid pB72C from *Citrobacter* B72, while a *cysS*^{pBT1}-like gene was found in the chromosome of the same strain *Citrobacter* B72 as well as many other *Citrobacter* species. The data suggests that the plasmid pBT1 originated from *Citrobacter* species by the acquisition of the *cysS*^{pBT1} gene from the chromosome, followed by transfer to *Salmonella* that infects reptiles, and finally transferred into African *S.* Typhimurium ST313. These findings offer a possible explanation of an early stage of the evolution of the pBT1 plasmid in context of African *S.* Typhimurium ST313. As more genome sequencing data accumulates over time, the validity of this explanation may become clearer.

It was unfortunate that further experiments were curtailed by the impact of the Covid pandemic and associated lockdowns in the UK in 2020 and 2021.

Chapter 7.
General discussion

7.1 Summary

S. Typhimurium ST313-associated iNTS in sub-Saharan Africa has been a serious health issue (Marchello *et al.*, 2019). To find out the genomic difference between invasive ST313 and well-characterised ST19 that causes gastroenteritis, a previous study identified unique components of the accessory genome of African ST313 L2, including 2 novel prophages BTP1 and BTP5 and 3 novel plasmids pBT1, pBT2 and pBT3 (Kingsley *et al.*, 2009). Mobile genetic elements (MGEs) including prophages and plasmids and the *Salmonella* pathogenicity islands (SPIs) play a key role in the evolution and pathogenesis of *Salmonella* (Mills *et al.*, 1995; Shea *et al.*, 1996; Shintani, 2017). To learn more about pathogenesis and success of ST313 in Africa, I used genomic and transcriptomic data from African L2 representative strain D23580 in comparison with closely related *S.* Typhimurium strains carrying a different accessory genomes. The aim was to investigate the impact of the accessory genome upon the biology of *S.* Typhimurium ST313.

In Chapter 3, the accessory genome of D23580 was compared to the UK variants of ST313. Totaly 4 homologs of BTP1 and 3 homologs of BTP5 were identified. According to previous study (Owen *et al.*, 2017), BTP1 has a extremently high level of spontaneous induction. The CI repressor plays a key role in prophage spontaneous induction (Czyz *et al.*, 2001), and there was speculation that BTP1 might carry an unstable CI repressor. However, the homolog $BTP1^{UK3}$ carried a totally different CI repressor protein, and showed the same high level of spontaneous

induction level as BTP1. Taken together, these findings suggested that the high spontaneous induction level of BTP1 was not directly associated with a particular characteristic of the CI repressor. Because the functions of most of the UK-BTP1 specific genes are still unknown, more research on BTP1 spontaneous induction is required.

BTP5 is incapable of forming visible phage plaques on derivatives of D23580 (Owen *et al.*, 2017). In contrast, it was identified homologs of BTP5 that were capable of plaque formation, with a broad host range. In future, the function of BTP5 and its homologs deserve to be characterized in more detail.

In Chapter 4, D23580 was compared with the novel ST313 L2.2 sublineage which has replaced L2 in Malawian iNTS patients in recent years. My experiment confirmed the reduced level of motility of L2.2 strain D37712, but whether reduced expression of flagellin has any functional significance during the infection of L2.2 remains to be determined.

The properties of D37712 in NonSPI2 growth conditions were the focus of my next sets of experiments. The flow cytometry data showed that, in NonSPI2 media, D37712 had a higher level of SPI2 gene expression and a higher proportion of SPI2-positive cells than D23580. The SPI2 locus encodes Type 3 Secretion System which is important for *Salmonella* pathogenesis, particularly intra-macrophage

survival.

To consider the mechanistic basic for the high expression of these SPI2-associated genes under non-inducing conditions, I built on the current understanding of the control of SPI2 gene expression by a variety of regulatory systems (Figure 7.1; Jiang *et al.*, 2019). The transcriptomic data of D37712 for all the key regulatory genes listed in Figure 7.1 in the NonSPI2 condition are available at SalComD37712-SPI2Regulators (B. Kumwenda *Pers. Comm.*). Several genes were up-regulated by between 3.39 to 1.15-fold compared with D23580 The SPI2 activator proteins were most highly up-regulated in D37712, compared with D23580, including the two-component regulators *ompR*, *phoP* and *ssrB*.

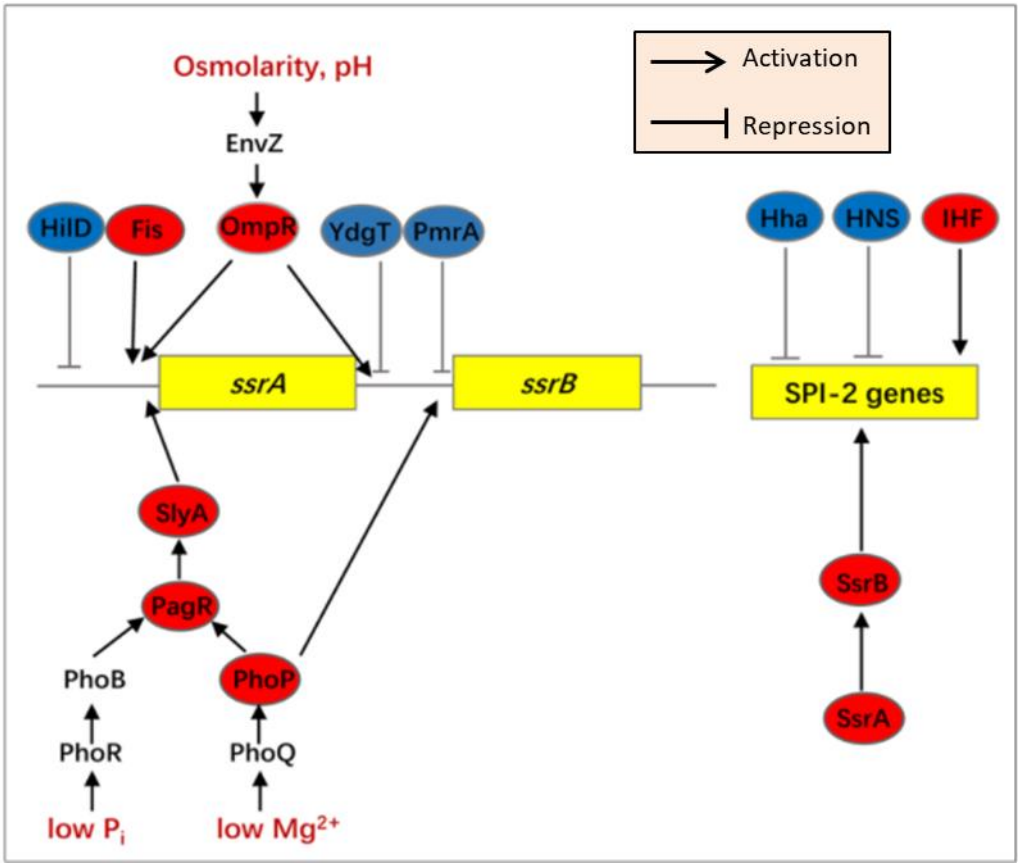


Figure 7.1. Schematic representation of the routes of SPI2 regulation. The figure is modified from Jiang *et al.*, 2019.

The genetic basis for the up-regulation of the SPI2 regulation of D37712 under non-inducing conditions remains to be established. Previous genomic comparison between D23580 and D37712 identified just 26 SNP differences between two strains, but none of them were located in the SPI2 locus. The data suggests that the high expression level of D37712 SPI2 genes in NonSPI2 media is not related to the core-genome associated nucleotide difference between D23580 and D37712, raising the possibility that up-regulation of the SPI2 regulon is driven by differences at the level of the accessory genome. More experiments will be needed to establish the molecular basis of the aberrant high expression of SPI2 genes.

The hypothesis was that the aberrant expression of SPI2 genes in NonSPI2 media would cause a fitness cost of D37712 in NonSPI2 media. In fact, my three different experimental approaches confirmed that D37712 had an increased competitive index in NonSPI2 media, compared with D23580. To conclude, the increased fitness of D37712 cannot be attributed to that differential expression of SPI2 genes. Because genomic comparisons identified no difference in the prophage repertoire between D23580 and D37712 (B. Kumwenda *Pers. Comm.*), the data suggests that the competitive advantage of D37712 over D23580 is associated with the accessory genome and either reflects the presence of the novel plasmid pBT4, or the absence of pBT1, in D37712.

Therefore, in the remaining two chapters, the function of the uncharacterized pBT1 plasmid was investigated, focusing on the functional replacement of the essential chromosomal *cysS*^{chr} gene by a pBT1-encoded paralog *cysS*^{pBT1}. In Chapter 5, the data proved that the expression of *cysS*^{chr} was repressed by pBT1 and identified several potential pBT1-encoded regulators that infulenced *cysS*^{chr} expression.

In Chapter 6, a plasmid-borne GFP reporter fusion syetem was used to show that the pBT1-encoded protein pBT1-0081 influenced the promoter activity of *cysS*^{chr}. My subsequent experiments suggest that the *pBT1-0081*-associated locus was involved in the regulation of *cysS*^{chr} transcription. In conclusion, a complex pBT1-encoded regulatory system controls *cysS*^{chr} transcription which is worthy of future

experimental investigation.

One hypothesis is that there is an unknown pBT1-encoded toxin/antitoxin system that represses *cysS*^{chr}. A toxin/antitoxin (TA) system contains a toxin gene and its cognate antitoxin (Singh *et al.*, 2021; Unterholzner *et al.*, 2013) and TA systems are normally more highly expressed during stress conditions. This hypothesis is supported by a recent study which identified a TA system that targeted another tRNA-synthetase gene, where toxin HipA inhibits the Glu-tRNA-synthetase gene (*gltX*) and an antitoxin HipB inactivates HipA by direct binding (Renbarger *et al.*, 2017).

7.2 Perspective

My findings highlighted the value of studying the accessory genome of *S.* Typhimurium ST313 to understand the success of this pathogen. Clearly, both the prophage BTP1 and the plasmid pBT1 are playing important roles. It was recently reported that BTP1 encoded the BstA phage defense system, and it has been speculated that BstA might provide a selective advantage in certain ecologically niche that would promote the survival of ST313 L2 (Owen *et al.*, 2021). In this context, the high spontaneous induction level of BTP1 represents a key phenotype which needs to be studied further. For example, it has been reported that spontaneous induction of the Stx prophage can directly contribute to pathogenesis of *E. coli* (Livny and Friedman, 2004). The data from my competitive index experiment

that involved ST313 L2 and L2.2 suggests that the pBT1 plasmid is associated with the fitness advantage of L2.2. The possibility that certain pBT1-encoded proteins provide a fitness advantage was supported by a global transposon mutagenesis experiment with D23580 which showed that 36 of the 94 pBT1 plasmid genes were "required" for growth (Canals *et al.*, 2019a). In contrast, only 2 of the 122 genes carried by the pSLT-BT virulence plasmid were reported for growth of *S.* Typhimurium (Canals *et al.*, 2019a).

It has been suggested that the alternate pBT1-encoded aminoacyl-tRNA synthetase might protect ST313 L2 against nutrient starvation during infection (Canals *et al.*, 2019a), but no experimental evidence is available to support this hypothesis. The discovery of plasmid-encoded aminoacyl-tRNA synthetase genes across a wide spectrum of Gram-positive and Gram-negative bacteria (Canals *et al.*, 2019a) suggests that the role of plasmid-associated essential genes in bacterial physiology deserves further investigation.

My investigation of the accessory genome advances our understanding of the evolution of iNTS ST313 in Africa. The comparison of the prophage repertoire of global ST313 revealed that the BTP1^{UK1} prophage was carried by African ST313 L3, UK-ST313 and Brazilian ST313 isolates. This finding is consistent with historical inter-continental transmission of *S.* Typhimurium ST313, an event that has been suggested by others (Seribelli *et al.*, 2020). My comparative genomic analysis on the

ST313 L2-associated pBT1 plasmid has suggests that the *S.* Typhimurium pBT1 plasmid may have originated in *Citrobacter* species before being transferred to reptile-associated *Salmonella*. This possibility would be consistent with the complex evolution of *S.* Typhimurium ST313 (Pulford *et al.*, 2021) which is associated with the epidemic of iNTS disease.

My PhD project focused on linking the accessory genome to the biology of *S.* Typhimurium, to contribute to our understanding the evolution and success of *S.* Typhimurium ST313 in Africa. It is anticipated that an awareness of the bacterial drivers of iNTS disease development will benefit the ongoing vaccine and novel therapeutic approaches being developed to target iNTS in Africa (Piccini and Montomoli, 2020). I am hoping that some of the insights I have gained will prove to be a useful basis for future research.

Bibliography

Achtman, M., Wain, J., Weill, F.-X., Nair, S., Zhou, Z., Sangal, V., Krauland, M. G., Hale, J. L., Harbottle, H., Uesbeck, A., Dougan, G., Harrison, L. H., Brisse, S., & the, S. e. M. s. g. (2012). Multilocus Sequence Typing as a Replacement for Serotyping in *Salmonella* enterica. *PLoS Pathogens*, 8(6), e1002776. https://doi.org/10.1371/journal.ppat.1002776

Ackermann, M., Stecher, B., Freed, N. E., Songhet, P., Hardt, W.-D., & Doebeli, M. (2008). Self-destructive cooperation mediated by phenotypic noise. *Nature*, 454(7207), 987-990. https://doi.org/10.1038/nature07067

Adler, B. A., Kazakov, A. E., Zhong, C., Liu, H., Kutter, E., Lui, L. M., Nielsen, T. N., Carion, H., Deutschbauer, A. M., Mutalik, V. K., & Arkin, A. P. (2021). The genetic basis of phage susceptibility, cross-resistance and host-range in *Salmonella*. *Microbiology*, 167(12). https://doi.org/https://doi.org/10.1099/mic.0.001126

Agirrezabala, X., Fernández, I. S., Kelley, A. C., Cartón, D. G., Ramakrishnan, V., & Valle, M. (2013). The ribosome triggers the stringent response by RelA via a highly distorted tRNA. *EMBO Reports*, 14(9), 811-816. https://doi.org/10.1038/embor.2013.106

Alikhan, N.-F., Petty, N. K., Ben Zakour, N. L., & Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics*, 12(1), 402. https://doi.org/10.1186/1471-2164-12-402

Alikhan, N.-F., Zhou, Z., Sergeant, M. J., & Achtman, M. (2018). A genomic overview of the population structure of *Salmonella*. *PLoS Genetics*, 14(4), e1007261. https://doi.org/10.1371/journal.pgen.1007261

Almeida, F., Seribelli, A. A., da Silva, P., Medeiros, M. I. C., Dos Prazeres Rodrigues, D., Moreira, C. G., Allard, M. W., & Falcão, J. P. (2017). Multilocus sequence typing of *Salmonella* Typhimurium reveals the presence of the highly invasive ST313 in Brazil. *Infection, Genetics and Evolution*, 51, 41-44. https://doi.org/10.1016/j.meegid.2017.03.009

Ashton, P. M., Owen, S. V., Kaindama, L., Rowe, W. P. M., Lane, C. R., Larkin, L., Nair, S., Jenkins, C., de Pinna, E. M., Feasey, N. A., Hinton, J. C. D., & Dallman, T. J. (2017). Public health surveillance in the UK revolutionises our understanding of the invasive *Salmonella* Typhimurium epidemic in Africa. *Genome Medicine*, 9(1), 92. https://doi.org/10.1186/s13073-017-0480-7

Aulicino, A., Antanaviciute, A., Frost, J., Sousa Geros, A., Mellado, E., Attar, M., Jagielowicz, M., Hublitz, P., Sinz, J., Preciado-Llanes, L., Napolitani, G., Bowden, R., Koohy, H., Drakesmith, H., & Simmons, A. (2022). Dual RNA sequencing reveals dendritic cell

reprogramming in response to typhoidal *Salmonella* invasion. *Communications Biology*, 5(1), 111. https://doi.org/10.1038/s42003-022-03038-z

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., Pusch, G. D., Reich, C., Stevens, R., Vassieva, O., Vonstein, V., Wilke, A., & Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, 9(1), 75. https://doi.org/10.1186/1471-2164-9-75

Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L., & Mori, H. (2006). Construction of *Escherichia coli* K-12 in-frame, single-gene knockout mutants: the Keio collection. *Molecular Systems Biology*, 2, 2006.0008. https://doi.org/10.1038/msb4100050

Barquist, L., Langridge, G. C., Turner, D. J., Phan, M. D., Turner, A. K., Bateman, A., Parkhill, J., Wain, J., & Gardner, P. P. (2013). A comparison of dense transposon insertion libraries in the *Salmonella* serovars Typhi and Typhimurium. *Nucleic Acids Research*, 41(8), 4549-4564. https://doi.org/10.1093/nar/gkt148

Baumgart, D. C. & Dignass, A. U. (2002). Intestinal barrier function. *Current Opinion in Clinical Nutrition and Metabolic Care*, 5(6), 685-694. doi: 10.1097/00075197-200211000-00012.

Berg, D. E., & Berg, C. M. (1983). The Prokaryotic Transposable Element Tn*5*. *Bio/Technology*, 1(5), 417-435. https://doi.org/10.1038/nbt0783-417

Beuzón, C. R., & Holden, D. W. (2001). Use of mixed infections with *Salmonella* strains to study virulence genes and their interactions *in vivo*. *Microbes and Infection*, 3(14-15), 1345-1352. https://doi.org/10.1016/s1286-4579(01)01496-4

Bobay, L. M., Rocha, E. P., & Touchon, M. (2013). The adaptation of temperate bacteriophages to their host genomes. *Molecular Biology Evolution*, 30(4), 737-751. https://doi.org/10.1093/molbev/mss279

Bobay, L.-M., Touchon, M., & Rocha, E. P. C. (2014). Pervasive domestication of defective prophages by bacteria. *Proceedings of the National Academy of Sciences*, 111(33), 12127-12132. https://doi.org/doi:10.1073/pnas.1405336111

Branchu, P., Bawn, M., & Kingsley, R. A. (2018). Genome Variation and Molecular Epidemiology of *Salmonella enterica* Serovar Typhimurium Pathovariants. *Infection and Immunity*, 86(8). https://doi.org/10.1128/iai.00079-18

Brown, N. F., Rogers, L. D., Sanderson, K. L., Gouw, J. W., Hartland, E. L., & Foster, L. J.

(2014). A horizontally acquired transcription factor coordinates *Salmonella* adaptations to host microenvironments. *mBio*, 5(5), e01727-01714. https://doi.org/10.1128/mBio.01727-14

Brüssow, H. (2007). Bacteria between protists and phages: from antipredation strategies to the evolution of pathogenicity. *Molecular Microbiology*, 65(3), 583-589. https://doi.org/https://doi.org/10.1111/j.1365-2958.2007.05826.x

Buckle, G. C., Walker, C. L. F., & Black, R. E. (2012). Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. *Journal of Global Health*, 2(1), 010401-010401. https://doi.org/10.7189/jogh.02.010401

Canals, R., Chaudhuri, R. R., Steiner, R. E., Owen, S. V., Quinones-Olvera, N., Gordon, M. A., Baym, M., Ibba, M., & Hinton, J. C. D. (2019a). The fitness landscape of the African *Salmonella* Typhimurium ST313 strain D23580 reveals unique properties of the pBT1 plasmid. *PLoS Pathogens*, 15(9), e1007948. https://doi.org/10.1371/journal.ppat.1007948

Canals, R., Hammarlöf, D. L., Kröger, C., Owen, S. V., Fong, W. Y., Lacharme-Lora, L., Zhu, X., Wenner, N., Carden, S. E., Honeycutt, J., Monack, D. M., Kingsley, R. A., Brownridge, P., Chaudhuri, R. R., Rowe, W. P. M., Predeus, A. V., Hokamp, K., Gordon, M. A., & Hinton, J. C. D. (2019b). Adding function to the genome of African *Salmonella* Typhimurium ST313 strain D23580. *PLoS Biology*, 17(1), e3000059. https://doi.org/10.1371/journal.pbio.3000059

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., & Brüssow, H. (2003). Prophage genomics. *Microbiology Molecular Biology Reviews*, 67(2), 238-276. https://doi.org/10.1128/mmbr.67.2.238-276.2003

Carden, S., Okoro, C., Dougan, G., & Monack, D. (2015). Non-typhoidal *Salmonella* Typhimurium ST313 isolates that cause bacteremia in humans stimulate less inflammasome activation than ST19 isolates associated with gastroenteritis. *Pathogens and Disease*, 73(4). https://doi.org/10.1093/femspd/ftu023

Carroll, A. C., & Wong, A. (2018). Plasmid persistence: costs, benefits, and the plasmid paradox. *Canadian Journal of Microbiology*, 64(5), 293-304. https://doi.org/10.1139/cjm-2017-0609

Carver, T., Berriman, M., Tivey, A., Patel, C., Böhme, U., Barrell, B. G., Parkhill, J., & Rajandream, M.-A. (2008). Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics*, 24(23), 2672-2676. https://doi.org/10.1093/bioinformatics/btn529

Casjens, S. R., Gilcrease, E. B., Winn-Stapley, D. A., Schicklmaier, P., Schmieger, H.,

Pedulla, M. L., Ford, M. E., Houtz, J. M., Hatfull, G. F., & Hendrix, R. W. (2005). The Generalized Transducing *Salmonella* Bacteriophage ES18: Complete Genome Sequence and DNA Packaging Strategy. *Journal of Bacteriology*, 187(3), 1091-1104. https://doi.org/doi:10.1128/JB.187.3.1091-1104.2005

CDC (2015). *Serotypes and the Importance of Serotyping Salmonella.* Available at: https://www.cdc.gov/salmonella/reportspubs/salmonella-atlas/serotyping-importance.html (accessed February 8, 2019)

Chen, Y., Golding, I., Sawai, S., Guo, L., & Cox, E. C. (2005). Population Fitness and the Regulation of *Escherichia coli* Genes by Bacterial Viruses. *PLoS Biology*, 3(7), e229. https://doi.org/10.1371/journal.pbio.0030229

Choi, J.-W., Lee, J., Nishi, K., Kim, Y.-S., Jung, C.-H., & Kim, J.-S. (2008). Crystal Structure of a Minimal Nitroreductase, ydjA, from *Escherichia coli* K12 with and without FMN Cofactor. *Journal of Molecular Biology*, 377, 258-267. https://doi.org/10.1016/j.jmb.2008.01.004

Clare, B. (2021). Inflammasome activation by *Salmonella*. *Current Opinion in Microbiology*, 64, 27-32. https://doi.org/10.1016/j.mib.2021.09.004

Clokie, M. R., Millard, A. D., Letarov, A. V., & Heaphy, S. (2011). Phages in nature. *Bacteriophage*, 1(1), 31-45. https://doi.org/10.4161/bact.1.1.14942

Cortes, M. G., Krog, J., & Balázsi, G. (2019). Optimality of the spontaneous prophage induction rate. *Journal of Theoretical Biology*, 483, 110005. https://doi.org/https://doi.org/10.1016/j.jtbi.2019.110005

Cox, N. A., Berrang, M. E., & Cason, J. A. (2000). *Salmonella* penetration of egg shells and proliferation in broiler hatching eggs–a review. *Poultry Science*, 79(11), 1571-1574. https://doi.org/https://doi.org/10.1093/ps/79.11.1571

Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., Parkhill, J., & Harris, S. R. (2014). Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Research*, 43(3), e15-e15. https://doi.org/10.1093/nar/gku1196

Czyz, A., Los, M., Wrobel, B., & Wegrzyn, G. (2001). Inhibition of spontaneous induction of lambdoid prophages in *Escherichia coli* cultures: simple procedures with possible biotechnological applications. *BMC Biotechnology*, 1, 1-1. https://doi.org/10.1186/1472-6750-1-1

Datsenko, K. A., & Wanner, B. L. (2000). One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proceedings of the National Academy of Sciences*, 97(12), 6640-6645. https://doi.org/doi:10.1073/pnas.120163297

De Cesare, A. (2018). *Salmonella* in Foods: A Reemerging Problem. *Advances in Food and Nutrition Research*, 86, 137-179. https://doi.org/10.1016/bs.afnr.2018.02.007

de Jong, H. K., Parry, C. M., van der Poll, T., & Wiersinga, W. J. (2012). Host-pathogen interaction in invasive Salmonellosis. *PLoS pathogens*, 8(10), e1002933. https://doi.org/10.1371/journal.ppat.1002933

Diep, B., Barretto, C., Portmann, A.-C., Fournier, C., Karczmarek, A., Voets, G., Li, S., Deng, X., & Klijn, A. (2019). *Salmonella* Serotyping; Comparison of the Traditional Method to a Microarray-Based Method and an in silico Platform Using Whole Genome Sequencing Data. *Frontiers in Microbiology*, 10, 2554. https://doi.org/10.3389/fmicb.2019.02554

Doolittle, R. F., Feng, D.-F., Tsang, S., Cho, G., & Little, E. (1996). Determining Divergence Times of the Major Kingdoms of Living Organisms with a Protein Clock. *Science*, 271(5248), 470-477. https://doi.org/doi:10.1126/science.271.5248.470

dos Santos, A. M. P., Ferrari, R. G., & Conte-Junior, C. A. (2019). Virulence Factors in *Salmonella* Typhimurium: The Sagacity of a Bacterium. *Current Microbiology*, 76(6), 762-773. https://doi.org/10.1007/s00284-018-1510-4

Dostál, L., & Schildbach, J. F. (2010). Single-stranded DNA binding by F TraI relaxase and helicase domains is coordinately regulated. *Journal of Bacteriology*, 192(14), 3620-3628. https://doi.org/10.1128/jb.00154-10

Doublet, B., Douard, G., Targant, H., Meunier, D., Madec, J. Y., & Cloeckaert, A. (2008). Antibiotic marker modifications of lambda Red and FLP helper plasmids, pKD46 and pCP20, for inactivation of chromosomal genes using PCR products in multidrug-resistant strains. *Journal of Microbiological Methods*, 75(2), 359-361. https://doi.org/10.1016/j.mimet.2008.06.010

Down, T. A., Piipari, M., & Hubbard, T. J. P. (2011). Dalliance: interactive genome viewing on the web. *Bioinformatics*, 27(6), 889-890. https://doi.org/10.1093/bioinformatics/btr020

Dróżdż, M., Małaszczuk, M., Paluch, E., & Pawlak, A. (2021). Zoonotic potential and prevalence of *Salmonella* serovars isolated from pets. *Infection Ecology and Epidemiology*, 11(1), 1975530. https://doi.org/10.1080/20008686.2021.1975530

Edwards, R. A., Helm, R. A., & Maloy, S. R. (1999). Increasing DNA Transfer Efficiency by Temporary Inactivation of Host Restriction. *BioTechniques*, 26(5), 892-900. https://doi.org/10.2144/99265st02

Eisenreich, W., Dandekar, T., Heesemann, J., & Goebel, W. (2010). Carbon metabolism of intracellular bacterial pathogens and possible links to virulence. *Nature Reviews*

Microbiology, 8(6), 401-412. https://doi.org/10.1038/nrmicro2351

Emamalipour, M., Seidi, K., Zununi Vahed, S., Jahanban-Esfahlan, A., Jaymand, M., Majdi, H., Amoozgar, Z., Chitkushev, L. T., Javaheri, T., Jahanban-Esfahlan, R., & Zare, P. (2020). Horizontal Gene Transfer: From Evolutionary Flexibility to Disease Progression. *Frontiers in Cell and Developmental Biology*, 8, 229. https://doi.org/10.3389/fcell.2020.00229

Eng, S.-K., Pusparajah, P., Ab Mutalib, N.-S., Ser, H.-L., Chan, K.-G., & Lee, L.-H. (2015). *Salmonella*: A review on pathogenesis, epidemiology and antibiotic resistance. *Frontiers in Life Science*, 8(3), 284-293. https://doi.org/10.1080/21553769.2015.1051243

Fàbrega, A., & Vila, J. (2013). *Salmonella enterica* serovar Typhimurium skills to succeed in the host: virulence and regulation. *Clinical Microbiology Reviews*, 26(2), 308-341. https://doi.org/10.1128/cmr.00066-12

Fang, Ferric C., Frawley, Elaine R., Tapscott, T., & Vázquez-Torres, A. (2016). Bacterial Stress Responses during Host Infection. *Cell Host & Microbe*, 20(2), 133-143. https://doi.org/https://doi.org/10.1016/j.chom.2016.07.009

Fass, E., & Groisman, E. A. (2009). Control of *Salmonella* pathogenicity island-2 gene expression. *Current Opinion in Microbiology*, 12(2), 199-204. https://doi.org/10.1016/j.mib.2009.01.004

Feasey, N. A., Dougan, G., Kingsley, R. A., Heyderman, R. S., & Gordon, M. A. (2012). Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa.*The Lancet*, 379(9835), 2489-2499. https://doi.org/10.1016/S0140-6736(11)61752-2

Feasey, N. A., Hadfield, J., Keddy, K. H., Dallman, T. J., Jacobs, J., Deng, X., Wigley, P., Barquist, L., Langridge, G. C., Feltwell, T., Harris, S. R., Mather, A. E., Fookes, M., Aslett, M., Msefula, C., Kariuki, S., Maclennan, C. A., Onsare, R. S., Weill, F. X., Le Hello, S., Smith, A. M., McClelland, M., Desai, P., Parry, C. M., Cheesbrough, J., French, N., Campos, J., Chabalgoity, J. A., Betancor, L., Hopkins, K. L., Nair, S., Humphrey, T. J., Lunguya, O., Cogan, T. A., Tapia, M. D., Sow, S. O., Tennant, S. M., Bornstein, K., Levine, M. M., Lacharme-Lora, L., Everett, D. B., Kingsley, R. A., Parkhill, J., Heyderman, R. S., Dougan, G., Gordon, M. A., & Thomson, N. R. (2016). Distinct *Salmonella* Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. *Nature Genetics*, 48(10), 1211-1217. https://doi.org/10.1038/ng.3644

Figueira, R., & Holden, D. W. (2012). Functions of the *Salmonella* pathogenicity island 2 (SPI-2) type III secretion system effectors. *Microbiology (Reading)*, 158(Pt 5), 1147-1161. https://doi.org/10.1099/mic.0.058115-0

Figueroa-Bossi, N., & Bossi, L. (1999). Inducible prophages contribute to *Salmonella* virulence in mice. *Molecular Microbiology*, 33(1), 167-176. https://doi.org/https://doi.org/10.1046/j.1365-2958.1999.01461.x

Figueroa-Bossi, N., & Bossi, L. (2004). Resuscitation of a Defective Prophage in *Salmonella* Cocultures. *Journal of Bacteriology*, 186(12), 4038-4041. https://doi.org/doi:10.1128/JB.186.12.4038-4041.2004

Fitzsimmons, L. F., Liu, L., Kim, J.-S., Jones-Carson, J., Vázquez-Torres, A., Aballay, A., Swanson, M., & Richardson, A. (2018). *Salmonella* Reprograms Nucleotide Metabolism in Its Adaptation to Nitrosative Stress. *mBio*, 9(1), e00211-00218. https://doi.org/doi:10.1128/mBio.00211-18

Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., & et al. (1995). Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science*, 269(5223), 496-512. https://doi.org/10.1126/science.7542800

Fookes, M., Schroeder, G., Langridge, G., Blondel, C., Mammina, C., Connor, T., Seth-Smith, H., Vernikos, G., Robinson, K., Sanders, M., Petty, N., Kingsley, R., Baumler, A., Nuccio, S.-P., Contreras, I., Santiviago, C., Maskell, D., Barrow, P., Humphrey, T., & Thomson, N. (2011). *Salmonella* bongori Provides Insights into the Evolution of the *Salmonellae*. *PLoS Pathogens*, 7, e1002191. https://doi.org/10.1371/journal.ppat.1002191

Galán, J. E. (2021). *Salmonella* Typhimurium and inflammation: a pathogen-centric affair. *Nature Reviews Microbiology*, 19(11), 716-725. https://doi.org/10.1038/s41579-021-00561-4

Gal-Mor, O., Boyle, E. C., & Grassl, G. A. (2014). Same species, different diseases: how and why typhoidal and non-typhoidal *Salmonella enterica* serovars differ. *Frontiers in Microbiology*, 5, 391-391. https://doi.org/10.3389/fmicb.2014.00391

Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.

Garrity, G., Brenner, D.J., Krieg, N.R. & Staley, J.R. (2007). *Bergey's Manual of Systematic Bacteriology*: Volume 2: The Proteobacteria, Part B: The Gammaproteobacteria. Springer US.

Gilchrist, J. J. & Maclennan, C. A. (2019). Invasive Nontyphoidal *Salmonella* Disease in Afica. *EcoSal Plus*, 8(2). doi: 10.1128/ecosalplus.ESP-0007-2018.

Gogoi, M., Shreenivas, M. M., & Chakravortty, D. (2019). Hoodwinking the Big-Eater to Prosper: The *Salmonella*-Macrophage Paradigm. *Journal of Innate Immunity*, 11(3),

289-299. https://doi.org/10.1159/000490953

Grimont, P., & Weill, F.-X. (2007). *Antigenic Formulae of the Salmonella serovars*, (9th ed.) Paris: WHO Collaborating Centre for Reference and Research on *Salmonella*. Institute Pasteur., 1-166.

Guiney, D. G., & Fierer, J. (2011). The Role of the *spv* Genes in *Salmonella* Pathogenesis. *Frontiers in Microbiology*, 2, 129. https://doi.org/10.3389/fmicb.2011.00129

Günster, R. A., Matthews, S. A., Holden, D. W., Thurston, T. L. M., & Bäumler, A. J. (2017). SseK1 and SseK3 Type III Secretion System Effectors Inhibit NF-κB Signaling and Necroptotic Cell Death in *Salmonella*-Infected Macrophages. *Infection and Immunity*, 85(3), e00010-00017. https://doi.org/doi:10.1128/IAI.00010-17

Hacker, J., & Kaper, J. B. (2000). Pathogenicity islands and the evolution of microbes. *Annual Review of Microbiology*, 54, 641-679. https://doi.org/10.1146/annurev.micro.54.1.641

Hammarlöf, D. L., Kröger, C., Owen, S. V., Canals, R., Lacharme-Lora, L., Wenner, N., Schager, A. E., Wells, T. J., Henderson, I. R., Wigley, P., Hokamp, K., Feasey, N. A., Gordon, M. A., & Hinton, J. C. D. (2018). Role of a single noncoding nucleotide in the evolution of an epidemic African clade of *Salmonella*. *Proceedings of the National Academy of Sciences*, 115(11), E2614-E2623. https://doi.org/doi:10.1073/pnas.1714718115

Haseltine, W. A., & Block, R. (1973). Synthesis of Guanosine Tetra- and Pentaphosphate Requires the Presence of a Codon-Specific, Uncharged Transfer Ribonucleic Acid in the Acceptor Site of Ribosomes. *Proceedings of the National Academy of Sciences*, 70(5), 1564-1568. https://doi.org/doi:10.1073/pnas.70.5.1564

Hautefort, I., Proença, M. J., & Hinton, J. C. (2003). Single-copy green fluorescent protein gene fusions allow accurate measurement of *Salmonella* gene expression in vitro and during infection of mammalian cells. *Applied Environmental Microbiology*, 69(12), 7480-7491. https://doi.org/10.1128/aem.69.12.7480-7491.2003

Hayashi, F., Smith, K. D., Ozinsky, A., Hawn, T. R., Yi, E. C., Goodlett, D. R., Eng, J. K., Akira, S., Underhill, D. M., & Aderem, A. (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor 5. *Nature*, 410(6832), 1099-1103. https://doi.org/10.1038/35074106

Heilbron, K., Toll-Riera, M., Kojadinovic, M., & MacLean, R. C. (2014). Fitness is strongly influenced by rare mutations of large effect in a microbial mutation accumulation experiment. *Genetics*, 197(3), 981-990. https://doi.org/10.1534/genetics.114.163147

Hensel, M., Shea, J. E., Gleeson, C., Jones, M. D., Dalton, E., & Holden, D. W. (1995).

Simultaneous identification of bacterial virulence genes by negative selection. *Science*, 269(5222), 400-403. https://doi.org/10.1126/science.7618105

Herrero-Fresno, A., Wallrodt, I., Leekitcharoenphon, P., Olsen, J. E., Aarestrup, F. M., & Hendriksen, R. S. (2014). The Role of the *st313-td* Gene in Virulence of *Salmonella* Typhimurium ST313. *PloS One*, 9(1), e84566. https://doi.org/10.1371/journal.pone.0084566

Hersch, S. J., Radan, B., Ilyas, B., Lavoie, P., Navarre, W. W., & Metcalf, W. W. (2021). Stress-Induced Block in Dicarboxylate Uptake and Utilization in *Salmonella enterica* Serovar Typhimurium. *Journal of Bacteriology*, 203(9), e00487-00420. https://doi.org/doi:10.1128/JB.00487-20

Hibbing, M. E., Fuqua, C., Parsek, M. R., & Peterson, S. B. (2010). Bacterial competition: surviving and thriving in the microbial jungle. *Nature Reviews of Microbiology*, 8(1), 15-25. https://doi.org/10.1038/nrmicro2259

Hiyoshi H., Tiffany C. R., Bronner D. N. & Bäumler, A. J. (2018) Typhoidal *Salmonella* serovars: ecological opportunity and the evolution of a new pathovar, *FEMS Microbiology Reviews*, 42(4), 527-541, https://doi.org/10.1093/femsre/fuy024

Ho, T. D., Figueroa-Bossi, N., Wang, M., Uzzau, S., Bossi, L., & Slauch, J. M. (2002). Identification of GtgE, a novel virulence factor encoded on the Gifsy-2 bacteriophage of *Salmonella enterica* serovar Typhimurium. *Journal of Bacteriology*, 184(19), 5234-5239. https://doi.org/10.1128/JB.184.19.5234-5239.2002

Hoess, R. H., & Landy, A. (1978). Structure of the lambda att sites generated by int-dependent deletions. *Proceedings of the National Academy of Sciences*, 75(11), 5437-5441. https://doi.org/doi:10.1073/pnas.75.11.5437

Honeycutt, J. D., Wenner, N., Li, Y., Brewer, S. M., Massis, L. M., Brubaker, S. W., Chairatana, P., Owen, S. V., Canals, R., Hinton, J. C. D., & Monack, D. M. (2020). Genetic variation in the MacAB-TolC efflux pump influences pathogenesis of invasive *Salmonella* isolates from Africa. *PLoS Pathogens*, 16(8), e1008763. https://doi.org/10.1371/journal.ppat.1008763

Hutchison, C. A., Merryman, C., Sun, L., Assad-Garcia, N., Richter, R. A., Smith, H. O., Glass, J. I., & Silhavy, T. J. (2019). Polar Effects of Transposon Insertion into a Minimal Bacterial Genome. *Journal of Bacteriology*, 201(19), e00185-00119. https://doi.org/doi:10.1128/JB.00185-19

Ilyas, B., Mulder, D. T., Little, D. J., Elhenawy, W., Banda, M. M., Pérez-Morales, D., Tsai, C. N., Chau, N. Y. E., Bustamante, V. H., & Coombes, B. K. (2018). Regulatory Evolution

Drives Evasion of Host Inflammasomes by *Salmonella* Typhimurium. *Cell Reports*, 25(4), 825-832.e825. https://doi.org/https://doi.org/10.1016/j.celrep.2018.09.078

Ilyas, B., Tsai, C. N., & Coombes, B. K. (2017). Evolution of *Salmonella*-Host Cell Interactions through a Dynamic Bacterial Genome. *Frontiers in Cellular and Infection Microbiology*, 7, 428. https://doi.org/10.3389/fcimb.2017.00428

Jackson, R. W., Vinatzer, B., Arnold, D. L., Dorus, S., & Murillo, J. (2011). The influence of the accessory genome on bacterial pathogen evolution. *Mobile Genetic Elements*, 1(1), 55-65. https://doi.org/10.4161/mge.1.1.16432

Jajere, S. M. (2019) A review of Salmonella enterica with particular focus on the pathogenicity and virulence factors, host specificity and antimicrobial resistance including multidrug resistance. *Veterinary world*, 12(4), 504–521. https://doi.org/10.14202/vetworld.2019.504-521

Jiang, W., Bikard, D., Cox, D., Zhang, F., & Marraffini, L. A. (2013). RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature Biotechnology*, 31(3), 233-239. https://doi.org/10.1038/nbt.2508

Johnson, T. J., & Nolan, L. K. (2009). Pathogenomics of the virulence plasmids of *Escherichia coli*. *Microbiology and Molecular Biology Reviews*, 73(4), 750-774. https://doi.org/10.1128/mmbr.00015-09

Khashe, S., & Janda, J. M. (1996). Iron utilization studies in *Citrobacter* species. *FEMS Microbiology Letters*, 137(2-3), 141-146. https://doi.org/10.1111/j.1574-6968.1996.tb08096.x

Kingsley, R. A., Msefula, C. L., Thomson, N. R., Kariuki, S., Holt, K. E., Gordon, M. A., Harris, D., Clarke, L., Whitehead, S., Sangal, V., Marsh, K., Achtman, M., Molyneux, M. E., Cormican, M., Parkhill, J., MacLennan, C. A., Heyderman, R. S., & Dougan, G. (2009). Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Research*, 19(12), 2279-2287. https://doi.org/10.1101/gr.091017.109

Kintz, E., Davies, M. R., Hammarlöf, D. L., Canals, R., Hinton, J. C., & van der Woude, M. W. (2015). A BTP1 prophage gene present in invasive non-typhoidal *Salmonella* determines composition and length of the O-antigen of the lipopolysaccharide. *Molecular Microbiology*, 96(2), 263-275. https://doi.org/10.1111/mmi.12933

Koskiniemi, S., Pränting, M., Gullberg, E., Näsvall, J., & Andersson, D. I. (2011). Activation of cryptic aminoglycoside resistance in *Salmonella enterica*. *Molecular Microbiology*, 80(6), 1464-1478. https://doi.org/https://doi.org/10.1111/j.1365-2958.2011.07657.x

Kottara, A., Hall, J. P. J., Harrison, E., & Brockhurst, M. A. (2018). Variable plasmid fitness effects and mobile genetic element dynamics across *Pseudomonas* species. *FEMS Microbiology Ecology*, 94(1). https://doi.org/10.1093/femsec/fix172

Kozlov, A. M., Darriba, D., Flouri, T., Morel, B., & Stamatakis, A. (2019). RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*, 35(21), 4453-4455. https://doi.org/10.1093/bioinformatics/btz305

Kröger, C., Colgan, A., Srikumar, S., Händler, K., Sivasankaran, S. K., Hammarlöf, D. L., Canals, R., Grissom, J. E., Conway, T., Hokamp, K., & Hinton, J. C. (2013). An infection-relevant transcriptomic compendium for *Salmonella enterica* Serovar Typhimurium. *Cell Host & Microbe*, 14(6), 683-695. https://doi.org/10.1016/j.chom.2013.11.010

Kropinski, A. M., Sulakvelidze, A., Konczy, P., & Poppe, C. (2007). *Salmonella* phages and prophages--genomics and practical aspects. *Methods in Molecular Biology*, 394, 133-175. https://doi.org/10.1007/978-1-59745-512-1_9

Kutter E. (2009) *Phage Host Range and Efficiency of Plating*. In: Clokie M.R., Kropinski A.M. (eds) Bacteriophages. Methods in Molecular Biolog, vol 501. Humana Press. https://doi.org/10.1007/978-1-60327-164-6_14

Lamas, A., Miranda, J. M., Regal, P., Vázquez, B., Franco, C. M., & Cepeda, A. (2018). A comprehensive review of non-enterica subspecies of *Salmonella enterica*. *Microbiological Research*, 206, 60-73. https://doi.org/https://doi.org/10.1016/j.micres.2017.09.010

Langridge, G. C., Phan, M.-D., Turner, D. J., Perkins, T. T., Parts, L., Haase, J., Charles, I., Maskell, D. J., Peters, S. E., Dougan, G., Wain, J., Parkhill, J., & Turner, A. K. (2009). Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Research*, 19(12), 2308-2316. https://doi.org/10.1101/gr.097097.109

Lara-Tejero, M., & Galán, J. E. (2009). *Salmonella enterica* Serovar Typhimurium Pathogenicity Island 1-Encoded Type III Secretion System Translocases Mediate Intimate Attachment to Nonphagocytic Cells. *Infection and Immunity*, 77(7), 2635-2642. https://doi.org/doi:10.1128/IAI.00077-09

Larsen, R. A., Wilson, M. M., Guss, A. M., & Metcalf, W. W. (2002). Genetic analysis of pigment biosynthesis in Xanthobacter autotrophicus Py2 using a new, highly efficient transposon mutagenesis system that is functional in a wide variety of bacteria. *Archives of Microbiology*, 178(3), 193-201. https://doi.org/10.1007/s00203-002-0442-2

Laurent, F., Lelièvre, H., Cornu, M., Vandenesch, F., Carret, G., Etienne, J., & Flandrois, J.-P. (2001). Fitness and competitive growth advantage of new gentamicin-susceptible MRSA clones spreading in French hospitals. *Journal of Antimicrobial Chemotherapy*, 47(3),

277-283. https://doi.org/10.1093/jac/47.3.277

Lauritsen, I., Porse, A., Sommer, M. O. A., & Nørholm, M. H. H. (2017). A versatile one-step CRISPR-Cas9 based approach to plasmid-curing. *Microbial Cell Factories*, 16(1), 135. https://doi.org/10.1186/s12934-017-0748-z

Lemire, S., Figueroa-Bossi, N., & Bossi, L. (2008). A Singular Case of Prophage Complementation in Mutational Activation of *recET* Orthologs in *Salmonella enterica* Serovar Typhimurium. *Journal of Bacteriology*, 190(20), 6857-6866. https://doi.org/doi:10.1128/JB.00769-08

Letunic, I., & Bork, P. (2007). Interactive Tree Of Life (iTOL): an online tool for phylogenetic tree display and annotation. *Bioinformatics*, 23(1), 127-128. https://doi.org/10.1093/bioinformatics/btl529

Lim, S., Kim, B., Choi, H.-S., Lee, Y., & Ryu, S. (2006). Fis is required for proper regulation of ssaG expression in *Salmonella enterica* serovar Typhimurium. *Microbial Pathogenesis*, 41(1), 33-42. https://doi.org/https://doi.org/10.1016/j.micpath.2006.03.005

Lindberg, A. A., Sarvas, M., & Mäkelä, P. H. (1970). Bacteriophage attachment to the somatic antigen of *salmonella*: effect of o-specific structures in leaky R mutants and s, t1 hybrids. *Infection and Immunity*, 1(1), 88-97. https://doi.org/10.1128/iai.1.1.88-97.1970

Livny, J., & Friedman, D. I. (2004). Characterizing spontaneous induction of Stx encoding phages using a selectable reporter system. *Molecular Microbiology*, 51(6), 1691-1704. https://doi.org/https://doi.org/10.1111/j.1365-2958.2003.03934.x

Llagostera, M., Barbé, J., & Guerrero, R. (1986). Characterization of SE1, a New General Transducing Phage of *Salmonella* typhimurium. *Microbiology*, 132(4), 1035-1041. https://doi.org/https://doi.org/10.1099/00221287-132-4-1035

Löber, S., Jäckel, D., Kaiser, N., & Hensel, M. (2006). Regulation of *Salmonella* pathogenicity island 2 genes by independent environmental signals. *International Journal of Medical Microbiology*, 296(7), 435-447.

Löffler, F. (1892). Ueber Epidemieen unter den im hygienischen Institute zu Greifswald gehaltenen Mäusen und über die Bekämpfung der Feldmausplage. *Zentralblatt für Bakteriologie, Parasitenkunde und Infektionskrankheiten*.11(5), 130–141.

Lofton, H., Anwar, N., Rhen, M., & Andersson, D. I. (2014). Fitness of *Salmonella* mutants resistant to antimicrobial peptides. *Journal of Antimicrobial Chemotherapy*, 70(2), 432-440. https://doi.org/10.1093/jac/dku423

Lucchini, S., Rowley, G., Goldberg, M. D., Hurd, D., Harrison, M., & Hinton, J. C. D. (2006).

H-NS Mediates the Silencing of Laterally Acquired Genes in Bacteria. *PLoS Pathogens*, 2(8), e81. https://doi.org/10.1371/journal.ppat.0020081

MacKenzie, K. D., Wang, Y., Musicha, P., Hansen, E. G., Palmer, M. B., Herman, D. J., Feasey, N. A., & White, A. P. (2019). Parallel evolution leading to impaired biofilm formation in invasive *Salmonella* strains. *PLOS Genetics*, 15(6), e1008233. https://doi.org/10.1371/journal.pgen.1008233

Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M., & Spratt, B. G. (1998). Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences*, 95(6), 3140-3145. https://doi.org/doi:10.1073/pnas.95.6.3140

Majowicz, S. E., Musto, J., Scallan, E., Angulo, F. J., Kirk, M., O'Brien, S. J., Jones, T. F., Fazil, A., Hoekstra, R. M., & Studies, f. t. I. C. o. E. D. B. o. I. (2010). The Global Burden of Nontyphoidal *Salmonella* Gastroenteritis. *Clinical Infectious Diseases*, 50(6), 882-889. https://doi.org/10.1086/650733

Makino, K., Ishii, K., Yasunaga, T., Hattori, M., Yokoyama, K., Yutsudo, C. H., Kubota, Y., Yamaichi, Y., Iida, T., Yamamoto, K., Honda, T., Han, C.-G., Ohtsubo, E., Kasamatsu, M., Hayashi, T., Kuhara, S., & Shinagawa, H. (1998). Complete Nucleotide Sequences of 93-kb and 3.3-kb Plasmids of an Enterohemorrhagic *Escherichia coli* O157:H7 Derived from Sakai Outbreak. *DNA Research*, 5(1), 1-9. https://doi.org/10.1093/dnares/5.1.1

Maloy, S.R. (1990). *Experimental techniques in bacterial genetics*. Jones and Bartlett, Boston.

Marchello, C. S., Dale, A. P., Pisharody, S., Rubach, M. P., & Crump, J. A. (2019). A Systematic Review and Meta-analysis of the Prevalence of Community-Onset Bloodstream Infections among Hospitalized Patients in Africa and Asia. *Antimicrobial Agents and Chemotherapy*, 64(1). https://doi.org/10.1128/aac.01974-19

Mateus, A., Shah, M., Hevler, J., Kurzawa, N., Bobonis, J., Typas, A., Savitski, M. M., & Sanchez, L. M. (2021). Transcriptional and Post-Transcriptional Polar Effects in Bacterial Gene Deletion Libraries. *mSystems*, 6(5), e00813-00821. https://doi.org/doi:10.1128/mSystems.00813-21

Mather, A. E., Phuong, T. L. T., Gao, Y., Clare, S., Mukhopadhyay, S., Goulding, D. A., Hoang, N. T. D., Tuyen, H. T., Lan, N. P. H., Thompson, C. N., Trang, N. H. T., Carrique-Mas, J., Tue, N. T., Campbell, J. I., Rabaa, M. A., Thanh, D. P., Harcourt, K., Hoa, N. T., Trung, N. V., Schultsz, C., Perron, G. G., Coia, J. E., Brown, D. J., Okoro, C., Parkhill, J., Thomson, N. R., Chau, N. V. V., Thwaites, G. E., Maskell, D. J., Dougan, G.,

Kenney, L. J., & Baker, S. (2018). New Variant of Multidrug-Resistant *Salmonella enterica* Serovar Typhimurium Associated with Invasive Disease in Immunocompromised Patients in Vietnam. *mBio*, 9(5), e01056-01018. https://doi.org/10.1128/mBio.01056-18

McClelland, M., Sanderson, K. E., Spieth, J., Clifton, S. W., Latreille, P., Courtney, L., Porwollik, S., Ali, J., Dante, M., Du, F., Hou, S., Layman, D., Leonard, S., Nguyen, C., Scott, K., Holmes, A., Grewal, N., Mulvaney, E., Ryan, E., Sun, H., Florea, L., Miller, W., Stoneking, T., Nhan, M., Waterston, R., & Wilson, R. K. (2001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*, 413(6858), 852-856. https://doi.org/10.1038/35101614

McLaughlin, K. J., Nash, R. P., & Redinbo, M. R. (2014). Unique helicase determinants in the essential conjugative TraI factor from *Salmonella enterica* serovar Typhimurium plasmid pCU1. *Journal of Bacteriology*, 196(17), 3082-3090. https://doi.org/10.1128/JB.01496-14

McQuiston, J., Herrera-Leon, S., Wertheim, B., Doyle, J., Fields, P., Tauxe, R., & Logsdon, J. (2008). Molecular Phylogeny of the *Salmonellae*: Relationships among Salmonella Species and Subspecies Determined from Four Housekeeping Genes and Evidence of Lateral Gene Transfer Events. *Journal of Bacteriology*, 190, 7060-7067. https://doi.org/10.1128/JB.01552-07

Melnyk, A. H., Wong, A., & Kassen, R. (2015). The fitness costs of antibiotic resistance mutations. *Evolutionary Applications*, 8(3), 273-283. https://doi.org/10.1111/eva.12196

Mills, D. M., Bajaj, V., & Lee, C. A. (1995). A 40 kb chromosomal fragment encoding *Salmonella* typhimurium invasion genes is absent from the corresponding region of the *Escherichia coli* K-12 chromosome. *Molecular Microbiology*, 15(4), 749-759. https://doi.org/10.1111/j.1365-2958.1995.tb02382.x

Mogasale, V., Maskery, B., Ochiai, R. L., Lee, J. S., Mogasale, V. V., Ramani, E., Kim, Y. E., Park, J. K., & Wierzba, T. F. (2014). Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment. *The Lancet Global Health*, 2(10), e570-e580. https://doi.org/https://doi.org/10.1016/S2214-109X(14)70301-8

Msefula, C., Kingsley, R., Gordon, M., Molyneux, E., Molyneux, M., Maclennan, C., Dougan, G., & Heyderman, R. (2012). Genotypic Homogeneity of Multidrug Resistant S. Typhimurium Infecting Distinct Adult and Childhood Susceptibility Groups in Blantyre, Malawi. *PloS One*, 7, e42085. https://doi.org/10.1371/journal.pone.0042085

Newberry, K. J., Hou, Y.-M., & Perona, J. J. (2002). Structural origins of amino acid selection without editing by cysteinyl-tRNA synthetase. *The EMBO Journal*, 21(11), 2778-2787. https://doi.org/10.1093/emboj/21.11.2778

Okoro, C. K., Kingsley, R. A., Connor, T. R., Harris, S. R., Parry, C. M., Al-Mashhadani, M. N., Kariuki, S., Msefula, C. L., Gordon, M. A., de Pinna, E., Wain, J., Heyderman, R. S., Obaro, S., Alonso, P. L., Mandomando, I., MacLennan, C. A., Tapia, M. D., Levine, M. M., Tennant, S. M., Parkhill, J., & Dougan, G. (2012). Intracontinental spread of human invasive *Salmonella* Typhimurium pathovariants in sub-Saharan Africa. *Nature Genetics*, 44(11), 1215-1221. https://doi.org/10.1038/ng.2423

Orlek, A., Stoesser, N., Anjum, M. F., Doumith, M., Ellington, M. J., Peto, T., Crook, D., Woodford, N., Walker, A. S., Phan, H., & Sheppard, A. E. (2017). Plasmid Classification in an Era of Whole-Genome Sequencing: Application in Studies of Antibiotic Resistance Epidemiology. *Frontiers in Microbiology*, 8, 182. https://doi.org/10.3389/fmicb.2017.00182

O'Toole, G. A., & Kolter, R. (1998). Flagellar and twitching motility are necessary for Pseudomonas aeruginosa biofilm development. *Molecular Microbiology*, 30(2), 295-304. https://doi.org/https://doi.org/10.1046/j.1365-2958.1998.01062.x

Owen, S. V.(2017). Exploring the prophage biology of *Salmonella enterica* serovar Typhimurium SR313. PhD Thesis. University of Liverpool. https://livrepository.liverpool.ac.uk/3011773/

Owen, S. V., Wenner, N., Canals, R., Makumi, A., Hammarlöf, D. L., Gordon, M. A., Aertsen, A., Feasey, N. A., & Hinton, J. C. D. (2017). Characterization of the Prophage Repertoire of African *Salmonella* Typhimurium ST313 Reveals High Levels of Spontaneous Induction of Novel Phage BTP1. *Frontiers in Microbiology*, 8, 235-235. https://doi.org/10.3389/fmicb.2017.00235

Owen, S. V., Wenner, N., Dulberger, C. L., Rodwell, E. V., Bowers-Barnard, A., Quinones-Olvera, N., Rigden, D. J., Rubin, E. J., Garner, E. C., Baym, M., & Hinton, J. C. D. (2021). Prophages encode phage-defense systems with cognate self-immunity. *Cell Host & Microbe*, 29(11), 1620-1633.e1628. https://doi.org/https://doi.org/10.1016/j.chom.2021.09.002

Pal, C., Bengtsson-Palme, J., Kristiansson, E., & Larsson, D. G. J. (2015). Co-occurrence of resistance genes to antibiotics, biocides and metals reveals novel insights into their co-selection potential. *BMC Genomics*, 16, 964-964. https://doi.org/10.1186/s12864-015-2153-5

Parkhill, J., Dougan, G., James, K. D., Thomson, N. R., Pickard, D., Wain, J., Churcher, C., Mungall, K. L., Bentley, S. D., Holden, M. T., Sebaihia, M., Baker, S., Basham, D., Brooks, K., Chillingworth, T., Connerton, P., Cronin, A., Davis, P., Davies, R. M., Dowd, L., White, N., Farrar, J., Feltwell, T., Hamlin, N., Haque, A., Hien, T. T., Holroyd, S., Jagels, K., Krogh, A., Larsen, T. S., Leather, S., Moule, S., O'Gaora, P., Parry, C., Quail, M., Rutherford, K., Simmonds, M., Skelton, J., Stevens, K., Whitehead, S., & Barrell, B.

G. (2001). Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18. *Nature*, 413(6858), 848-852. https://doi.org/10.1038/35101607

Patel, S., & McCormick, B. A. (2014). Mucosal Inflammatory Response to *Salmonella* typhimurium Infection. *Frontiers in Immunology*, 5, 311. https://doi.org/10.3389/fimmu.2014.00311

Pedersen, T., Tellevik, M., Kommedal, O., Lindemann, P., Moyo, S., Janice, J., Blomberg, B., Samuelsen, Ø., & Langeland, N. (2020). Horizontal Plasmid Transfer among Klebsiella pneumoniae Isolates Is the Key Factor for Dissemination of Extended-Spectrum β-Lactamases among Children in Tanzania. *mSphere*, 5. https://doi.org/10.1128/mSphere.00428-20

Petersen, E., & Miller, S. I. (2019). The cellular microbiology of *Salmonella*e interactions with macrophages. *Cellular Microbiology*, 21(11), e13116. https://doi.org/10.1111/cmi.13116

Perez-Sepulveda, B. M., Heavens, D., Pulford, C. V., Predeus, A. V., Low, R., Webster, H., Dykes, G. F., Schudoma, C., Rowe, W., Lipscombe, J., Watkins, C., Kumwenda, B., Shearer, N., Costigan, K., Baker, K. S., Feasey, N. A., Hinton, J. C. D., Hall, N., Perez-Sepulveda, B. M., Heavens, D., Pulford, C. V., Acuña, M. T., Antic, D., Antonio, M., Baker, K. S., Bernal, J., Bolaños, H., Chattaway, M., Cheesbrough, J., Chirambo, A., Costigan, K., Darboe, S., Díaz, P., Donado, P., Duarte, C., Duarte, F., Everett, D., Fanning, S., Feasey, N. A., Feglo, P., Ferreira, A. M., Floyd, R., Gavilán, R. G., Gordon, M. A., Hall, N., Hernandes, R. T., Hernández-Mora, G., Hinton, J. C. D., Hurley, D., Kasumba, I. N., Kumwenda, B., Kwambana-Adams, B., Lipscombe, J., Low, R., Mattar, S., Montaño, L. A., Moreira, C. G., Moreno, J., Muthappa, D. M., Nair, S., Parry, C. M., Peno, C., Permala-Booth, J., Petrović, J., Predeus, A. V., Puente, J. L., Rebrie, G., Redway, M., Rowe, W., Sadatsune, T., Schudoma, C., Shearer, N., Silva, C., Smith, A. M., Tennant, S., Tran-Dien, A., Watkins, C., Webster, H., Weill, F.-X., Wiesner, M., Wilson, C., & The, K. S. G. c. (2021, 2021/12/21). An accessible, efficient and global approach for the large-scale sequencing of bacterial genomes. *Genome Biology*, 22(1), 349. https://doi.org/10.1186/s13059-021-02536-3

Piccini, G., & Montomoli, E. (2020). Pathogenic signature of invasive non-typhoidal *Salmonella* in Africa: implications for vaccine development. *Human Vaccines & Immunotherapeutics*, 16(9), 2056-2071. https://doi.org/10.1080/21645515.2020.1785791

Pilla, G., & Tang, C. M. (2018). Going around in circles: virulence plasmids in enteric pathogens. *Nature Reviews Microbiology*, 16(8), 484-495. https://doi.org/10.1038/s41579-018-0031-2

Potrykus, K., & Cashel, M. (2008). (p)ppGpp: still magical? *Annual Review of Microbiology*, 62, 35-51. https://doi.org/10.1146/annurev.micro.62.081307.162903

Pradhan, D., & Devi Negi, V. (2019). Stress-induced adaptations in *Salmonella*: A ground for shaping its pathogenesis. *Microbiological Research*, 229, 126311. https://doi.org/https://doi.org/10.1016/j.micres.2019.126311

Preciado-Llanes, L., Aulicino, A., Canals, R., Moynihan, P. J., Zhu, X., Jambo, N., Nyirenda, T. S., Kadwala, I., Gerós, A. S., Owen, S. V., Jambo, K. C., Kumwenda, B., Veerapen, N., Besra, G. S., Gordon, M. A., Hinton, J. C. D., Napolitani, G., Salio, M., & Simmons, A. (2020). Evasion of MAIT cell recognition by the African *Salmonella* Typhimurium ST313 pathovar that causes invasive disease. *Proceedings of the National Academy of Sciences*, 117(34), 20717-20728. https://doi.org/doi:10.1073/pnas.2007472117

Price, M. N., Dehal, P. S., & Arkin, A. P. (2010). FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. *PloS One*, 5(3), e9490. https://doi.org/10.1371/journal.pone.0009490

Ptashne, M. (2004). A genetic switch : phage lambda revisited. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Pulford, C. V., Perez-Sepulveda, B. M., Canals, R., Bevington, J. A., Bengtsson, R. J., Wenner, N., Rodwell, E. V., Kumwenda, B., Zhu, X., Bennett, R. J., Stenhouse, G. E., Malaka De Silva, P., Webster, H. J., Bengoechea, J. A., Dumigan, A., Tran-Dien, A., Prakash, R., Banda, H. C., Alufandika, L., Mautanga, M. P., Bowers-Barnard, A., Beliavskaia, A. Y., Predeus, A. V., Rowe, W. P. M., Darby, A. C., Hall, N., Weill, F.-X., Gordon, M. A., Feasey, N. A., Baker, K. S., & Hinton, J. C. D. (2021). Stepwise evolution of *Salmonella* Typhimurium ST313 causing bloodstream infection in Africa. *Nature Microbiology*, 6(3), 327-338. https://doi.org/10.1038/s41564-020-00836-1

Pulford, C. V., Wenner, N., Redway, M. L., Rodwell, E. V., Webster, H. J., Escudero, R., Kröger, C., Canals, R., Rowe, W., Lopez, J., Hall, N., Rowley, P. D., Timofte, D., Harrison, R. A., Baker, K. S., & Hinton, J. C. D. (2019). The diversity, evolution and ecology of *Salmonella* in venomous snakes. *PLoS Neglected Tropical Diseases*, 13(6), e0007169. https://doi.org/10.1371/journal.pntd.0007169

Rajer, F., Sandegren, L., & Cooper, V. S. (2022). The Role of Antibiotic Resistance Genes in the Fitness Cost of Multiresistance Plasmids. *mBio*, 13(1), e03552-03521. https://doi.org/doi:10.1128/mbio.03552-21

Ramachandran, G., Perkins, D. J., Schmidlein, P. J., Tulapurkar, M. E., & Tennant, S. M. (2015). Invasive *Salmonella* Typhimurium ST313 with Naturally Attenuated Flagellin Elicits Reduced Inflammation and Replicates within Macrophages. *PLoS Neglected Tropical Diseases*, 9(1), e3394. https://doi.org/10.1371/journal.pntd.0003394

Rankin, D. J., Rocha, E. P. C., & Brown, S. P. (2011). What traits are carried on mobile

genetic elements, and why? *Heredity*, 106(1), 1-10. https://doi.org/10.1038/hdy.2010.24

Rankin, J. D., & Taylor, R. J. (1966). The estimation of doses of *Salmonella* typhimurium suitable for the experimental production of disease in calves. *Veterinary Record*, 78(21), 706-707. https://doi.org/10.1136/vr.78.21.706

Renbarger, T. L., Baker, J. M., & Sattley, W. M. (2017). Slow and steady wins the race: an examination of bacterial persistence. *AIMS Microbiology*, 3(2), 171-185. https://doi.org/10.3934/microbiol.2017.2.171

Richardson, D. L., Jr., Aoyama, A., & Hayashi, M. (1988). Proteolysis of bacteriophage phi X174 prohead protein gpB by a protease located in the *Escherichia coli* outer membrane. *Journal of Bacteriology*, 170(12), 5564-5571. https://doi.org/10.1128/jb.170.12.5564-5571.1988

Rodríguez-Beltrán, J., DelaFuente, J., León-Sampedro, R., MacLean, R. C., & San Millán, Á. (2021). Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. *Nature Reviews Microbiology*, 19(6), 347-359. https://doi.org/10.1038/s41579-020-00497-1

Rodwell, E. V., Wenner, N., Pulford, C. V., Cai, Y., Bowers-Barnard, A., Beckett, A., Rigby, J., Picton, D. M., Blower, T. R., Feasey, N. A., Hinton, J. C. D., & Perez-Sepulveda, B. M. (2021). Isolation and Characterisation of Bacteriophages with Activity against Invasive Non-Typhoidal *Salmonella* Causing Bloodstream Infection in Malawi. *Viruses*, 13(3), 478. https://doi.org/10.3390/v13030478

Römling, U., Bian, Z., Hammar, M., Sierralta, W. D., & Normark, S. (1998). Curli fibers are highly conserved between *Salmonella* typhimurium and *Escherichia coli* with respect to operon structure and regulation. *Journal of Bacteriology*, 180(3), 722-731. https://doi.org/10.1128/jb.180.3.722-731.1998

Ruan, B., Nakano, H., Tanaka, M., Mills, J. A., DeVito, J. A., Min, B., Low, K. B., Battista, J. R., & Söll, D. (2004). Cysteinyl-tRNA^{Cys} Formation in *Methanocaldococcus jannaschii*: the Mechanism Is Still Unknown. *Journal of Bacteriology*, 186(1), 8-14. https://doi.org/doi:10.1128/JB.186.1.8-14.2004

Rukambile, E., Sintchenko, V., Muscatello, G., Kock, R., & Alders, R. (2019). Infection, colonization and shedding of Campylobacter and *Salmonella* in animals and their contribution to human disease: A review. *Zoonoses Public Health*, 66(6), 562-578. https://doi.org/10.1111/zph.12611

Sabbagh, S. C., Forest, C. G., Lepage, C., Leclerc, J.-M., & Daigle, F. (2010). So similar, yet so different: uncovering distinctive features in the genomes of *Salmonella* enterica serovars Typhimurium and Typhi. *FEMS Microbiology Letters*, 305(1), 1-13.

https://doi.org/10.1111/j.1574-6968.2010.01904.x

Saha, S., Haggård-Ljungquist, E., & Nordström, K. (1987). The cox protein of bacteriophage P2 inhibits the formation of the repressor protein and autoregulates the early operon. *The EMBO Journal*, 6(10), 3191-3199.

Salmonella Subcommittee of the Nomenclature Committee of the International Society for Microbiology (1934). The Genus *Salmonella* Lignières, 1900. *Journal of Hygiene*, 34(3), 333-350. https://doi.org/10.1017/s0022172400034677

Santiago, F., Doscher, E., Kim, J., Camps, M., Meza, J., Sindi, S., & Barlow, M. (2020). Growth rate assays reveal fitness consequences of β-lactamases. *PloS One*, 15(1), e0228240. https://doi.org/10.1371/journal.pone.0228240

Sauer, R. T., Ross, M. J. & Ptashne M. (1982) Cleavage of the lambda and P22 repressors by recA protein. *Journal of Biological Chemistry*, 257(8), 4458-62. https://doi.org/10.1016/S0021-9258(18)34744-6

Schadich E, Dzubak P & Hajduch M. (2016) Role of Salmonella Typhi Vi Antigen and Secretory Systems on Immune Response. *Current Pharmaceutical Design.*22(41):6251-6260. https://doi: 10.2174/1381612822666160829142308.

Schlechter, R. O., Jun, H., Bernach, M., Oso, S., Boyd, E., Muñoz-Lintz, D. A., Dobson, R. C. J., Remus, D. M., & Remus-Emsermann, M. N. P. (2018). Chromatic Bacteria – A Broad Host-Range Plasmid and Chromosomal Insertion Toolbox for Fluorescent Protein Expression in Bacteria. *Frontiers in Microbiology*, 9, 3052. https://doi.org/10.3389/fmicb.2018.03052

Schmieger, H. (1972). Phage P22-mutants with increased or decreased transduction abilities. *Molecular and General Genetics*, 119(1), 75-88.

Scholz, O., Thiel, A., Hillen, W., & Niederweis, M. (2000). Quantitative analysis of gene expression with an improved green fluorescent protein. *European Journal of Biochemistry*, 267(6), 1565-1570. https://doi.org/10.1046/j.1432-1327.2000.01170.x

Schulte, M., Sterzenbach, T., Miskiewicz, K., Elpers, L., Hensel, M., & Hansmeier, N. (2019). A versatile remote control system for functional expression of bacterial virulence genes based on the tetA promoter. *International Journal of Medical Microbiology*, 309(1), 54-65. https://doi.org/https://doi.org/10.1016/j.ijmm.2018.11.001

Seribelli, A. A., Gonzales, J. C., de Almeida, F., Benevides, L., Cazentini Medeiros, M. I., Dos Prazeres Rodrigues, D., de, C. S. S., Allard, M. W., & Falcão, J. P. (2020). Phylogenetic analysis revealed that Salmonella Typhimurium ST313 isolated from humans and food in Brazil presented a high genomic similarity. *Brazilian Journal of Microbiology*, 51(1),

53-64. https://doi.org/10.1007/s42770-019-00155-6

Shah, D. H., Elder, J. R., Chiok, K. L., & Paul, N. C. (2017). Genetic Basis of *Salmonella* Enteritidis Pathogenesis in Chickens. Producing Safe Eggs, 187-208. https://doi.org/10.1016/B978-0-12-802582-6.00010-0

Shao, Q., Trinh, J. T., & Zeng, L. (2019). High-resolution studies of lysis-lysogeny decision-making in bacteriophage lambda. *The Journal of Biological Chemistry*, 294(10), 3343-3349. https://doi.org/10.1074/jbc.TM118.003209

Shea, J. E., Hensel, M., Gleeson, C., & Holden, D. W. (1996). Identification of a virulence locus encoding a second type III secretion system in Salmonella typhimurium. *Proceedings of the National Academy of Sciences of the United States of America*, 93(6), 2593-2597. https://doi.org/10.1073/pnas.93.6.2593

Shearwin, K. E., Brumby, A. M., & Egan, J. B. (1998). The Tum protein of coliphage 186 is an antirepressor. *Journal of Biological Chemistry*, 273(10), 5708-5715.

Shintani, M. (2017). The behavior of mobile genetic elements (MGEs) in different environments. *Bioscience Biotechnology and Biochemistry*, 81. https://doi.org/10.1080/09168451.2016.1270743

Shintani, M., Sanchez, Z. K., & Kimbara, K. (2015). Genomics of microbial plasmids: classification and identification based on replication and transfer systems and host taxonomy. *Frontiers in Microbiology*, 6, 242-242. https://doi.org/10.3389/fmicb.2015.00242

Sievers, F., & Higgins, D. G. (2014). Clustal Omega, accurate alignment of very large numbers of sequences. *Methods in Molecular Biology*, 1079, 105-116. https://doi.org/10.1007/978-1-62703-646-7_6

Silva, C., Calva, E., Fernández-Mora, M., Puente, J. L., & Vinuesa, P. (2019). Population analysis of D6-like plasmid prophage variants associated with specific IncC plasmid types in the emerging *Salmonella* Typhimurium ST213 genotype. *PloS One*, 14(10), e0223975. https://doi.org/10.1371/journal.pone.0223975

Silva, C., Puente, J. L., & Calva, E. (2017). *Salmonella* virulence plasmid: pathogenesis and ecology. *Pathogens and Disease*, 75(6). https://doi.org/10.1093/femspd/ftx070

Singh, G., Yadav, M., Ghosh, C., & Rathore, J. S. (2021). Bacterial toxin-antitoxin modules: classification, functions, and association with persistence. *Current Research in Microbial Sciences*, 2, 100047. https://doi.org/https://doi.org/10.1016/j.crmicr.2021.100047

Singletary, L. A., Karlinsey, J. E., Libby, S. J., Mooney, J. P., Lokken, K. L., Tsolis, R. M.,

Byndloss, M. X., Hirao, L. A., Gaulke, C. A., Crawford, R. W., Dandekar, S., Kingsley, R. A., Msefula, C. L., Heyderman, R. S., & Fang, F. C. (2016). Loss of Multicellular Behavior in Epidemic African Nontyphoidal *Salmonella enterica* Serovar Typhimurium ST313 Strain D23580. *mBio*, 7(2), e02265. https://doi.org/10.1128/mBio.02265-15

Spiegelhauer, M. R., García, V., Guerra, P. R., Olsen, J. E., & Herrero-Fresno, A. (2020). Association of the prophage BTP1 and the prophage-encoded gene, bstA, with antivirulence of *Salmonella* Typhimurium ST313. *Pathogens and Disease*, 78(3), ftaa019.

Spöring, I., Felgner, S., Preuße, M., Eckweiler, D., Rohde, M., Häussler, S., Weiss, S., Erhardt, M., Gunn, J. S., & Kaufmann, S. H. E. (2018). Regulation of Flagellum Biosynthesis in Response to Cell Envelope Stress in *Salmonella enterica* Serovar Typhimurium. *mBio*, 9(3), e00736-00717. https://doi.org/doi:10.1128/mBio.00736-17

Srivatsan, A., Han, Y., Peng, J., Tehranchi, A. K., Gibbs, R., Wang, J. D., & Chen, R. (2008). High-Precision, Whole-Genome Sequencing of Laboratory Strains Facilitates Genetic Studies. *PLOS Genetics*, 4(8), e1000139. https://doi.org/10.1371/journal.pgen.1000139

Stanaway, J.D., Parisi, A., Sarkar, K., Blacker, B.F., Reiner, R.C., Hay, S.I., Nixon, M.R., Dolecek, C., James, S.L., Mokdad, A.H., Abebe, G., Ahmadian, E., Alahdab, F., Alemnew, B.T.T., Alipour, V., Allah Bakeshei, F., Animut, M.D., Ansari, F., Arabloo, J., Asfaw, E.T., Bagherzadeh, M., Bassat, Q., Belayneh, Y.M.M., Carvalho, F., Daryani, A., Demeke, F.M., Demis, A.B.B., Dubey, M., Duken, E.E., Dunachie, S.J., Eftekhari, A., Fernandes, E., Fouladi Fard, R., Gedefaw, G.A., Geta, B., Gibney, K.B., Hasanzadeh, A., Hoang, C.L., Kasaeian, A., Khater, A., Kidanemariam, Z.T., Lakew, A.M., Malekzadeh, R., Melese, A., Mengistu, D.T., Mestrovic, T., Miazgowski, B., Mohammad, K.A., Mohammadian, M., Mohammadian-Hafshejani, A., Nguyen, C.T., Nguyen, L.H., Nguyen, S.H., Nirayo, Y.L., Olagunju, A.T., Olagunju, T.O., Pourjafar, H., Qorbani, M., Rabiee, M., Rabiee, N., Rafay, A., Rezapour, A., Samy, A.M., Sepanlou, S.G., Shaikh, M.A., Sharif, M., Shigematsu, M., Tessema, B., Tran, B.X., Ullah, I., Yimer, E.M., Zaidi, Z., Murray, C.J.L., Crump, J.A., (2019a). The global burden of non-typhoidal *Salmonella* invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Dis*eases, 19, 1312–1324. https://doi.org/10.1016/S1473-3099(19)30418-9

Stanaway, J.D., Reiner, R.C., Blacker, B.F., Goldberg, E.M., Khalil, I.A., Troeger, C.E., Andrews, J.R., Bhutta, Z.A., Crump, J.A., Im, J., Marks, F., Mintz, E., Park, S.E., Zaidi, A.K.M., Abebe, Z., Abejie, A.N., Adedeji, I.A., Ali, B.A., Amare, A.T., Atalay, H.T., Avokpaho, E.F.G.A., Bacha, U., Barac, A., Bedi, N., Berhane, A., Browne, A.J., Chirinos, J.L., Chitheer, A., Dolecek, C., El Sayed Zaki, M., Eshrati, B., Foreman, K.J., Gemechu, A., Gupta, R., Hailu, G.B., Henok, A., Hibstu, D.T., Hoang, C.L., Ilesanmi, O.S., Iyer, V.J., Kahsay, A., Kasaeian, A., Kassa, T.D., Khan, E.A., Khang, Y.-H., Magdy Abd El Razek, H., Melku, M., Mengistu, D.T., Mohammad, K.A., Mohammed, S., Mokdad, A.H., Nachega, J.B., Naheed, A., Nguyen, C.T., Nguyen, H.L.T., Nguyen, L.H., Nguyen, N.B.,

Nguyen, T.H., Nirayo, Y.L., Pangestu, T., Patton, G.C., Qorbani, M., Rai, R.K., Rana, S.M., Ranabhat, C.L., Roba, K.T., Roberts, N.L.S., Rubino, S., Safiri, S., Sartorius, B., Sawhney, M., Shiferaw, M.S., Smith, D.L., Sykes, B.L., Tran, B.X., Tran, T.T., Ukwaja, K.N., Vu, G.T., Vu, L.G., Weldegebreal, F., Yenit, M.K., Murray, C.J.L., Hay, S.I., (2019b). The global burden of typhoid and 167 paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet Infectious Dis*eases. 19, 369–381. https://doi.org/10.1016/S1473-3099(18)30685-6

Stanley, T. L., Ellermeier, C. D., & Slauch, J. M. (2000). Tissue-specific gene expression identifies a gene in the lysogenic phage Gifsy-1 that affects *Salmonella enterica* serovar typhimurium survival in Peyer's patches. *Journal of Bacteriology*, 182(16), 4406-4413. https://doi.org/10.1128/jb.182.16.4406-4413.2000

Stathopoulos, C., Kim, W., Li, T., Anderson, I., Deutsch, B., Palioura, S., Whitman, W., & Soll, D. (2002). Cysteinyl-tRNA synthetase is not essential for viability of the archaeon Methanococcus maripaludis. *Proceedings of the National Academy of Sciences of the United States of America*, 98, 14292-14297. https://doi.org/10.1073/pnas.201540498

Stevens, M. P., Humphrey, T. J., & Maskell, D. J. (2009). Molecular insights into farm animal and zoonotic *Salmonella* infections. Philosophical transactions of the Royal Society of London. *Philosophical Transactions of the Royal Society B*, 364(1530), 2709-2723. https://doi.org/10.1098/rstb.2009.0094

Sun, D., Jeannot, K., Xiao, Y., & Knapp, C. W. (2019). Editorial: Horizontal Gene Transfer Mediated Bacterial Antibiotic Resistance. *Frontiers in Microbiology*, 10, 1933. https://doi.org/10.3389/fmicb.2019.01933

Susskind, M. M., Wright, A., & Botstein, D. (1974). Superinfection exclusion by P22 prophage in lysogens of *Salmonella* typhimurium: IV. Genetics and physiology of *sieB* exclusion. *Virology*, 62(2), 367-384. https://doi.org/https://doi.org/10.1016/0042-6822(74)90399-7

Tang, K., Wang, W., Sun, Y., Zhou, Y., Wang, P., Guo, Y., & Wang, X. (2021). Prophage Tracer: precisely tracing prophages in prokaryotic genomes using overlapping split-read alignment. *Nucleic Acids Research*, 49(22), e128-e128. https://doi.org/10.1093/nar/gkab824

Tang, S., Orsi, R. H., Luo, H., Ge, C., Zhang, G., Baker, R. C., Stevenson, A., & Wiedmann, M. (2019). Assessment and Comparison of Molecular Subtyping and Characterization Methods for *Salmonella*. *Frontiers in Microbiology*, 10, 1591. https://doi.org/10.3389/fmicb.2019.01591

Taylor, K., & Wegrzyn, G. (1995). Replication of coliphage lambda DNA. *FEMS Microbiology Reviews*, 17(1-2), 109-119. https://doi.org/10.1111/j.1574-6976.1995.tb00192.x

Thompson, A., Rolfe, M. D., Lucchini, S., Schwerk, P., Hinton, J. C. D., & Tedin, K. (2006). The Bacterial Signal Molecule, ppGpp, Mediates the Environmental Regulation of Both the Invasion and Intracellular Virulence Gene Programs of *Salmonella**. *Journal of Biological Chemistry*, 281, 30112 - 30121. https://doi.org/10.1074/jbc.M605616200

Tomljenovic-Berube, A. M., Henriksbo, B., Porwollik, S., Cooper, C. A., Tuinema, B. R., McClelland, M., & Coombes, B. K. (2013). Mapping and regulation of genes within *Salmonella* pathogenicity island 12 that contribute to in vivo fitness of *Salmonella enterica* Serovar Typhimurium. *Infection and Immunity*, 81(7), 2394-2404. https://doi.org/10.1128/IAI.00067-13

Tsai, C. N., & Coombes, B. K. (2021). Emergence of invasive *Salmonella* in Africa. *Nature Microbiology*, 6(3), 273-274. https://doi.org/10.1038/s41564-021-00864-5

Tsolis, R. M., Xavier, M. N., Santos, R. L., & Bäumler, A. J. (2011). How to become a top model: impact of animal experimentation on human *Salmonella* disease research. *Infection and Immunity*, 79(5), 1806-1814. https://doi.org/10.1128/iai.01369-10

Tsolis, R. M., Young, G. M., Solnick, J. V., & Bäumler, A. J. (2008). From bench to bedside: stealth of enteroinvasive pathogens. *Nature Reviews Microbiology*, 6(12), 883-892. https://doi.org/10.1038/nrmicro2012

Uche, I. V., MacLennan, C. A., & Saul, A. (2017). A Systematic Review of the Incidence, Risk Factors and Case Fatality Rates of Invasive Nontyphoidal *Salmonella* (iNTS) Disease in Africa (1966 to 2014). *PLoS Neglected Tropical Diseases*, 11(1), e0005118-e0005118. https://doi.org/10.1371/journal.pntd.0005118

Unterholzner, S. J., Poppenberger, B., & Rozhon, W. (2013). Toxin–antitoxin systems: biology, identification, and application. *Mobile Genetic Elements*, 3(5), e26219.

Uzzau, S., Brown, D. J., Wallis, T., Rubino, S., Leori, G., Bernard, S., Casadesús, J., Platt, D. J., & Olsen, J. E. (2000). Host adapted serotypes of *Salmonella* enterica. *Epidemiology and Infection*, 125(2), 229-255. https://doi.org/10.1017/s0950268899004379

Van Puyvelde, S., Pickard, D., Vandelannoote, K., Heinz, E., Barbé, B., de Block, T., Clare, S., Coomber, E. L., Harcourt, K., Sridhar, S., Lees, E. A., Wheeler, N. E., Klemm, E. J., Kuijpers, L., Mbuyi Kalonji, L., Phoba, M.-F., Falay, D., Ngbonda, D., Lunguya, O., Jacobs, J., Dougan, G., & Deborggraeve, S. (2019). An African *Salmonella* Typhimurium ST313 sublineage with extensive drug-resistance and signatures of host adaptation. *Nature Communications*, 10(1), 4280. https://doi.org/10.1038/s41467-019-11844-z

Vershon, A. K., Liao, S. M., McClure, W. R., & Sauer, R. T. (1987). Bacteriophage P22 Mnt repressor. DNA binding and effects on transcription in vitro. *Journal of Molecular Biology*, 195(2), 311-322. https://doi.org/10.1016/0022-2836(87)90652-8

Wahl, A., Battesti, A., & Ansaldi, M. (2019). Prophages in *Salmonella* enterica: a driving force in reshaping the genome and physiology of their bacterial host? *Molecular Microbiology*, 111(2), 303-316. https://doi.org/10.1111/mmi.14167

Waterman, S. R., & Holden, D. W. (2003). Functions and effectors of the *Salmonella* pathogenicity island 2 type III secretion system. *Cellular Microbiology*, 5(8), 501-511. https://doi.org/10.1046/j.1462-5822.2003.00294.x

Weening, E. H., Barker, J. D., Laarakker, M. C., Humphries, A. D., Tsolis, R. M., & Bäumler, A. J. (2005). The *Salmonella* enterica serotype Typhimurium lpf, bcf, stb, stc, std, and sth fimbrial operons are required for intestinal persistence in mice. *Infection and Immunity*, 73(6), 3358-3366. https://doi.org/10.1128/IAI.73.6.3358-3366.2005

White, A. P., Gibson, D. L., Kim, W., Kay, W. W., & Surette, M. G. (2006). Thin Aggregative Fimbriae and Cellulose Enhance Long-Term Survival and Persistence of *Salmonella*. *Journal of Bacteriology*, 188(9), 3219-3227. https://doi.org/doi:10.1128/JB.188.9.3219-3227.2006

Wiser, M. J., & Lenski, R. E. (2015). A Comparison of Methods to Measure Fitness in *Escherichia coli*. *PloS One*, 10(5), e0126210. https://doi.org/10.1371/journal.pone.0126210

Yamaguchi, T., Toma, S., Terahara, N., Miyata, T., Ashihara, M., Minamino, T., Namba, K., & Kato, T. (2020). Structural and functional comparison of *Salmonella* flagellar filaments composed of FljB and FliC. *Biomolecules*, 10(2), 246.

Yang, H. L., & Kessler, D. P. (1974). Genetic analysis of the leucine region in *Escherichia coli* B-r: gene-enzyme assignments. *Journal of Bacteriology*, 117(1), 63-72. https://doi.org/10.1128/jb.117.1.63-72.1974

Yang, J., Barrila, J., Roland, K. L., Kilbourne, J., Ott, C. M., Forsyth, R. J., & Nickerson, C. A. (2015). Characterization of the Invasive, Multidrug Resistant Non-typhoidal *Salmonella* Strain D23580 in a Murine Model of Infection. *PLoS Neglected Tropical Diseases*, 9(6), e0003839. https://doi.org/10.1371/journal.pntd.0003839

Yang, X., Huang, J., Wu, Q., Zhang, J., Yang, S., Wang, J., Ding, Y., Chen, M., Xue, L., Wu, S., Gu, Q., Zhang, Y., & Wei, X. (2022). Occurrence, serovars and antibiotic resistance of *Salmonella* spp. in retail ready-to-eat food products in some Chinese provinces. *LWT*, 154, 112699. https://doi.org/https://doi.org/10.1016/j.lwt.2021.112699

Yu, X. J., Grabe, G. J., Liu, M., Mota, L. J., & Holden, D. W. (2018). SsaV Interacts with SsaL to Control the Translocon-to-Effector Switch in the *Salmonella* SPI-2 Type Three Secretion System. *mBio*, 9(5). https://doi.org/10.1128/mBio.01149-18

Yu, Y., Ouyang, Y., & Yao, W. (2017). shinyCircos: an R/Shiny application for interactive creation of Circos plot. *Bioinformatics*, 34(7), 1229-1231. https://doi.org/10.1093/bioinformatics/btx763

Zhang, T., Zhu, J., Wei, S., Luo, Q., Li, L., Li, S., Tucker, A., Shao, H., & Zhou, R. (2016). The roles of RelA/(p)ppGpp in glucose-starvation induced adaptive response in the zoonotic Streptococcus suis. *Scientific Reports*, 6(1), 27169. https://doi.org/10.1038/srep27169

Zhou, Y., Liang, Y., Lynch, K. H., Dennis, J. J., & Wishart, D. S. (2011). PHAST: a fast phage search tool. *Nucleic Acids Research*, 39, W347-352. https://doi.org/10.1093/nar/gkr485

Zinke, M., Schröder, G. F., & Lange, A. (2022). Major tail proteins of bacteriophages of the order Caudovirales. *Journal of Biology Chemistry*, 298(1), 101472. https://doi.org/10.1016/j.jbc.2021.101472

Appendix I

All genomic data used in this thesis are listed below:

strains	Sequence type	Accession Number/Barcode	Source	Country	Reference
Typhi_CT18	NA	AL513382.1	GenBank	NA	Parkhill et al. (2001)
4/74	ST19	CP002487.1	GenBank	UK	Richardson et al. (2011)
D23580	ST313	GCA_000027025.1	GenBank	Malawi	Kingsley et al. (2009)
STm_44	ST313	LVFU00000000	GenBank	Brazil	Almeida et al. (2017)
STm_47	ST313	LUIX00000000	GenBank	Brazil	Almeida et al. (2017)
STm_39	ST313	LUJB00000000	GenBank	Brazil	Almeida et al. (2017)
STm_34	ST313	LVGA00000000	GenBank	Brazil	Almeida et al. (2017)
STm_35	ST313	LVFZ00000000	GenBank	Brazil	Almeida et al. (2017)
STm_30	ST313	LVGD00000000	GenBank	Brazil	Almeida et al. (2017)
U82	ST313	SRR5451272	SRA	UK	Ashton et al. (2017)
U31	ST313	SRR1967938	SRA	UK	Ashton et al. (2017)
U52	ST313	SRR3284779	SRA	UK	Ashton et al. (2017)
U15	ST313	SRR1645388	SRA	UK	Ashton et al. (2017)
U11	ST313	SRR1965387	SRA	UK	Ashton et al. (2017)
U56	ST313	SRR5451268	SRA	UK	Ashton et al. (2017)
U48	ST313	SRR5451252	SRA	UK	Ashton et al. (2017)
U75	ST313	SRR5451257	SRA	UK	Ashton et al. (2017)
U74	ST313	SRR5451280	SRA	UK	Ashton et al. (2017)
U4	ST313	SRR1645545	SRA	UK	Ashton et al. (2017)
U10	ST313	SRR1965263	SRA	UK	Ashton et al. (2017)
U9	ST313	SRR1968363	SRA	UK	Ashton et al. (2017)
U7	ST313	SRR1966408	SRA	UK	Ashton et al. (2017)
U33	ST313	SRR1968269	SRA	UK	Ashton et al. (2017)
U32	ST313	SRR1969492	SRA	UK	Ashton et al. (2017)
U25	ST313	SRR1958036	SRA	UK	Ashton et al. (2017)
U51	ST313	SRR3323011	SRA	UK	Ashton et al. (2017)
U84	ST313	SRR5451256	SRA	UK	Ashton et al. (2017)
U3	ST313	SRR1646238	SRA	UK	Ashton et al. (2017)
U80	ST313	SRR5451253	SRA	UK	Ashton et al. (2017)
U8	ST313	SRR1966656	SRA	UK	Ashton et al. (2017)
U44	ST313	SRR5451287	SRA	UK	Ashton et al. (2017)
U34	ST313	SRR1966635	SRA	UK	Ashton et al. (2017)
U36	ST313	SRR1966593	SRA	UK	Ashton et al. (2017)
U38	ST313	SRR1967438	SRA	UK	Ashton et al. (2017)
U37	ST313	SRR1968962	SRA	UK	Ashton et al. (2017)
U12	ST313	SRR1967419	SRA	UK	Ashton et al. (2017)
U16	ST313	SRR1645171	SRA	UK	Ashton et al. (2017)
U46	ST313	SRR5451286	SRA	UK	Ashton et al. (2017)

U23	ST313	SRR1969854	SRA	UK	Ashton et al. (2017)
U68	ST313	SRR5451264	SRA	UK	Ashton et al. (2017)
U86	ST313	SRR5451274	SRA	UK	Ashton et al. (2017)
U70	ST313	SRR5451259	SRA	UK	Ashton et al. (2017)
U50	ST313	SRR5451285	SRA	UK	Ashton et al. (2017)
U66	ST313	SRR3322379	SRA	UK	Ashton et al. (2017)
U24	ST313	SRR1967447	SRA	UK	Ashton et al. (2017)
U40	ST313	SRR1968061	SRA	UK	Ashton et al. (2017)
U54	ST313	SRR5193685	SRA	UK	Ashton et al. (2017)
U22	ST313	SRR3048892	SRA	UK	Ashton et al. (2017)
U57	ST313	SRR5451279	SRA	UK	Ashton et al. (2017)
U83	ST313	SRR5451267	SRA	UK	Ashton et al. (2017)
U27	ST313	SRR1958303	SRA	UK	Ashton et al. (2017)
U58	ST313	SRR5451270	SRA	UK	Ashton et al. (2017)
U81	ST313	SRR5215816	SRA	UK	Ashton et al. (2017)
U53	ST313	SRR3285477	SRA	UK	Ashton et al. (2017)
U62	ST313	SRR5451283	SRA	UK	Ashton et al. (2017)
U2	ST313	SRR1646230	SRA	UK	Ashton et al. (2017)
U61	ST313	SRR5451275	SRA	UK	Ashton et al. (2017)
U14	ST313	SRR1966418	SRA	UK	Ashton et al. (2017)
U49	ST313	SRR5451278	SRA	UK	Ashton et al. (2017)
U29	ST313	SRR1968084	SRA	UK	Ashton et al. (2017)
U77	ST313	SRR5451260	SRA	UK	Ashton et al. (2017)
U59	ST313	SRR5451254	SRA	UK	Ashton et al. (2017)
U55	ST313	SRR3322345	SRA	UK	Ashton et al. (2017)
U72	ST313	SRR5451284	SRA	UK	Ashton et al. (2017)
U60	ST313	SRR5451282	SRA	UK	Ashton et al. (2017)
U1	ST313	SRR1646227	SRA	UK	Ashton et al. (2017)
U13	ST313	SRR1966509	SRA	UK	Ashton et al. (2017)
U35	ST313	SRR1966023	SRA	UK	Ashton et al. (2017)
U42	ST313	SRR3285401	SRA	UK	Ashton et al. (2017)
U41	ST313	SRR3286620	SRA	UK	Ashton et al. (2017)
U85	ST313	SRR5451269	SRA	UK	Ashton et al. (2017)
U67	ST313	SRR3323046	SRA	UK	Ashton et al. (2017)
U45	ST313	SRR5451277	SRA	UK	Ashton et al. (2017)
U18	ST313	SRR1966068	SRA	UK	Ashton et al. (2017)
U6	ST313	SRR1967248	SRA	UK	Ashton et al. (2017)
U19	ST313	SRR1969775	SRA	UK	Ashton et al. (2017)
U20	ST313	SRR3048536	SRA	UK	Ashton et al. (2017)
U17	ST313	SRR1967883	SRA	UK	Ashton et al. (2017)
U73	ST313	SRR5451273	SRA	UK	Ashton et al. (2017)
U71	ST313	SRR5451271	SRA	UK	Ashton et al. (2017)
A130	ST313	ERA000075	SRA	UK	Kingsley et al. (2009)

D37712	ST313	ERR023634	EBI	Malawi	Msefula et al. (2012)
D36233	ST313	ERR023642	EBI	Malawi	Msefula et al. (2012)
D37601	ST313	ERR023636	EBI	Malawi	Msefula et al. (2012)
D36435	ST313	ERR023638	EBI	Malawi	Msefula et al. (2012)
A50315	ST313	ERR023657	EBI	Malawi	Msefula et al. (2012)
A50070	ST313	ERR023656	EBI	Malawi	Msefula et al. (2012)
D36448	ST313	ERR023643	EBI	Malawi	Msefula et al. (2012)
D15132	ST313	ERR023620	EBI	Malawi	Msefula et al. (2012)
D15330	ST313	ERR023630	EBI	Malawi	Msefula et al. (2012)
C5158	ST313	ERR023626	EBI	Malawi	Msefula et al. (2012)
D36807	ST313	ERR023640	EBI	Malawi	Msefula et al. (2012)
A22804	ST313	ERR023654	EBI	Malawi	Msefula et al. (2012)
D15176	ST313	ERR023624	EBI	Malawi	Msefula et al. (2012)
C5371	ST313	ERR023627	EBI	Malawi	Msefula et al. (2012)
D15759	ST313	ERR023625	EBI	Malawi	Msefula et al. (2012)
D36457	ST313	ERR023639	EBI	Malawi	Msefula et al. (2012)
D19828	ST313	ERR023622	EBI	Malawi	Msefula et al. (2012)
D16287	ST313	ERR023631	EBI	Malawi	Msefula et al. (2012)
A16802	ST313	ERR023650	EBI	Malawi	Msefula et al. (2012)
A19741	ST313	ERR023653	EBI	Malawi	Msefula et al. (2012)
D36632	ST313	ERR023644	EBI	Malawi	Msefula et al. (2012)
269DRC	ST313	ERS009048	EBI	DRC	Okoro et al. (2012)
356DRC	ST313	ERS004905	EBI	DRC	Okoro et al. (2012)
270DRC	ST313	ERS009046	EBI	DRC	Okoro et al. (2012)
265DRC	ST313	ERS009034	EBI	DRC	Okoro et al. (2012)
334U	ST313	ERS009044	EBI	Uganda	Okoro et al. (2012)
958U	ST313	ERS009035	EBI	Uganda	Okoro et al. (2012)
5577	ST313	ERS004958	EBI	Kenya	Okoro et al. (2012)
5897U	ST313	ERS009037	EBI	Uganda	Okoro et al. (2012)
1905U	ST313	ERS009036	EBI	Uganda	Okoro et al. (2012)
A357	ST313	ERS007448	EBI	Malawi	Okoro et al. (2012)
A018	ST313	ERS007446	EBI	Malawi	Okoro et al. (2012)
A13198	ST313	ERS007457	EBI	Malawi	Okoro et al. (2012)
M1253289	ST313	ERS009022	EBI	Mozambique	Okoro et al. (2012)
M1175849	ST313	ERS009051	EBI	Mozambique	Okoro et al. (2012)
M1776464	ST313	ERS007624	EBI	Mozambique	Okoro et al. (2012)
M1174516	ST313	ERS007629	EBI	Mozambique	Okoro et al. (2012)
D22889	ST313	ERS005403	EBI	Malawi	Okoro et al. (2012)
5580	ST313	ERS004896	EBI	Kenya	Okoro et al. (2012)
5576	ST313	ERS004955	EBI	Kenya	Okoro et al. (2012)
5597B	ST313	ERS004898	EBI	Kenya	Okoro et al. (2012)
5579	ST313	ERS004962	EBI	Kenya	Okoro et al. (2012)
A4447	ST313	ERS007453	EBI	Malawi	Okoro et al. (2012)

A3800	ST313	ERS007451	EBI	Malawi	Okoro et al. (2012)
A4283	ST313	ERS007452	EBI	Malawi	Okoro et al. (2012)
C2167	ST313	ERS007455	EBI	Malawi	Okoro et al. (2012)
A082	ST313	ERS007447	EBI	Malawi	Okoro et al. (2012)
Q175F6	ST313	ERS033126	EBI	Malawi	Okoro et al. (2012)
D25646	ST313	ERS004895	EBI	Malawi	Okoro et al. (2012)
C2364	ST313	ERS007456	EBI	Malawi	Okoro et al. (2012)
D25248	ST313	ERS004899	EBI	Malawi	Okoro et al. (2012)
A680	ST313	ERS007450	EBI	Malawi	Okoro et al. (2012)
C2110	ST313	ERS007454	EBI	Malawi	Okoro et al. (2012)
DT2B	ST313	ERS007580	EBI	UK	Okoro et al. (2012)
254DRC	ST313	ERS009040	EBI	DRC	Okoro et al. (2012)
14DRC	ST313	ERS009029	EBI	DRC	Okoro et al. (2012)
415DRC	ST313	ERS009039	EBI	DRC	Okoro et al. (2012)
373DRC	ST313	ERS009047	EBI	DRC	Okoro et al. (2012)
146U	ST313	ERS009045	EBI	Uganda	Okoro et al. (2012)
198U	ST313	ERS009032	EBI	Uganda	Okoro et al. (2012)
666U	ST313	ERS004903	EBI	Uganda	Okoro et al. (2012)
A39129	ST313	ERS007473	EBI	Malawi	Okoro et al. (2012)
A39155	ST313	ERS007474	EBI	Malawi	Okoro et al. (2012)
D14916	ST313	ERS007461	EBI	Malawi	Okoro et al. (2012)
Q340A	ST313	ERS033119	EBI	Malawi	Okoro et al. (2012)
D25023	ST313	ERS004897	EBI	Malawi	Okoro et al. (2012)
Q255A	ST313	ERS033115	EBI	Malawi	Okoro et al. (2012)
Q363A	ST313	ERS033120	EBI	Malawi	Okoro et al. (2012)
Q134F9	ST313	ERS033123	EBI	Malawi	Okoro et al. (2012)
Q367F2	ST313	ERS033133	EBI	Malawi	Okoro et al. (2012)
Q367A	ST313	ERS033121	EBI	Malawi	Okoro et al. (2012)
Q255F4	ST313	ERS033127	EBI	Malawi	Okoro et al. (2012)
Q258A	ST313	ERS033116	EBI	Malawi	Okoro et al. (2012)
Q258F4	ST313	ERS033128	EBI	Malawi	Okoro et al. (2012)
Q363F3	ST313	ERS033132	EBI	Malawi	Okoro et al. (2012)
Q285A	ST313	ERS033117	EBI	Malawi	Okoro et al. (2012)
Q285F5	ST313	ERS033129	EBI	Malawi	Okoro et al. (2012)
Q303F5	ST313	ERS033130	EBI	Malawi	Okoro et al. (2012)
D25734	ST313	ERS004900	EBI	Malawi	Okoro et al. (2012)
Q134A	ST313	ERS033111	EBI	Malawi	Okoro et al. (2012)
J20	ST313	ERS007482	EBI	Mali	Okoro et al. (2012)
J17	ST313	ERS007480	EBI	Mali	Okoro et al. (2012)
P62	ST313	ERS007476	EBI	Mali	Okoro et al. (2012)
P78	ST313	ERS007475	EBI	Mali	Okoro et al. (2012)
P73	ST313	ERS007477	EBI	Mali	Okoro et al. (2012)
M2815310	ST313	ERS009050	EBI	Mozambique	Okoro et al. (2012)

M1605206	ST313	ERS007627	EBI	Mozambique	Okoro et al. (2012)
M2907772	ST313	ERS007619	EBI	Mozambique	Okoro et al. (2012)
PO609	ST313	ERS009024	EBI	Nigeria	Okoro et al. (2012)
PO1276	ST313	ERS009027	EBI	Nigeria	Okoro et al. (2012)
PO1124	ST313	ERS009025	EBI	Nigeria	Okoro et al. (2012)
PO1140	ST313	ERS009026	EBI	Nigeria	Okoro et al. (2012)
J27	ST313	ERS007481	EBI	Mali	Okoro et al. (2012)
I32	ST313	ERS007484	EBI	Mali	Okoro et al. (2012)
J3	ST313	ERS007479	EBI	Mali	Okoro et al. (2012)
P51	ST313	ERS007478	EBI	Mali	Okoro et al. (2012)
I7	ST313	ERS007486	EBI	Mali	Okoro et al. (2012)
I45	ST313	ERS007485	EBI	Mali	Okoro et al. (2012)
I3	ST313	ERS007483	EBI	Mali	Okoro et al. (2012)
Q18A	ST313	ERS033110	EBI	Malawi	Okoro et al. (2012)
Q175A	ST313	ERS033114	EBI	Malawi	Okoro et al. (2012)
Q18F3_S	ST313	ERS033124	EBI	Malawi	Okoro et al. (2012)
A24910	ST313	ERS007464	EBI	Malawi	Okoro et al. (2012)
D23769A	ST313	ERS007467	EBI	Malawi	Okoro et al. (2012)
A32751	ST313	ERS007465	EBI	Malawi	Okoro et al. (2012)
A38589	ST313	ERS007469	EBI	Malawi	Okoro et al. (2012)
A32773	ST313	ERS007466	EBI	Malawi	Okoro et al. (2012)
A39051	ST313	ERS007472	EBI	Malawi	Okoro et al. (2012)
A38596	ST313	ERS007470	EBI	Malawi	Okoro et al. (2012)
C13184	ST313	ERS007471	EBI	Malawi	Okoro et al. (2012)
2643	ST313	ERS1698169	SRA	DRC	Van Puyvelde et al. (2019)
5232_4	ST313	ERS1698258	SRA	DRC	Van Puyvelde et al. (2019)
6948_11	ST313	ERS3501163	SRA	DRC	Van Puyvelde et al. (2019)
16755_3	ST313	ERS1698201	SRA	DRC	Van Puyvelde et al. (2019)
16927_3	ST313	ERS1698202	SRA	DRC	Van Puyvelde et al. (2019)
11711_3	ST313	ERS3501155	SRA	DRC	Van Puyvelde et al. (2019)
9541_3	ST313	ERS3501175	SRA	DRC	Van Puyvelde et al. (2019)
5390_4	ST313	ERS1698262	SRA	DRC	Van Puyvelde et al. (2019)
4225_3	ST313	ERS1698248	SRA	DRC	Van Puyvelde et al. (2019)
97	ST313	ERS1698078	SRA	DRC	Van Puyvelde et al. (2019)
12146_3	ST313	ERS3501158	SRA	DRC	Van Puyvelde et al. (2019)
10102_3	ST313	ERS3501148	SRA	DRC	Van Puyvelde et al. (2019)
7017_3	ST313	ERS1698274	SRA	DRC	Van Puyvelde et al. (2019)
12299_3	ST313	ERS3501161	SRA	DRC	Van Puyvelde et al. (2019)
8944_3	ST313	ERS3501168	SRA	DRC	Van Puyvelde et al. (2019)
9936_3	ST313	ERS3501181	SRA	DRC	Van Puyvelde et al. (2019)
9412_3	ST313	ERS3501173	SRA	DRC	Van Puyvelde et al. (2019)
9	ST313	ERS1698075	SRA	DRC	Van Puyvelde et al. (2019)
3198_3	ST313	ERS1698221	SRA	DRC	Van Puyvelde et al. (2019)

706	ST313	ERS1698149	SRA	DRC	Van Puyvelde et al. (2019)
3974_3	ST313	ERS1698242	SRA	DRC	Van Puyvelde et al. (2019)
5708_3	ST313	ERS1698265	SRA	DRC	Van Puyvelde et al. (2019)
2101	ST313	ERS3501145	SRA	DRC	Van Puyvelde et al. (2019)
9266_3	ST313	ERS3501172	SRA	DRC	Van Puyvelde et al. (2019)
6634_3	ST313	ERS1698272	SRA	DRC	Van Puyvelde et al. (2019)
1577	ST313	ERS1698160	SRA	DRC	Van Puyvelde et al. (2019)
1582	ST313	ERS1698161	SRA	DRC	Van Puyvelde et al. (2019)
2735	ST313	ERS1698170	SRA	DRC	Van Puyvelde et al. (2019)
8820_13	ST313	ERS3501165	SRA	DRC	Van Puyvelde et al. (2019)
9060_3	ST313	ERS3501170	SRA	DRC	Van Puyvelde et al. (2019)
10055_3	ST313	ERS3501146	SRA	DRC	Van Puyvelde et al. (2019)
9849_3	ST313	ERS3501180	SRA	DRC	Van Puyvelde et al. (2019)
4174_4	ST313	ERS1698246	SRA	DRC	Van Puyvelde et al. (2019)
18221_3	ST313	ERS1698227	SRA	DRC	Van Puyvelde et al. (2019)
17072_3	ST313	ERS1698206	SRA	DRC	Van Puyvelde et al. (2019)
15552_3	ST313	ERS1698195	SRA	DRC	Van Puyvelde et al. (2019)
10100_3	ST313	ERS3501147	SRA	DRC	Van Puyvelde et al. (2019)
8616_3	ST313	ERS1698310	SRA	DRC	Van Puyvelde et al. (2019)
11892_3	ST313	ERS1698183	SRA	DRC	Van Puyvelde et al. (2019)
11658_3	ST313	ERS3501144	SRA	DRC	Van Puyvelde et al. (2019)
11949_3	ST313	ERS3501156	SRA	DRC	Van Puyvelde et al. (2019)
11988_3	ST313	ERS3501157	SRA	DRC	Van Puyvelde et al. (2019)
9085_3	ST313	ERS1698341	SRA	DRC	Van Puyvelde et al. (2019)
10308_3	ST313	ERS3501152	SRA	DRC	Van Puyvelde et al. (2019)
9114_3	ST313	ERS3501171	SRA	DRC	Van Puyvelde et al. (2019)
10142_3	ST313	ERS3501149	SRA	DRC	Van Puyvelde et al. (2019)
11979_3	ST313	ERS1698184	SRA	DRC	Van Puyvelde et al. (2019)
11522_3	ST313	ERS3501143	SRA	DRC	Van Puyvelde et al. (2019)
10293_3	ST313	ERS3501151	SRA	DRC	Van Puyvelde et al. (2019)
9780_3	ST313	ERS3501179	SRA	DRC	Van Puyvelde et al. (2019)
9963_3	ST313	ERS3501182	SRA	DRC	Van Puyvelde et al. (2019)
8874_3	ST313	ERS3501167	SRA	DRC	Van Puyvelde et al. (2019)
10958_3	ST313	ERS3501137	SRA	DRC	Van Puyvelde et al. (2019)
11185_3	ST313	ERS3501138	SRA	DRC	Van Puyvelde et al. (2019)
18228_3	ST313	ERS1698228	SRA	DRC	Van Puyvelde et al. (2019)
12306_3	ST313	ERS3501162	SRA	DRC	Van Puyvelde et al. (2019)
17399_3	ST313	ERS1698207	SRA	DRC	Van Puyvelde et al. (2019)
17568_3	ST313	ERS1698210	SRA	DRC	Van Puyvelde et al. (2019)
8106_3	ST313	ERS1698295	SRA	DRC	Van Puyvelde et al. (2019)
7905_3	ST313	ERS1698291	SRA	DRC	Van Puyvelde et al. (2019)
7772_3	ST313	ERS1698286	SRA	DRC	Van Puyvelde et al. (2019)
12228_3	ST313	ERS3501160	SRA	DRC	Van Puyvelde et al. (2019)

10952_3	ST313	ERS3501136	SRA	DRC	Van Puyvelde et al. (2019)
11279_3	ST313	ERS3501141	SRA	DRC	Van Puyvelde et al. (2019)
11428_3	ST313	ERS3501142	SRA	DRC	Van Puyvelde et al. (2019)
12155_3	ST313	ERS3501159	SRA	DRC	Van Puyvelde et al. (2019)
9673_3	ST313	ERS3501176	SRA	DRC	Van Puyvelde et al. (2019)
11198_3	ST313	ERS3501139	SRA	DRC	Van Puyvelde et al. (2019)
10433_3	ST313	ERS3501154	SRA	DRC	Van Puyvelde et al. (2019)
10155_3	ST313	ERS3501150	SRA	DRC	Van Puyvelde et al. (2019)
9460_3	ST313	ERS3501174	SRA	DRC	Van Puyvelde et al. (2019)
9745_3	ST313	ERS3501178	SRA	DRC	Van Puyvelde et al. (2019)
10393_3	ST313	ERS3501153	SRA	DRC	Van Puyvelde et al. (2019)
8977_3	ST313	ERS3501169	SRA	DRC	Van Puyvelde et al. (2019)
8866_3	ST313	ERS3501166	SRA	DRC	Van Puyvelde et al. (2019)
9713_3	ST313	ERS3501177	SRA	DRC	Van Puyvelde et al. (2019)
8351_3_R	ST313	ERS3501164	SRA	DRC	Van Puyvelde et al. (2019)
10799_3	ST313	ERS3501135	SRA	DRC	Van Puyvelde et al. (2019)
10587_3	ST313	ERS3501134	SRA	DRC	Van Puyvelde et al. (2019)
11214_3	ST313	ERS3501140	SRA	DRC	Van Puyvelde et al. (2019)
FR10242002	ST313	SRR12152304	SRA	DRC	Pulford et al.(2021)
FR10242017	ST313	SRR12152348	SRA	DRC	Pulford et al.(2021)
FR10241936	ST313	SRR12152311	SRA	DRC	Pulford et al.(2021)
FR10241947	ST313	SRR12152334	SRA	DRC	Pulford et al.(2021)
FR10242053	ST313	SRR12152344	SRA	DRC	Pulford et al.(2021)
FR10242030	ST313	SRR12152316	SRA	DRC	Pulford et al.(2021)
FR10241978	ST313	SRR12152309	SRA	DRC	Pulford et al.(2021)
FR10241965	ST313	SRR12152329	SRA	DRC	Pulford et al.(2021)
FR10242004	ST313	SRR12152319	SRA	DRC	Pulford et al.(2021)
RM9437	ST19	CP012985.1	GenBank	USA	NA
138736	ST19	CP007581.1	GenBank	Israel	NA
DT104	ST19	HF937208.1	GenBank	UK	NA
ATCC_13311	ST19	CP009102.1	GenBank	UK	NA
U288	ST19	NC_021151.1	GenBank	UK	NA
T000240	ST19	AP011957.1	GenBank	Japan	NA
DT2	ST19	NC_022544.1	GenBank	UK	NA
14028s	ST19	NC_016856.1	GenBank	USA	NA
ST931R	ST19	CP016385.1	GenBank	USA	NA
UK1	ST19	NC_016863.1	GenBank	USA	NA
798	ST19	CP003386.1	GenBank	UK	NA
SL1344	ST19	FQ312003.1	GenBank	USA	NA
LT2	ST19	AE006468.2	GenBank	USA	NA

Appendix II

ORF function prediction for BTP1 and BTP5-like UK prophages. The annotation was performed by RAST server in 2019 (Aziz *et al.*, 2008).

BTP1^{UK1}

Feature ID	Start	Stop	Length (bp)	Function
BTP1UK1-1	13	1251	1239	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
BTP1UK1-2	2620	1457	1164	Mobile element protein
BTP1UK1-3	3479	2850	630	Phage EaC protein
BTP1UK1-4	3759	3580	180	Phage EaG protein
BTP1UK1-5	4385	3852	534	Phage EaA protein
BTP1UK1-6	5053	4382	672	Phage EaA protein
BTP1UK1-7	5194	5057	138	Phage protein
BTP1UK1-8	6182	5277	906	Phage protein
BTP1UK1-9	6703	6179	525	Phage EaE protein
BTP1UK1-10	6870	6700	171	Phage protein
BTP1UK1-11	7174	6881	294	FIG01047205: hypothetical protein
BTP1UK1-12	7505	7221	285	Phage anti-RecBCD 1
BTP1UK1-13	8212	7505	708	Putative DNA-binding protein Erf
BTP1UK1-14	8352	8209	144	DNA recombinase, phage-associated
BTP1UK1-15	8530	8342	189	FIG01045708: hypothetical protein
BTP1UK1-16	8669	8511	159	Phage protein
BTP1UK1-17	9004	8879	126	hypothetical protein
BTP1UK1-18	9291	9004	288	Phage protein
BTP1UK1-19	9825	9325	501	Phage-associated homing endonuclease
BTP1UK1-20	10018	10605	588	superinfection exclusion
BTP1UK1-21	10920	10618	303	Phage antitermination protein N
BTP1UK1-22	11977	11273	705	FIG017723: Putative cI prophage repressor protein
BTP1UK1-23	12054	12269	216	putative repressor protein
BTP1UK1-24	12380	12661	282	cII
BTP1UK1-25	12835	13668	834	Replication protein O
BTP1UK1-26	13665	15041	1377	Phage replicative DNA helicase (EC 3.6.1.-)
BTP1UK1-27	15038	15307	270	ORF30
BTP1UK1-28	15379	15651	273	Phage protein
BTP1UK1-29	15661	15870	210	Phage protein
BTP1UK1-30	15882	16175	294	hypothetical protein
BTP1UK1-31	16178	16366	189	hypothetical protein
BTP1UK1-32	16323	16769	447	Phage NinB DNA recombination
BTP1UK1-33	17085	17255	171	Phage NinF
BTP1UK1-34	17248	17859	612	Phage NinG rap recombination
BTP1UK1-35	17856	18080	225	Phage NinY

BTP1UK1-36	18077	18280	204	Phage NinH
BTP1UK1-37	18261	18440	180	Phage Nin protein
BTP1UK1-38	18437	18955	519	Phage antitermination protein Q
BTP1UK1-39	19420	19623	204	Phage holin, class II
BTP1UK1-40	19595	20098	504	Phage lysin (EC 3.2.1.17) # Phage lysozyme or muramidase (EC 3.2.1.17)
BTP1UK1-41	20187	20624	438	Phage endopeptidase
BTP1UK1-42	20774	21379	606	Phage Rha protein
BTP1UK1-43	21903	21643	261	hypothetical protein
BTP1UK1-44	22211	22453	243	Orf80
BTP1UK1-45	22455	22634	180	FIG01047621: hypothetical protein
BTP1UK1-46	22658	23146	489	Phage terminase, small subunit
BTP1UK1-47	23124	24623	1500	Phage terminase, large subunit
BTP1UK1-48	24623	26800	2178	Phage portal protein
BTP1UK1-49	26814	27725	912	Phage capsid and scaffold
BTP1UK1-50	27725	29017	1293	Phage capsid protein
BTP1UK1-51	29058	29618	561	ORF54
BTP1UK1-52	29602	30102	501	DNA stabilization, phage-associated
BTP1UK1-53	30062	31480	1419	DNA stabilization, phage-associated
BTP1UK1-54	31484	32185	702	DNA stabilization, phage-associated
BTP1UK1-55	32185	32640	456	Phage capsid and scaffold
BTP1UK1-56	32643	33332	690	Phage DNA transfer protein
BTP1UK1-57	33342	34676	1335	Phage DNA transfer protein
BTP1UK1-58	34676	36652	1977	Phage DNA transfer protein
BTP1UK1-59	36791	37084	294	lipoprotein, putative
BTP1UK1-60	37353	37105	249	Mnt
BTP1UK1-61	37489	39492	2004	Phage tail spike protein
BTP1UK1-62	41008	39551	1458	O-antigen conversion: translocase
BTP1UK1-63	41192	40998	195	Polymyxin resistance protein ArnC, glycosyl transferase (EC 2.4.-.-)

BTP1^{UK2}

Feature ID	Start	Stop	Length (bp)	Function
BTP1UK2-1	1	1251	1251	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
BTP1UK2-2	2620	1457	1164	Mobile element protein
BTP1UK2-3	2987	2850	138	Formate dehydrogenase N alpha subunit (EC 1.2.1.2) @ selenocysteine-containing
BTP1UK2-4	3586	3059	528	Phage EaA protein
BTP1UK2-5	3779	3588	192	unknown
BTP1UK2-6	4620	3781	840	Phage protein
BTP1UK2-7	5342	4617	726	Phage eae protein
BTP1UK2-8	5729	5436	294	FIG01047205: hypothetical protein
BTP1UK2-9	6060	5776	285	Phage anti-RecBCD 1
BTP1UK2-10	6767	6060	708	Putative DNA-binding protein Erf
BTP1UK2-11	6907	6764	144	DNA recombinase, phage-associated
BTP1UK2-12	7085	6897	189	FIG01045708: hypothetical protein
BTP1UK2-13	7826	7542	285	Phage protein
BTP1UK2-14	8015	7863	153	antirestriction protein
BTP1UK2-15	8271	8858	588	superinfection exclusion
BTP1UK2-16	9173	8871	303	Phage antitermination protein N
BTP1UK2-17	9537	9746	210	hypothetical protein
BTP1UK2-18	10893	9787	1107	Chromosome (plasmid) partitioning protein ParA
BTP1UK2-19	11340	10999	342	Phage protein
BTP1UK2-20	12096	11407	690	COG2932: Predicted transcriptional regulator
BTP1UK2-21	12207	12422	216	Cro
BTP1UK2-22	12533	12814	282	cII
BTP1UK2-23	12988	13887	900	FIG00640106: hypothetical protein
BTP1UK2-24	13877	15313	1437	DNA helicase (EC 3.6.1.-), phage-associated
BTP1UK2-25	15390	15584	195	hypothetical protein
BTP1UK2-26	15584	15874	291	hypothetical protein
BTP1UK2-27	15877	16173	297	Phage protein
BTP1UK2-28	16130	16576	447	Phage NinB DNA recombination
BTP1UK2-29	16713	16889	177	Protein NinE
BTP1UK2-30	16892	17233	342	Phage NinX
BTP1UK2-31	17226	17402	177	Phage NinF
BTP1UK2-32	17395	18000	606	Phage NinG rap recombination
BTP1UK2-33	17997	18221	225	Phage NinY
BTP1UK2-34	18218	18421	204	Phage NinH
BTP1UK2-35	18402	18581	180	Phage Nin protein
BTP1UK2-36	18578	19351	774	Phage antitermination protein Q
BTP1UK2-37	19782	19985	204	Phage holin, class II

BTP1UK2-38	19957	20460	504	Phage lysin (EC 3.2.1.17)
BTP1UK2-39	20457	20924	468	Phage outer membrane lytic protein Rz; Endopeptidase (EC 3.4.-.-)
BTP1UK2-40	21137	21664	528	FIG01048207: hypothetical protein
BTP1UK2-41	21773	22129	357	Orf80
BTP1UK2-42	22133	22522	390	Phage protein
BTP1UK2-43	22522	22926	405	Phage tail fibers
BTP1UK2-44	22930	23418	489	Phage terminase, small subunit
BTP1UK2-45	23396	24895	1500	Phage terminase, large subunit
BTP1UK2-46	24895	27072	2178	Phage portal protein
BTP1UK2-47	27086	27997	912	Phage capsid and scaffold
BTP1UK2-48	27997	29289	1293	Phage capsid protein
BTP1UK2-49	29330	29890	561	ORF54
BTP1UK2-50	29874	30374	501	DNA stabilization, phage-associated
BTP1UK2-51	30367	31752	1386	DNA stabilization, phage-associated
BTP1UK2-52	31756	32457	702	DNA stabilization, phage-associated
BTP1UK2-53	32457	32912	456	Phage capsid and scaffold
BTP1UK2-54	32915	33604	690	Phage DNA transfer protein
BTP1UK2-55	33615	34919	1305	Phage DNA transfer protein
BTP1UK2-56	34919	36832	1914	Phage DNA transfer protein
BTP1UK2-57	37179	36850	330	hypothetical protein
BTP1UK2-58	37210	37575	366	generated by GeneMarkS; orf_3
BTP1UK2-59	37768	37589	180	Phage protein
BTP1UK2-60	38119	37868	252	Mnt
BTP1UK2-61	38255	40258	2004	Phage tail spike protein
BTP1UK2-62	41549	40317	1233	O-antigen conversion: translocase
BTP1UK2-63	42696	41764	933	Polymyxin resistance protein ArnC, glycosyl transferase (EC 2.4.-.-)
BTP1UK2-64	43055	42693	363	Bactoprenol-linked glucose translocase

BTP1^{UK3}

Feature ID	Start	Stop	Length (bp)	Function
BTP1UK3-1	1	1251	1251	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
BTP1UK3-2	2620	1457	1164	Mobile element protein
BTP1UK3-3	2999	2850	150	Formate dehydrogenase N alpha subunit (EC 1.2.1.2) @ selenocysteine-containing
BTP1UK3-4	3453	3058	396	Phage EaA protein
BTP1UK3-5	3930	3457	474	putative methyltransferase
BTP1UK3-6	4379	3930	450	Phage protein
BTP1UK3-7	4680	4381	300	Phage protein
BTP1UK3-8	5075	4677	399	Phage eae protein
BTP1UK3-9	5236	5072	165	Phage protein
BTP1UK3-10	5540	5247	294	FIG01047205: hypothetical protein
BTP1UK3-11	5871	5587	285	Phage anti-RecBCD 1
BTP1UK3-12	6578	5871	708	Putative DNA-binding protein Erf
BTP1UK3-13	6718	6575	144	DNA recombinase, phage-associated
BTP1UK3-14	6896	6708	189	FIG01045708: hypothetical protein
BTP1UK3-15	7035	6877	159	Phage protein
BTP1UK3-16	7367	7242	126	hypothetical protein
BTP1UK3-17	7981	7367	615	Pentapeptide repeat family protein
BTP1UK3-18	8174	8752	579	superinfection exclusion
BTP1UK3-19	9075	8773	303	Phage antitermination protein N
BTP1UK3-20	10079	9429	651	FIG017723: Putative cI prophage repressor protein
BTP1UK3-21	10160	10345	186	FIG00640419: hypothetical protein
BTP1UK3-22	10452	10730	279	cII
BTP1UK3-23	10904	11719	816	Origin specific replication initiation factor
BTP1UK3-24	11716	13092	1377	Phage replicative DNA helicase (EC 3.6.1.-)
BTP1UK3-25	13089	13367	279	Phage protein
BTP1UK3-26	13440	13646	207	Phage protein
BTP1UK3-27	13658	13855	198	hypothetical protein
BTP1UK3-28	13867	14130	264	hypothetical protein
BTP1UK3-29	14133	14351	219	Phage protein
BTP1UK3-30	14308	14754	447	Phage NinB DNA recombination
BTP1UK3-31	14720	14842	123	hypothetical protein
BTP1UK3-32	14891	15067	177	Protein NinE
BTP1UK3-33	15070	15402	333	Phage NinX
BTP1UK3-34	15395	15571	177	Phage NinF
BTP1UK3-35	15564	16175	612	Phage NinG rap recombination
BTP1UK3-36	16172	16396	225	Phage NinY
BTP1UK3-37	16393	16596	204	Phage NinH

BTP1UK3-38	16577	16756	180	Phage Nin protein
BTP1UK3-39	16753	17376	624	Phage antitermination protein Q
BTP1UK3-40	17811	18137	327	Phage holin
BTP1UK3-41	18121	18558	438	Phage endolysin
BTP1UK3-42	18555	19025	471	FIG01046574: hypothetical protein
BTP1UK3-43	19183	19058	126	hypothetical protein
BTP1UK3-44	19238	19759	522	FIG00641140: hypothetical protein
BTP1UK3-45	19969	20325	357	Orf80
BTP1UK3-46	20328	20732	405	Phage tail fibers
BTP1UK3-47	20736	21224	489	Phage terminase, small subunit
BTP1UK3-48	21202	22701	1500	Phage terminase, large subunit
BTP1UK3-49	22701	24878	2178	Phage portal protein
BTP1UK3-50	24892	25803	912	Phage capsid and scaffold
BTP1UK3-51	25803	27095	1293	Phage capsid protein
BTP1UK3-52	27136	27696	561	ORF54
BTP1UK3-53	27680	28180	501	DNA stabilization, phage-associated
BTP1UK3-54	28140	29558	1419	DNA stabilization, phage-associated
BTP1UK3-55	29562	30263	702	DNA stabilization, phage-associated
BTP1UK3-56	30263	30718	456	Phage capsid and scaffold
BTP1UK3-57	30721	31410	690	Phage DNA transfer protein
BTP1UK3-58	31420	32790	1371	Phage DNA transfer protein
BTP1UK3-59	32790	34607	1818	Phage DNA transfer protein
BTP1UK3-60	34954	34625	330	hypothetical protein
BTP1UK3-61	35620	35015	606	Putative prophage repressor protein
BTP1UK3-62	35756	35929	174	hypothetical protein
BTP1UK3-63	35916	36641	726	Phage Rha protein
BTP1UK3-64	36629	37498	870	Phage Rha protein
BTP1UK3-65	37610	39580	1971	Phage tail fibers
BTP1UK3-66	41538	39616	1923	Putative lipopolysaccharide modification acyltransferase
BTP1UK3-67	41746	41865	120	hypothetical protein

BTP1^{UK4}

Feature ID	Start	Stop	Length (bp)	Function
BTP1UK4-1	13	1251	1239	Gamma-glutamyl phosphate reductase (EC 1.2.1.41)
BTP1UK4-2	2620	1457	1164	Mobile element protein
BTP1UK4-3	3451	2858	594	hypothetical protein
BTP1UK4-4	3807	3523	285	FIG01049143: hypothetical protein
BTP1UK4-5	4765	3800	966	Eaa protein
BTP1UK4-6	5531	4776	756	FIG01048215: hypothetical protein
BTP1UK4-7	6175	5531	645	Phage protein
BTP1UK4-8	6582	6172	411	HNH homing endonuclease # Phage intron
BTP1UK4-9	6749	6579	171	Phage protein
BTP1UK4-10	7053	6760	294	FIG01047205: hypothetical protein
BTP1UK4-11	7384	7100	285	Phage anti-RecBCD 1
BTP1UK4-12	8091	7384	708	Putative DNA-binding protein Erf
BTP1UK4-13	8231	8088	144	DNA recombinase, phage-associated
BTP1UK4-14	8409	8221	189	FIG01045708: hypothetical protein
BTP1UK4-15	8866	8741	126	hypothetical protein
BTP1UK4-16	9165	8866	300	Phage protein
BTP1UK4-17	9405	9205	201	antirestriction protein
BTP1UK4-18	9816	9484	333	Phage antitermination protein N
BTP1UK4-19	10180	10389	210	hypothetical protein
BTP1UK4-20	11348	10425	924	FIG01047270: hypothetical protein
BTP1UK4-21	12126	11437	690	COG2932: Predicted transcriptional regulator
BTP1UK4-22	12237	12452	216	Cro
BTP1UK4-23	12563	12844	282	cII
BTP1UK4-24	12879	13025	147	Phage protein
BTP1UK4-25	13018	13851	834	Replication protein O
BTP1UK4-26	13848	15224	1377	Replicative DNA helicase (EC 3.6.1.-)
BTP1UK4-27	15221	15490	270	ORF30
BTP1UK4-28	15562	15834	273	Phage protein
BTP1UK4-29	15844	16053	210	Phage protein
BTP1UK4-30	16065	16358	294	hypothetical protein
BTP1UK4-31	16361	16549	189	hypothetical protein
BTP1UK4-32	16506	16952	447	Phage NinB DNA recombination
BTP1UK4-33	17268	17438	171	Phage NinF
BTP1UK4-34	17431	18042	612	Phage NinG rap recombination
BTP1UK4-35	18039	18263	225	Phage NinY
BTP1UK4-36	18260	18463	204	Phage NinH
BTP1UK4-37	18444	18623	180	Phage Nin protein
BTP1UK4-38	18620	19138	519	Phage antitermination protein Q

BTP1UK4-39	19603	19806	204	Phage holin, class II
BTP1UK4-40	19778	20281	504	Phage lysin (EC 3.2.1.17) # Phage lysozyme or muramidase (EC 3.2.1.17)
BTP1UK4-41	20340	20807	468	Phage endopeptidase
BTP1UK4-42	20957	21562	606	Phage Rha protein
BTP1UK4-43	22085	21825	261	hypothetical protein
BTP1UK4-44	22393	22635	243	Orf80
BTP1UK4-45	22637	22816	180	FIG01047621: hypothetical protein
BTP1UK4-46	22840	23328	489	Phage terminase, small subunit
BTP1UK4-47	23306	24805	1500	Phage terminase, large subunit
BTP1UK4-48	24805	26982	2178	Phage portal protein
BTP1UK4-49	26996	27907	912	Phage capsid and scaffold
BTP1UK4-50	27907	29199	1293	Phage capsid protein
BTP1UK4-51	29249	29800	552	ORF54
BTP1UK4-52	29784	30284	501	DNA stabilization, phage-associated
BTP1UK4-53	30277	31662	1386	DNA stabilization, phage-associated
BTP1UK4-54	31666	32367	702	DNA stabilization, phage-associated
BTP1UK4-55	32367	32822	456	Phage capsid and scaffold
BTP1UK4-56	32825	33514	690	Phage DNA transfer protein
BTP1UK4-57	33524	34894	1371	Phage DNA transfer protein
BTP1UK4-58	34894	36852	1959	Phage DNA transfer protein
BTP1UK4-59	37008	38315	1308	Phage tail fibers

BTP5^{UK1}

Feature ID	Start	Stop	Length (bp)	Function
BTP5UK1-1	1187	1354	168	hypothetical protein
BTP5UK1-2	1400	1603	204	hypothetical protein
BTP5UK1-3	1802	1638	165	putative positive regulator of late gene transcription
BTP5UK1-4	3103	1934	1170	Gene D protein
BTP5UK1-5	3585	3100	486	Phage tail protein
BTP5UK1-6	6041	3600	2442	corresponds to STY4603 from Accession AL513382: Salmonella typhi CT18
BTP5UK1-7	6153	6034	120	putative phage tail protein
BTP5UK1-8	6521	6186	336	Tail protein
BTP5UK1-9	7102	6584	519	Phage major tail tube protein
BTP5UK1-10	8305	7118	1188	Phage tail sheath monomer
BTP5UK1-11	9064	8474	591	Phage tail fibers
BTP5UK1-12	11157	9064	2094	Phage tail fibers
BTP5UK1-13	11698	11168	531	Phage tail fibers
BTP5UK1-14	12599	11691	909	Baseplate assembly protein J
BTP5UK1-15	12953	12606	348	Phage baseplate assembly protein W
BTP5UK1-16	13591	12950	642	Baseplate assembly protein V
BTP5UK1-17	14109	13660	450	Phage tail completion protein
BTP5UK1-18	14569	14102	468	Phage tail protein # FIG072132 and FIG003673
BTP5UK1-19	14705	14532	174	Phage outer membrane lipoprotein Rz1
BTP5UK1-20	15090	14677	414	Phage spanin Rz
BTP5UK1-21	15584	15087	498	Phage endolysin
BTP5UK1-22	15867	15571	297	Phage holin
BTP5UK1-23	16059	15871	189	Phage tail completion protein
BTP5UK1-24	16580	16074	507	Phage head completion-stabilization protein
BTP5UK1-25	17423	16674	750	Phage terminase, endonuclease subunit
BTP5UK1-26	18494	17427	1068	Phage major capsid protein
BTP5UK1-27	19425	18571	855	Phage capsid scaffolding protein
BTP5UK1-28	19591	21360	1770	Phage terminase, ATPase subunit
BTP5UK1-29	21360	22406	1047	Phage-related capsid packaging protein
BTP5UK1-30	22428	22559	132	hypothetical protein
BTP5UK1-31	22933	25152	2220	ATPase involved in DNA repair
BTP5UK1-32	27667	25400	2268	Phage replication protein
BTP5UK1-33	27932	27657	276	FIG00639324: hypothetical protein
BTP5UK1-34	28153	27929	225	Phage protein
BTP5UK1-35	28455	28153	303	FIG00640077: hypothetical protein
BTP5UK1-36	28679	28455	225	FIG00638705: hypothetical protein
BTP5UK1-37	29243	28743	501	Replication gene B protein
BTP5UK1-38	29410	29240	171	FIG00641185: hypothetical protein

BTP5UK1-39	29685	29413	273	Cox
BTP5UK1-40	29820	30113	294	C protein
BTP5UK1-41	30183	31163	981	Phage integrase
BTP5UK1-42	31848	31348	501	P pilus assembly/Cpx signaling pathway, periplasmic inhibitor/zinc-resistance associated protein

BTP5^{UK2}

Feature ID	Start	Stop	Length (bp)	Function
BTP5UK2-1	774	1028	255	hypothetical protein
BTP5UK2-2	2318	1149	1170	Gene D protein
BTP5UK2-3	2800	2315	486	Phage tail protein
BTP5UK2-4	5226	2815	2412	corresponds to STY4603 from Accession AL513382: Salmonella typhi CT18
BTP5UK2-5	5368	5249	120	putative phage tail protein
BTP5UK2-6	5736	5401	336	Tail protein
BTP5UK2-7	6317	5799	519	Phage major tail tube protein
BTP5UK2-8	7520	6333	1188	Phage major tail sheath protein
BTP5UK2-9	8279	7689	591	Phage tail fibers
BTP5UK2-10	10390	8279	2112	Phage tail fibers
BTP5UK2-11	10931	10401	531	Phage tail fibers
BTP5UK2-12	11832	10924	909	Baseplate assembly protein J
BTP5UK2-13	12165	11839	327	Phage baseplate assembly protein W
BTP5UK2-14	12824	12183	642	Baseplate assembly protein V
BTP5UK2-15	13342	12893	450	Phage tail completion protein
BTP5UK2-16	13802	13335	468	Phage tail protein # FIG072132 and FIG003673
BTP5UK2-17	13923	13765	159	Phage outer membrane lipoprotein Rz1
BTP5UK2-18	14323	13910	414	Phage spanin Rz
BTP5UK2-19	14817	14320	498	Phage lysin, 1,4-beta-N-acetylmuramidase (EC 3.2.1.17) or lysozyme #Protein S in phage lambda
BTP5UK2-20	15100	14804	297	Phage holin
BTP5UK2-21	15307	15104	204	Phage tail completion protein
BTP5UK2-22	15813	15307	507	Phage head completion-stabilization protein
BTP5UK2-23	16629	15907	723	Phage terminase, endonuclease subunit
BTP5UK2-24	17727	16660	1068	Phage major capsid protein
BTP5UK2-25	18657	17803	855	Phage capsid scaffolding protein
BTP5UK2-26	18823	20592	1770	Phage terminase, ATPase subunit
BTP5UK2-27	20592	21638	1047	Phage-related capsid packaging protein
BTP5UK2-28	21867	21986	120	hypothetical protein
BTP5UK2-29	22867	21983	885	hypothetical protein
BTP5UK2-30	23816	23085	732	Orf97
BTP5UK2-31	24312	23899	414	SOS operon TUM protein
BTP5UK2-32	26677	24458	2220	Phage replication protein
BTP5UK2-33	27504	26674	831	hypothetical protein
BTP5UK2-34	27786	27508	279	Phage protein
BTP5UK2-35	28367	27783	585	Putative exonuclease
BTP5UK2-36	28591	28364	228	Hypothetical Zinc-finger containing protein
BTP5UK2-37	28773	28591	183	corresponds to STY3665 from Accession AL513382:

				Salmonella typhi CT18
BTP5UK2-38	29227	28889	339	Phage protein
BTP5UK2-39	29391	29191	201	FIG00731938: hypothetical protein
BTP5UK2-40	29908	29399	510	Regulatory protein CII
BTP5UK2-41	30202	29939	264	Phage regulatory protein
BTP5UK2-42	30295	30924	630	Repressor protein CI
BTP5UK2-43	30924	31967	1044	Phage integrase
BTP5UK2-44	32924	32157	768	iron-chelator utilization protein

BTP5^{UK3}

Feature ID	Start	Stop	Length (bp)	Function
BTP5UK3-1	1187	1594	408	hypothetical protein
BTP5UK3-2	1809	1645	165	putative positive regulator of late gene transcription
BTP5UK3-3	3110	1941	1170	Gene D protein
BTP5UK3-4	3592	3107	486	Phage tail protein
BTP5UK3-5	6024	3607	2418	corresponds to STY4603 from Accession AL513382: Salmonella typhi CT18
BTP5UK3-6	6163	6044	120	putative phage tail protein
BTP5UK3-7	6531	6196	336	Tail protein
BTP5UK3-8	7112	6594	519	Phage major tail tube protein
BTP5UK3-9	8315	7128	1188	Phage major tail sheath protein
BTP5UK3-10	9074	8484	591	Phage tail fibers
BTP5UK3-11	11185	9074	2112	Phage tail fibers
BTP5UK3-12	11726	11196	531	Phage tail fibers
BTP5UK3-13	12627	11719	909	Baseplate assembly protein J
BTP5UK3-14	12981	12634	348	Phage baseplate assembly protein W
BTP5UK3-15	13619	12978	642	Baseplate assembly protein V
BTP5UK3-16	14137	13688	450	Phage tail completion protein
BTP5UK3-17	14597	14130	468	Phage tail protein # FIG072132 and FIG003673
BTP5UK3-18	14718	14560	159	Phage outer membrane lipoprotein Rz1
BTP5UK3-19	15118	14705	414	Phage spanin Rz
BTP5UK3-20	15612	15115	498	Phage lysin, 1,4-beta-N-acetylmuramidase (EC 3.2.1.17) or lysozyme #Protein S in phage lambda
BTP5UK3-21	15895	15599	297	Phage holin
BTP5UK3-22	16087	15899	189	Phage tail completion protein
BTP5UK3-23	16608	16102	507	Phage head completion-stabilization protein
BTP5UK3-24	17451	16702	750	Phage terminase, endonuclease subunit
BTP5UK3-25	18522	17455	1068	Phage major capsid protein
BTP5UK3-26	19453	18599	855	Phage capsid scaffolding protein
BTP5UK3-27	19619	21391	1773	Phage terminase, ATPase subunit
BTP5UK3-28	21388	22134	747	Phage terminase, ATPase subunit
BTP5UK3-29	22131	23153	1023	Phage-related capsid packaging protein
BTP5UK3-30	24440	23205	1236	putative DNA-binding protein
BTP5UK3-31	24680	24453	228	hypothetical protein
BTP5UK3-32	24666	24797	132	hypothetical protein
BTP5UK3-33	25101	24865	237	Phage protein
BTP5UK3-34	25277	25101	177	FIG00640898: hypothetical protein
BTP5UK3-35	27571	25283	2289	Phage replication protein
BTP5UK3-36	27836	27561	276	FIG00639324: hypothetical protein
BTP5UK3-37	28057	27833	225	Phage protein

BTP5UK3-38	28358	28062	297	FIG00640077: hypothetical protein
BTP5UK3-39	28582	28358	225	FIG00638705: hypothetical protein
BTP5UK3-40	29146	28646	501	Replication gene B protein
BTP5UK3-41	29259	29137	123	hypothetical protein
BTP5UK3-42	29588	29316	273	Cox
BTP5UK3-43	29725	30018	294	C protein
BTP5UK3-44	30088	31068	981	Integrase
BTP5UK3-45	31753	31253	501	P pilus assembly/Cpx signaling pathway, periplasmic inhibitor/zinc-resistance associated protein

Appendix III

Accession number and metadata of the ST313 isolates from the Public Health England used in this thesis. Empty cells or N/A indicate no data available.

ID	Accession number	Lineage	PHE centre (if UK)	Travel to sSA	country associated with	travel info available	Sample Type	Isolate source	Report Date
U10	SRR1965263	UK	Yorkshire and Humber	N	UK	Y	Human	Faeces	03/06/2014
U11	SRR1965387	UK	West Midlands	N	UK	Y	Human	Faeces	05/06/2014
U12	SRR1967419	UK	London	N	UK	Y	Human	N/A	19/06/2014
U14	SRR1966418	UK	East of England	N	UK	Y	Human	Blood	26/08/2014
U15	SRR1645388	UK	East Midlands	N	UK	N	Human	Faeces	24/05/2012
U16	SRR1645171	UK	Yorkshire and Humber	N	UK	N	Human	Faeces	15/06/2012
U2	SRR1646230	UK	East of England	N	UK	Y	Human	Faeces	02/11/2012
U22	SRR3048892	UK	South East	N	UK	N	Human	Faeces	18/06/2014
U23	SRR1969854	UK	South East	N	UK	Y	Human	Faeces	10/10/2014
U24	SRR1967447	UK	London	N	UK	Y	Human	Faeces	23/10/2014
U25	SRR1958036	UK	London	N	UK	N	Human	Faeces	10/11/2014
U27	SRR1958303	UK	London	Y	Uganda	Y	Human	Faeces	07/11/2014
U28	SRR1960363	UK	None	N	UK	N	Human	Faeces	10/12/2014
U29	SRR1968084	UK	London	N	UK	N	Human	Faeces	26/01/2015
U3	SRR1646238	UK	East of England	N	UK	N	Human	Faeces	10/12/2012
U30	SRR1968221	UK	#N/A	N	UK	N	Food	Raw Produce	24/02/2015
U31	SRR1967938	UK	London	N	UK	N	Human	Faeces	18/02/2015

U32	SRR1969492	UK	London	N	UK	N	Human	Blood	25/02/2015
U34	SRR1966635	UK	West Midlands	N	UK	Y	Human	Faeces	05/03/2015
U36	SRR1966593	UK	London	N	UK	Y	Human	Faeces	05/03/2015
U37	SRR1968962	UK	East of England	N	UK	Y	Human	Faeces	12/03/2015
U38	SRR1967438	UK	Yorkshire and Humber	N	UK	Y	Human	Faeces	17/03/2015
U4	SRR1645545	UK	West Midlands	N	UK	Y	Human	Faeces	20/12/2012
U40	SRR1968061	UK	London	N	UK	N	Human	Faeces	18/03/2015
U43	SRR5451258	UK	Yorkshire and Humber	N	UK	N	Human	Faeces	10/06/2015
U44	SRR5451287	UK	London	N	Pakistan	Y	Human	Faeces	26/05/2015
U46	SRR5451286	UK	London	N	UK	Y	Human	Faeces	12/06/2015
U47	SRR5451276	UK	South West	N	UK	Y	Human	Faeces	16/06/2015
U48	SRR5451252	UK	Yorkshire and Humber	N	UK	Y	Human	Faeces	29/06/2015
U49	SRR5451278	UK	London	N	UK	N	Human	Faeces	30/06/2015
U5	SRR1645768	UK	Wales	N	UK	N	Human	Faeces	21/12/2012
U50	SRR5451285	UK	London	N	UK	N	Human	Faeces	02/07/2015
U51	SRR3323011	UK	London	N	UK	Y	Human	Faeces	06/08/2015
U52	SRR3284779	UK	London	N	UK	Y	Human	Faeces	06/08/2015
U53	SRR3285477	UK	East of England	N	UK	N	Human	Faeces	07/08/2015
U54	SRR5193685	UK	South East	N	UK	Y	Human	Faeces	10/08/2015
U55	SRR3322345	UK	North West	N	UK	Y	Human	Faeces	08/09/2015
U56	SRR5451268	UK	Yorkshire and Humber	N	Pakistan	Y	Human	Faeces	12/10/2015

U57	SRR5451279	UK	London	N	UK	Y	Human	Faeces	25/09/2015
U58	SRR5451270	UK	South West	N	UK	Y	Human	Faeces	01/10/2015
U59	SRR5451254	UK	East of England	N	UK	N	Human	Faeces	19/10/2015
U61	SRR5451275	UK	South East	N	UK	Y	Human	Faeces	27/10/2015
U62	SRR5451283	UK	East of England	N	UK	N	Human	Faeces	19/10/2015
U63	SRR5451281	UK	London	N	UK	N	Human	Blood	22/10/2015
U64	SRR5451266	UK	London	N	UK	Y	Human	Faeces	27/10/2015
U65	SRR5451255	UK	London	N	Portugal	Y	Human	Faeces	28/10/2015
U66	SRR3322379	UK	London	N	UK	N	Human	Faeces	03/11/2015
U69	SRR5451261	UK	#N/A	N	UK	N	Dog	Dog	22/12/2015
U7	SRR1966408	UK	East of England	N	UK	Y	Human	Faeces	14/05/2014
U70	SRR5451259	UK	London	N	UK	N	Human	Faeces	04/01/2016
U74	SRR5451280	UK	North West	N	Cape verde	Y	Human	Faeces	03/02/2016
U75	SRR5451257	UK	East of England	N	Belize	Y	Human	Faeces	26/02/2016
U76	SRR5451251	UK	South East	N	India	Y	Human	Faeces	02/03/2016
U77	SRR5451260	UK	East of England	N	UK	Y	Human	Faeces	29/02/2016
U78	SRR5451265	UK	West Midlands	N	India	Y	Human	Faeces	11/03/2016
U8	SRR1966656	UK	London	N	UK	Y	Human	Faeces	29/05/2014
U80	SRR5451253	UK	East of England	N	UK	Y	Human	Faeces	21/03/2016
U81	SRR5215816	UK	East of England	N	UK	Y	Human	Faeces	29/03/2016
U82	SRR5451272	UK	London	N	UK	Y	Human	Faeces	04/04/2016
U83	SRR5451267	UK	South East	N	UK	Y	Human	Faeces	22/04/2016
U84	SRR5451256	UK	East of England	N	UK	Y	Human	Faeces	22/04/2016
U86	SRR5451274	UK	South East	N	UK	Y	Human	Faeces	26/05/2016
U87	SRR5451262	UK	North West	N	India	Y	Human	Faeces	26/05/2016

U9	SRR1968363	UK	East of England	N	UK	Y	Human	N/A	30/05/2014
U1	SRR1646227	II	London	Y	Ghana	Y	Human	Blood	28/08/2012
U13	SRR1966509	II	London	Y	Nigeria	Y	Human	Blood	21/07/2014
U17	SRR1967883	II	London	Y	African continent (?)	Y	Human	Blood	07/05/2014
U35	SRR1966023	II	London	Y	Nigeria	Y	Human	Blood	04/03/2015
U42	SRR3285401	II	London	N	UK (African name)	N	Human	Blood	27/05/2015
U45	SRR5451277	II	London	N	UK (STD clinic)	N	Human	Blood	17/07/2015
U60	SRR5451282	II	London	Y	Kenya	Y	Human	Blood	08/10/2015
U67	SRR3323046	II	East Midlands	Y	Nigeria	Y	Human	Blood	23/11/2015
U68	SRR5451264	II	London	N	UK (African name)	N	Human	Pus	10/12/2015
U72	SRR5451284	II	London	N	UK	Y	Human	N/A	03/02/2016
U73	SRR5451273	II	London	N	UK (African name, African markets)	Y	Human	Blood	03/02/2016
U85	SRR5451269	II	London	Y	Ghana	Y	Human	Faeces	09/05/2016

Appendix IV

Sources and NCBI accession number of all genome sequences used in Chapter 6 are listed below:

NCBI Accession Number	Genome Sequence
LS997975.1	Salmonella enterica subsp. enterica serovar Typhimurium strain D23580 genome assembly, plasmid: D23580_liv_pBT1
CP029996.1	Salmonella enterica subsp. salamae serovar 56:b:[1,5] strain SA20053897 plasmid pSA20053897.1, complete sequence
CP022141.1	Salmonella enterica subsp. salamae serovar 55:k:z39 str. 1315K plasmid unnamed2, complete sequence
CP022036.1	Salmonella enterica subsp. enterica serovar Onderstepoort str. SA20060086 plasmid punamed1, complete sequence
JQ418539.1	Salmonella sp. 40 plasmid p40-95A, complete sequence
JQ418537.1	Salmonella sp. 14 plasmid p14-95A, complete sequence
CP053317.1	Salmonella enterica subsp. salamae serovar 6,8:a:z52 strain 62-3163 plasmid unnamed, complete sequence
CP034708.1	Salmonella enterica subsp. enterica serovar Waycross strain RSE24 plasmid pRSE24, complete sequence
CP022136.1	Salmonella enterica subsp. diarizonae serovar 65:c:z str. SA20044251 plasmid unnamed1, complete sequence
CP069804.1	Citrobacter sp. B72 plasmid pB72C, complete sequence
CP069801.1	Citrobacter sp. B72 chromosome, complete genome