The Molecular Epidemiology of Enteric Fever in South and Southeast Asia

Thesis

How to cite:

For guidance on citations see FAQs.

© 2018 The Author

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.
THE MOLECULAR EPIDEMIOLOGY OF ENTERIC FEVER IN
SOUTH AND SOUTHEAST ASIA

Pham Thanh Duy

A thesis submitted to the Open University U.K
For the degree of Doctor of Philosophy in the field of Life Sciences

Oxford University Clinical Research Unit
Hospital for Tropical Diseases
Ho Chi Minh City, Viet Nam
April, 2018
Abstract

Typhoid fever is a life-threatening systemic infection caused by *Salmonella enterica* sub-species *enterica* serovars Typhi (*S. Typhi*) and Paratyphi A (*S. Paratyphi A*). While the disease is mainly travel-associated in developed countries, it still causes significant burden in the poorest areas in developing countries where safe water and adequate sanitation and food hygiene remain limited. Typhoid management largely relies on antimicrobial therapy; however, antimicrobial resistance (AMR) in these causative pathogens has become a global threat, compromising the effectiveness of the treatment therapy and signifying the burden of this disease. Understanding the epidemiology of typhoid fever in different endemic settings as well as the impact of AMR on the disease outcome is crucial for disease control and management.

First, this thesis utilized whole genome sequences of *S. Typhi* combined with clinical data from a randomized controlled trial to investigate the impact of AMR and bacterial genotype on the disease outcome. A novel subclade of ciprofloxacin-resistant H58 *S. Typhi* associated with increased treatment failure was identified and these organisms were likely widespread in Indian subcontinent. Subsequently, this study combined bacterial genomics with conventional epidemiological tools to reveal the population structure and spatiotemporal dynamics of *S. Paratyphi A* isolates in Nepal. The Nepalese *S. Paratyphi A* population was highly dynamic with evidences of regular inter-country transmission, clonal expansion and replacement of distinct genotypes during the study period. A number of localized spatiotemporal clusters of
S. Paratyphi A cases were also identified. A molecular epidemiological investigation was also performed to provide insights into the AMR, epidemiological features and population structure and dynamics of S. Typhi in rural areas of Siem Reap, Cambodia. A substantial burden of pediatric typhoid fever was revealed and communes with high-risk of infection were identified. Multidrug resistant H58 S. Typhi with reduced susceptibility to fluoroquinolones was dominant in this setting. This study also investigated the phylogenetic relationship between acute and carriage S. Typhi isolates in Nepal and deciphered the genetic characteristics associated with carriage isolates. My study suggested that typhoid carriage was likely not an important source of new infections in endemic area.
Acknowledgements

First and foremost, I would like to express my deep gratitude to my PhD supervisor, Professor Stephen Baker, who has been supervising, teaching and guiding my scientific career since the first day I came into the science world. He has broadened my mind, helped me to expand my horizons, and constantly pushed me forward to become a better scientist. I have received timely and valuable encouragement, support, advice, and guidance from him during my PhD journey.

I have also received lots of guidance and support from my supervisor team, Dr Maia Rabaa, Professor Christopher Parry and Professor Gordon Dougan. Particularly, I would like to express my gratitude to Maia, who has taught me many useful things about phylogenetics and spent time chasing me up and editing my PhD thesis.

During my PhD, there are many colleagues and collaborators who have provided support for my research in terms of data analysis, project delivery, bioinformatics training, advice and guidance. I would like to extend my thanks to Dr Corrine Thompson, Dr Nhan Ho Thi (OUCRU-VN), Dr Abhilasha Karkey, Dr Sabina Dongol, Dr Buddah Basnyat (OUCRU-Nepal), Professor Nicholas Thompson (Wellcome Trust Sanger Institute), Dr Kathryn Holt (University of Melbourne) and Dr Paul Turner (COMRU).
I also really appreciate all the assistance from my colleagues in the enteric infections group (OUCRU-VN) during my work in the lab. Thank you for listening to me and supporting me with your encouraging messages throughout writing this thesis.

Finally, my deepest gratitude goes to my parents, who have wholeheartedly and unconditionally supported me. I would like to thank my dad for guiding me with his strength and diligence and my mom for her wisdom.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>µg</td>
<td>Microgram</td>
</tr>
<tr>
<td>µm</td>
<td>Micromole</td>
</tr>
<tr>
<td>µl</td>
<td>Microliter</td>
</tr>
<tr>
<td>AIDS</td>
<td>Acquired Immune Deficiency Syndrome</td>
</tr>
<tr>
<td>AHC</td>
<td>Angkor Hospital for Children</td>
</tr>
<tr>
<td>AMR</td>
<td>Antimicrobial resistance</td>
</tr>
<tr>
<td>ATCC</td>
<td>American Type Culture Collection</td>
</tr>
<tr>
<td>BiP</td>
<td>Biallelic Polymorphisms</td>
</tr>
<tr>
<td>bp</td>
<td>Base pairs</td>
</tr>
<tr>
<td>CDC</td>
<td>Centers for Disease Control and Prevention</td>
</tr>
<tr>
<td>CLSI</td>
<td>Clinical and Laboratory Standards Institute</td>
</tr>
<tr>
<td>cfu</td>
<td>Colony Forming Unit</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribose nucleic acid</td>
</tr>
<tr>
<td>ESBL</td>
<td>Extended Spectrum Beta Lactamase</td>
</tr>
<tr>
<td>FCT</td>
<td>Fever Clearance Time</td>
</tr>
<tr>
<td>FQ</td>
<td>Fluoroquinolone</td>
</tr>
<tr>
<td>GPS</td>
<td>Global Positioning System</td>
</tr>
<tr>
<td>GTR</td>
<td>General Time Reversible</td>
</tr>
<tr>
<td>HLA</td>
<td>Human Leukocyte Antigen</td>
</tr>
<tr>
<td>ICD</td>
<td>International Classification of Diseases</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>IVI</td>
<td>International Vaccine Institute</td>
</tr>
<tr>
<td>Kb</td>
<td>Kilobase pairs</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharides</td>
</tr>
<tr>
<td>LSMC</td>
<td>Lalitpur Sub-Metropolitan City</td>
</tr>
<tr>
<td>MDa</td>
<td>Megadalton</td>
</tr>
<tr>
<td>MDR</td>
<td>Multidrug resistant</td>
</tr>
<tr>
<td>MIC</td>
<td>Minimum Inhibitory Concentration</td>
</tr>
<tr>
<td>MLEE</td>
<td>Multilocus enzyme electrophoresis</td>
</tr>
<tr>
<td>MLST</td>
<td>Multilocus sequence typing</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>SNP</td>
<td>Single Nucleotide Polymorphism</td>
</tr>
<tr>
<td>SPI</td>
<td>Salmonella Pathogenicity Island</td>
</tr>
<tr>
<td>WGS</td>
<td>Whole genome sequencing</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>
Table of Contents

Abstract........................................................................................................................................ii
Acknowledgement................................................................................................................iv
Abbreviations........................................................................................................................vi
List of Figures..........................................................................................................................xii
List of Tables..........................................................................................................................xiv

Chapter 1 Introduction.............................................................................................................1
  1.1 The genus Salmonella .................................................................................................1
  1.2 Salmonella serology .............................................................................................2
  1.3 Salmonella nomenclature ......................................................................................4
  1.4 Salmonella, the host and disease .........................................................................5
  1.5 Typhoid fever ........................................................................................................6
    1.5.1 Clinical features of typhoid fever ................................................................6
    1.5.2 Pathogenesis of typhoid fever ......................................................................8
    1.5.3 Typhoid diagnosis ......................................................................................10
    1.5.4 Typhoid treatment .....................................................................................12
  1.6 The epidemiology of typhoid fever .........................................................................14
    1.6.1 Epidemiological features of typhoid fever .................................................14
    1.6.2 Typhoid carriage ........................................................................................15
    1.6.3 The global burden of typhoid fever ............................................................16
    1.6.4 Typhoid control and prevention ...............................................................18
  1.7 Antimicrobial resistance in Typhoidal Salmonella ..............................................19
    1.7.1 Global dissemination of multidrug resistant Typhoidal Salmonella .........19
    1.7.2 Emergence of fluoroquinolone resistance ..................................................21
    1.7.3 Resistance to third generation cephalosporins and macrolides ..........23
  1.8 The convergent evolution of Salmonella Typhi and Salmonella Paratyphi A 24
  1.9 Genomics and molecular epidemiology of typhoid fever ..................................26
1.9.1 Genomics and its role in molecular epidemiological studies in Asia and Africa ......................................................... 26
1.9.2 The origin and global dissemination of H58 Salmonella Typhi ................. 29
1.10 Aims of this study .......................................................................................................................... 31

Chapter 2 Methods .......................................................................................................................... 33

2.1 Study sites and settings ................................................................................................................. 33
  2.1.1 Patan Hospital in Kathmandu, Nepal ........................................................................ 33
  2.1.2 Angkor Hospital for Children in Siem Reap, Cambodia ........................................ 34

2.2 Bacterial identification and antimicrobial susceptibility testing ............................................ 34
  2.2.1 Blood culture .................................................................................................................. 34
  2.2.2 Bile and stool culture .................................................................................................... 35
  2.2.3 Antimicrobial susceptibility testing ............................................................................. 35

2.3 Vi agglutination assay ............................................................................................................... 36

2.4 Data sources and bacterial isolates .......................................................................................... 36
  2.4.1 Data sources and bacterial isolates in chapter 3 ....................................................... 36
  2.4.2 Data sources and bacterial isolates in chapter 4 ....................................................... 38
  2.4.3 Data sources and bacterial isolates in chapter 5 ....................................................... 39
  2.4.4 Data sources and bacterial isolates in chapter 6 ....................................................... 40

2.5 Whole genome sequencing ....................................................................................................... 40
  2.5.1 DNA extraction ............................................................................................................. 40
  2.5.2 DNA quantification ....................................................................................................... 42
  2.5.3 DNA library preparation ............................................................................................... 43
    2.5.3.1 Tagment genomic DNA .......................................................................................... 43
    2.5.3.2 Clean up Tagmented DNA ................................................................................ 44
    2.5.3.3 Amplify Tagmented DNA ...................................................................................... 45
    2.5.3.4 Clean up libraries ................................................................................................ 46
    2.5.3.5 Library quantification ........................................................................................... 47
    2.5.3.6 Normalize and Pool libraries .............................................................................. 48
  2.5.4 Perform a run on Miseq ................................................................................................. 49

2.6 Single Nucleotide Polymorphism (SNP) detection and analysis ......................................... 50
2.6.1 SNP detection and analysis for Salmonella Typhi genomes ....................... 50
  2.6.1.1 SNP detection and annotation ............................................................... 50
  2.6.1.2 Salmonella Typhi genotyping ............................................................... 51
  2.6.1.3 Functional analysis ................................................................................ 54
  2.6.1.4 Pairwise SNP distance .......................................................................... 54
2.6.2 SNP detection and analysis for Salmonella Paratyphi A genomes .......... 56
2.7 Phylogenetic analysis ..................................................................................... 56
  2.7.1 Phylogenetic analysis for chapter 3 ............................................................. 56
  2.7.2 Phylogenetic analysis for chapter 4 ............................................................. 57
  2.7.3 Phylogenetic analysis for chapter 5 ............................................................. 57
  2.7.4 Phylogenetic analysis for chapter 6 ............................................................. 58
2.8 Resistance gene and plasmid analysis ............................................................. 59
2.9 Pan-genome analysis ....................................................................................... 59
2.10 Spatiotemporal cluster analysis .................................................................... 60
  2.10.1 Cluster analysis for chapter 4 ................................................................. 60
  2.10.2 Spatiotemporal clustering detection for chapter 5 ..................................... 60
2.11 Statistical analysis ......................................................................................... 61
  2.11.1 Statistical analysis for chapter 3 ............................................................. 61
  2.11.2 Statistical analysis for chapter 5 ............................................................. 62

Chapter 3 Emergence of a novel ciprofloxacin-resistant subclade of H58
Salmonella Typhi associated with fluoroquinolone treatment failure in Nepal... 63

3.1 Introduction .................................................................................................... 63
3.2 Results .......................................................................................................... 65
  3.2.1 Salmonella Typhi whole genome sequencing .......................................... 65
  3.2.2 Clinical presentation of Salmonella Typhi infections ................................ 67
  3.2.3 Treatment failure and fever clearance times ............................................ 71
  3.2.4 The emergence of fluoroquinolone-resistant Salmonella Typhi ............ 78
3.3 Discussion ..................................................................................................... 80

Chapter 4 The phylogenetics and spatiotemporal dynamics of Salmonella
Paratyphi A in Kathmandu, Nepal ..................................................................... 83
4.1 Introduction ........................................................................................................ 83
4.2 Results ................................................................................................................ 87
  4.2.1 Baseline characteristics ............................................................................... 87
  4.2.2 Antimicrobial resistance .............................................................................. 90
  4.2.3 The population structure and dynamics of S. Paratyphi A isolates in Nepal .............................................................................................................................. 92
  4.2.4 Genetic relatedness between acute and carriage S. Paratyphi A isolates in Nepal ........................................................................................................................................ 100
  4.2.5 Spatial and spatiotemporal distribution of S. Paratyphi A genotypes ......... 102
    4.2.5.1 Genotypic subgrouping ....................................................................... 102
    4.2.5.2 Spatial and spatiotemporal mapping of S. Paratyphi A genotypes .... 108
4.3 Discussion ........................................................................................................ 111

Chapter 5 The molecular and spatial epidemiology of typhoid fever in rural Cambodia ........................................................................................................... 116
  5.1 Introduction ...................................................................................................... 116
  5.2 Results .............................................................................................................. 118
    5.2.1 Baseline characteristics.............................................................................. 118
    5.2.2 Spatiotemporal clustering of typhoid fever cases...................................... 122
    5.2.3 The population structure of Salmonella Typhi in Siem Reap province, Cambodia ......................................................................................................................... 125
    5.2.4 The spatiotemporal distribution of Salmonella Typhi genotypes.......... 127
    5.2.5 Population risk factors for typhoid fever................................................. 129
  5.3 Discussion ........................................................................................................ 131

Chapter 6 Genetic traits of Salmonella Typhi gallbladder carriage isolates and their role in disease transmission in Kathmandu, Nepal ......................... 135
  6.1 Introduction ...................................................................................................... 135
  6.2 Results .............................................................................................................. 137
    6.2.1 The phylogenetic structure of Nepalese acute and carrier Salmonella Typhi isolates between 2007 and 2010 ................................................................. 137
    6.2.2 Dissecting the genetic traits of Salmonella Typhi carriage isolates ......... 140
6.2.3 Positive selection associated with typhoid carriage .............................. 144
6.2.4 Estimating the role of typhoid carriage in disease transmission in Kathmandu, Nepal ................................................................. 146
6.3 Discussion ......................................................................................... 152

Chapter 7 General discussion ........................................................................ 159
Chapter 8 References .................................................................................. 164
Chapter 9 Appendices .................................................................................. 194
Appendix A  *Salmonella* Typhi isolates and their corresponding sequencing metadata in chapter 3 ................................................................. 194
Appendix B  *Salmonella* Paratyphi A isolates and their corresponding sequencing metadata in chapter 4 ......................................................... 200
Appendix C  *Salmonella* Typhi isolates and their corresponding sequencing metadata in chapter 5 ................................................................. 209
Appendix D  *Salmonella* Typhi isolates and their corresponding sequencing metadata in chapter 6 ................................................................. 221
Appendix E  Acute-specific nonsynonymous mutations and their functional classes ....................................................................................... 226
Appendix F  Carrier-specific nonsynonymous mutations and their functional classes ....................................................................................... 246
Appendix G  Papers published on aspects of this thesis ................................. 256
List of Figures

**Figure 3.1** The phylogenetic structure of 78 Nepali *Salmonella* Typhi isolated during a gatifloxacin versus ceftriaxone randomised controlled trial .........................................................66

**Figure 3.2** The association of *Salmonella* Typhi lineage and ciprofloxacin susceptibility with treatment failure and fever clearance time in patients randomised to gatifloxacin ..............................................................................................................77

**Figure 3.3** The phylogenetic structure of fluoroquinolone resistant *Salmonella* Typhi H58 in a regional context..................................................................................................................79

**Figure 4.1** Locations of *Salmonella* Paratyphi A isolates in Kathmandu ...............89

**Figure 4.2** Minimal inhibitory concentrations (MICs) of *S*. Paratyphi A isolates to various antimicrobials over time...............................................................................................................91

**Figure 4.3** Phylogenetics of Nepalese *S*. Paratyphi A isolates in a global context....94

**Figure 4.4** Novel sopE prophage of sub-lineage A1 ...................................................95

**Figure 4.5** Minimal spanning tree showing different clonal clusters within sub-lineages A1 and A2 .................................................................................................................................103

**Figure 4.6** Annual distribution of *S*. Paratyphi A genotypes in Nepal......................106

**Figure 4.7** Clustering of *S*. Paratyphi A genotypes over space between 2005 and 2014 ....................................................................................................................................110

**Figure 5.1** The annual and seasonal distribution of typhoid fever cases at Angkor Hospital for Children in Cambodia.................................................................121

**Figure 5.2** The spatial distribution of typhoid fever cases in Siem Reap province, Cambodia ......................................................................................................................124

**Figure 5.3** The phylogenetic structure of the H58 lineage of Cambodian *Salmonella* Typhi ........................................................................................................................................126

**Figure 5.4** The spatiotemporal distribution of the various *Salmonella* Typhi lineages/sublineages in Siem Reap province, Cambodia ..................................................128

**Figure 6.1** The phylogenetic structure of carriage and acute *S*. Typhi isolates collected between 2007 and 2010 .................................................................................................139
Figure 6.2 Top ten functional classes with highest prevalence of acute-specific nonsynonymous SNPs versus carrier-specific nonsynonymous SNPs ...................... 143

Figure 6.3 Phylogenetic structure of acute and carrier 4.3.1 Salmonella Typhi isolates from Nepal in the global context .............................................................. 148

Figure 6.4 Distribution of pairwise SNP distances within and between acute and carrier isolates ................................................................. 151
List of Tables

**Table 2.1** Canonical SNPs for genotyping *Salmonella* Typhi into Clades and Subclades ........................................................................................................................................52

**Table 2.2** *Salmonella* Typhi functional classification scheme .............................................55

**Table 3.1** Baseline characteristics by *Salmonella* Typhi lineage ........................................68

**Table 3.2** Baseline characteristics by *Salmonella* Typhi ciprofloxacin susceptibility ..........................69

**Table 3.3** Comparison of antimicrobial susceptibility by *Salmonella* Typhi lineage ..........70

**Table 3.4** Summary of time to treatment failure and fever clearance time by *Salmonella* Typhi lineage ...............................................................................................................................................73

**Table 3.5** Treatment failure in detail by *Salmonella* Typhi lineage in the gatifloxacin treatment group ......................................................................................................................................................74

**Table 3.6** Summary of time to treatment failure and fever clearance time by ciprofloxacin susceptibility .........................................................................................................................................................75

**Table 3.7** Treatment failure in detail by ciprofloxacin susceptibility in the gatifloxacin treatment group ......................................................................................................................................................76

**Table 4.1** Nonsynonymous mutations and indels associated with sub-lineage A1 ...............96

**Table 4.2** Nonsynonymous mutations and indels associated with sub-lineage A2 ...............99

**Table 4.3** Characteristics of two main sub-lineages A1 and A2 of *S. Paratyphi* A in Nepal ..............................................................................................................................................................................107

**Table 5.1** Baseline characteristics of all communes and those with at least one case of typhoid fever ..........................................................................................................................................................120

**Table 5.2** Regression results of highlighting factors associated with typhoid cases .................................130
Chapter 1

Introduction

1.1 The genus *Salmonella*

The genus *Salmonella* belongs to the family *Enterobacteriaceae*, which is a large and diverse group of bacteria found in soil, water, waste, plants, and the gastrointestinal tracts of animals \(^1\). This large family is composed of genetically and phenotypically closely related bacteria.

*Salmonella* are facultatively anaerobic, non-spore forming, Gram-negative bacilli, 2 to 5 μm long and 0.8 to 1.5 μm wide, usually motile by peritrichous flagella. The bacteria are catalase positive, oxidase negative, ferment glucose, mannitol and sorbitol to produce acid and gas, and can use citrate as a sole carbon source. They also do not hydrolyze urea or deaminate phenylalanine, but usually produce hydrogen sulfide on triple sugar iron agar. *Salmonella* yield negative Voges-Proskauer and positive methyl red tests and do not produce cytochrome oxide. Most *Salmonella* are positive for lysine decarboxylase and ornithine decarboxylase, except for *Salmonella* Typhi and *Salmonella* Paratyphi A. *Salmonella* is most closely related to *Escherichia coli* (E. coli); the geneses are thought to have shared a common ancestor more than 100 million years ago \(^2\). During their evolution, *E. coli* have become lactose positive while *Salmonella* are generally identified as being non-lactose fermenting \(^3\). *Salmonella* can be identified from biochemical tests based on lactose fermentation,
acid and gas production from glucose, mannitol, maltose, sorbitol, and the production of hydrogen sulfide.

1.2 *Salmonella* serology

The *Salmonella* are traditionally classified into different serotypes (also referred to as serovars) by the Kauffmann-White scheme, the system was established to categorise all identified *Salmonella* serotypes. The documentation of this scheme is updated annually by the World Health Organization’s (WHO) Collaborating Centre for Reference and Research on *Salmonella* at the Pasteur Institute, Paris, France. Up until 2002, The U.S. Centers for Disease Control and Prevention (CDC) used a marginally different version of the scheme, but have since also adopted the Kauffmann-White Scheme.

*Salmonella* serotyping is based on antibody agglutination reactions with the bacterial surface structures, the O antigen (somatic) and H antigen (flagella). The O antigen is a polysaccharide polymerized from O subunits, with each subunit typically comprised of four to six sugars depending on the O antigen. It is the outermost component of the lipopolysaccharides (LPS) found in the outer membrane of Gram-negative bacteria. Variation in the sugar components of the subunit, the covalent bond between the sugars of the subunit or the linkage between O subunits that form the O antigen polymer results in O antigen diversity. O antigens are classified as primary O factors that define O serogroups (O groups) and additional O factors that are often variably present or expressed in some O groups. For instance, *S. Typhi* possess type 9 and 12
O antigens and belong to serogroup O:9 or D1, whereas *S. Paratyphi* A carry type 1, 2 and 12 O antigens and belong to serogroup O:2 or A.

The H antigen is made up of protein subunits called flagellin and is the filamentous portion of the bacterial peritrichous flagellar. *Salmonella* is unique among enteric bacteria as it can express two forms (phases) of H antigens, which are encoded by different genes. The expression of these two genes is regulated so that only one flagellar antigen is expressed at a time in a single bacterial cell, a phenomenon known as phase variation. Most *Salmonella* are diphasic and can express both phase 1 and phase 2 H antigen; however, some *Salmonella* serovars like *S. Typhi*, *S. Paratyphi* A and *S. Enteritidis* are naturally monophasic and can only express a single flagellin type due to disruption of the gene encoding phase 1 or phase 2 antigen.

Identification can also be performed by agglutination with another surface antigen, such as the virulence (Vi) polysaccharide capsular antigen that can be found on some *Salmonella* serovars, including *S. Typhi*, *S. Paratyphi* C and *S. Dublin*. The Vi capsule typically blankets the O antigen and blocks O agglutination but can be selectively removed by heat treatment prior to the O agglutination assay. Vi agglutination is used widely to identify *S. Typhi* isolates. As of 2002, the Kauffmann-White scheme was comprised of 2,541 serotypes. This classification scheme is essential for laboratory and clinical identification of the *Salmonella* as well as for international communication between scientists, health officials, and the public. Although more advanced molecular typing techniques have been introduced, *Salmonella* serotyping
remains important in epidemiological surveillance and outbreak investigation and is widely used in most microbiology laboratories in the world\textsuperscript{9,10}.

1.3 \textit{Salmonella} nomenclature

The nomenclature of \textit{Salmonella} has been controversial as the original classification of the genus was not based on DNA relatedness; rather, names were originally given according to the particular disease caused, the animal from which the organism was isolated (e.g., \textit{S. Typhi} causing human typhoid and \textit{S. Typhimurium} causing typhoid-like in murine), or by the geographical location where the strain was first identified (e.g., \textit{S. Montevideo}, \textit{S. Newport}). Classification of \textit{Salmonella} evolved over time beginning from the one serotype-one species concept that was originally proposed by Kauffmann in 1966\textsuperscript{11}. However if this concept were used today, it would result in the distinction of over 2,500 species.

Several classical \textit{Salmonella} nomenclatural systems were subsequently proposed and inconsistently and confusingly divided the genus \textit{Salmonella} into species, subspecies, subgenera, groups, subgroups and serotypes\textsuperscript{12}. The turning point for \textit{Salmonella} nomenclature occurred in the early 1970s when DNA-DNA hybridization experiments demonstrated that all \textit{Salmonella} serotypes share greater than 85 percent of their genetic information and thus form a single species\textsuperscript{13}. In 1986, Le Minor and Popoff proposed to designate this single species as \textit{Salmonella enterica}, which became widely accepted\textsuperscript{14}. Based on DNA relatedness, \textit{Salmonella enterica} species were further divided into seven subgroups corresponding to seven subspecies (I, II, IIIa, IIIb, IV, V, VI), with subspecies V being the most distantly related of the
subgroups. In 1989, Reeves et al. published and classified subspecies V (*Salmonella enterica* subsp *bongori*) as a separate species (*Salmonella bongori*) 15.

According to the current *Salmonella* nomenclature used by CDC, the genus *Salmonella* comprises of two species, *Salmonella enterica* (*S. enterica*) and *Salmonella bongori*. *Salmonella enterica* is further divided into six subspecies, which are referred to by a Roman numeral and a name: I, *S. enterica* subsp. *enterica*; II, *S. enterica* subsp. *salamae*; IIIa, *S. enterica* subsp. *arizonae*; IIIb, *S. enterica* subsp. *diarizonae*; IV, *S. enterica* subsp. *houtenae*; and VI, *S. enterica* subsp. *indica* 12. The majority (59%) of the *Salmonella* serotypes belong to subspecies I (*S. enterica* subsp. *enterica*) 16. Within this subspecies I, the most common O serogroups are A, B, C1, C2, D and E, which cause approximately 99% of *Salmonella* infections in humans and warm-blooded animals 16. Serotypes of other subspecies and *Salmonella bongori* are usually found in cold-blooded animals and the environment, and are rarely associated with disease in humans 17.

### 1.4 *Salmonella*, the host and disease

*Salmonella* are widely distributed in nature and can colonize the gastrointestinal tracts of both cold- and warm-blooded animals, including humans. Even though there are more than 2,500 *Salmonella* serovars, most serotypes are not pathogenic in their natural hosts and the majority of infections in humans and animals are caused by a small number of serotypes within subspecies I 18. The manifestation of disease largely depends on both host susceptibility and the infecting *Salmonella enterica* serotype, and is generally associated with one of the four major syndromes: typhoid fever,
enterocolitis/diarrhea, bacteremia and chronic asymptomatic carriage. For instance, in humans, most *Salmonella* serotypes are associated with acute and self-limiting gastroenteritis, whereas other serotypes (e.g., *S*. Typhi, *S*. Paratyphi and *S*. Sendai) cause enteric fever and a few serotypes (e.g., *S*. Choleraesuis and *S*. Dublin) are more likely to cause bacteremia than diarrhea.

Regarding host adaptability, the majority of *Salmonella* serotypes (e.g., *S*. Typhimurium, *S*. Enteritidis) have a wide host range including animals and humans; however, other serotypes are largely adapted to specific animals and are infrequently found in humans (e.g., *S*. Dublin in cattle, *S*. Gallinarum in poultry, *S*. Abortusequi in horses, *S*. Abortusovis in sheep and *S*. Choleraesuis in pigs). Furthermore, a small number of serotypes such as *S*. Typhi, *S*. Paratyphi A, B, C and *S*. Sendai are fully adapted to humans and higher primates and are unable to cause disease in other hosts. Host adaptation in *Salmonella* can be defined as the ability of a pathogen to circulate and cause disease in a particular host population. Host-adapted *Salmonella* serovars (e.g., *S*. Typhi, *S*. Choleraesuis) tend to be more virulent and cause systemic infections with a higher mortality rates in their hosts compared to broad-host-range serovars (e.g., *S*. Typhimurium or *S*. Enteritidis), which are often associated with non-invasive infections.

### 1.5 Typhoid fever

#### 1.5.1 Clinical features of typhoid fever

Typhoid fever is a life-threatening systemic infection predominantly caused by *Salmonella enterica* subsp *enterica* serovar Typhi (*S*. Typhi) and *Salmonella enterica* subsp *enterica* serovar Typhimurium (*S*. Typhimurium).
subsp enterica serovar Paratyphi A (S. Paratyphi A). The onset of symptoms is marked with prolonged fever, headache, malaise, anorexia, nausea, dry cough and disturbances of bowel function (constipation or diarrhea), which typically occurs after the end of first week. A coated tongue, hepatomegaly and splenomegaly are also common. The fever is low grade during the first week but rises gradually; by the second week, the temperature increases, reaching a plateau of 39 to 40°C. Rose spots on the chest, abdomen and back are reported in 5 to 30 percent of cases. The clinical manifestations and severity of typhoid fever are highly variable due to factors such as duration of illness before appropriate therapy, virulence and antimicrobial resistance of the causative agents, selection of antimicrobial therapy, patient age, previous exposure or vaccination, inoculum size, host factors (HLA type, AIDS or other immune suppression) and antacid consumption 23–32. Depending on the clinical setting and quality of available heathcare, serious complications can occur in 10 to 15 percent of typhoid patients, particularly in those who have been sick for more than two weeks. A wide range of complications are described, of which gastrointestinal bleeding, intestinal perforation, and typhoid encephalopathy are the most common and likely associated with risk of death 33,34.

In the pre-antimicrobial era, case fatality rates were as high as 10 to 30 percent 35,36. Currently, with effective antimicrobial treatment, the average case fatality rate is usually less than 1 percent. However, case fatality rates vary significantly among different regions of the world, ranging from less than 2 percent in Pakistan and Vietnam to 30 to 50 percent in some areas of Indonesia and Papua New Guinea 37–40. Poor outcomes are highest among children less than one year of age and the elderly,
and often result from delayed treatment with effective antimicrobials. Relapse can occur 1 to 3 weeks after the patient recovers from the first episode, with milder symptoms than those experienced during the initial illness. The relapse rate is 5 to 10 percent in untreated cases but still occurs after antimicrobial treatment, especially in those treated with first line drugs (chloramphenicol, trimethoprim-sulfamethoxazole, and ampicillin). Relapse rates in patients treated with newer antimicrobials including fluoroquinolones (1.5%) or broad-spectrum cephalosporins (5%) are normally lower than those treated with first line drugs. Typically, relapses are caused by the same isolate as the original episode and can be distinguished from reinfection by molecular typing. Clinical features of *S. Paratyphi* A infections are commonly considered milder, with fewer complications than infections caused by *S. Typhi*. However, in the largest comparison to date there were no significant differences between clinical presentation, duration or outcome of typhoid caused by these organisms.

### 1.5.2 Pathogenesis of typhoid fever

As *S. Typhi* and *S. Paratyphi A* only cause disease in humans, pathogenesis studies of typhoid fever are hindered by the lack of a suitable animal model. Much of the understanding of pathogenic features of typhoid fever has arisen from *Salmonella enterica* serovar Typhimurium infection in a susceptible murine model that is thought to mimic human typhoid. The infectious dose has been determined in human challenge studies and varies between 1,000 and 1 million organisms, depending on the individual and the settings. After ingestion, the bacteria have to survive the gastric acid before reaching the small intestine. In the small intestine, bacteria can
adhere to the intestinal mucosa and subsequently invade the gut mucosa, probably through specialized epithelial cells overlaying the Peyer’s patches known as the M cells. After penetration, the invading bacteria translocate to the intestinal lymphoid follicles and are drained into mesenteric lymph nodes and the thoracic duct, and eventually into the bloodstream. In this primary bacteremia (within 24 hours of their ingestion), the organisms reach the liver, spleen, bone marrow and other parts of the mononuclear phagocyte systems where they can survive and multiply within the cells of monocytic lineage. After an incubation period of 7-14 days (depending on bacterial load, virulence and host response), bacteria are shed back into the bloodstream causing secondary, sustained bacteremia and marking the onset of the clinical symptoms. The most common sites of secondary infection are the spleen, liver, bone marrow, gallbladder and Peyer’s patches of the terminal ileum. Gallbladder colonisation occurs either directly from blood or by retrograde spread from the bile. Organisms are shed via the bile duct into the small intestine resulting in fecal shedding.

There are some notable differences in the pathogenesis of typhoid fever compared to infections caused by nontyphoidal Salmonella serovars. S. Typhi can translocate across the intestinal mucosa during an early phase of infection without causing any physical cellular damage and therefore without triggering a rapid acute inflammatory response as normally seen in gastrointestinal infections caused by nontyphoidal Salmonella. After infection, the incubation period may not always be followed by clinical symptoms. Furthermore, S. Typhi can survive and multiply in monocytes, which is essential for dissemination and persistence within the host, whereas
nontyphoidal *Salmonella* serovars, like Typhimurium, are effectively cleared by human monocytes. It is also noteworthy that the number of recoverable *S. Typhi* from patients with typhoid fever is low, with a median of 1 cfu/ml of blood and 10 cfu/ml of bone marrow, which has a negative impact on diagnostics.

1.5.3 Typhoid diagnosis

The clinical diagnosis of typhoid fever is challenging in endemic regions as it is difficult to distinguish typhoid from other acute febrile illnesses such as malaria or dengue. Currently, blood culture followed by microbiological identification is still the gold standard diagnostic method, with a sensitivity of approximately 80 percent. Even though blood culture is the most reliable method and can be standardized, it remains costly and requires specialist facilities and personnel and therefore is generally only utilised in major hospitals in developing countries. The sensitivity of blood culture largely relies on the volume of blood taken from typhoid patients due to a low number of circulating organisms. To achieve the highest recovery rates, 10-15 ml of blood from school children and adults and 2-4 ml from toddlers and preschool children are required. In some areas, it can be challenging to obtain such large volumes of blood, especially in children, which undermines diagnosis. Bone marrow culture is more sensitive than blood culture because the number of organisms in bone marrow is comparatively high. The culturing of bone marrow can be valuable for patients who have been treated with antimicrobials regardless of duration of illness. However, this is an invasive procedure and thus not widely accepted, particularly in children, and is rarely performed outside of specialist hospitals. Stool culture has a positive rate of 30 percent in acute typhoid patients and its sensitivity depends on the
amount of feces cultured as well as the duration of the illness. Stool culture can be used to detect typhoid carriers but requires that multiple samples be examined, as shedding is sporadic and may persist at low levels. Instead, a Vi agglutination assay has been used to identify typhoid carriers because these individuals may produce high levels of Vi antibodies over a longer period compared to acutely infected patients.

The Felix-Widal test was developed in 1896 and is the first serological method used to diagnose typhoid fever. This method is based on the measurement of agglutinating antibody levels against the O and H antigens of S. Typhi. The sensitivity and specificity of this test are moderate as S. Typhi share O and H antigens with other Salmonella serovars and also have cross-reacting epitopes with members of Enterobacteriaceae, which can result in a high false positive rate. Additionally, typhoid patients might not develop detectable antibodies or show demonstrable increases in antibody titers. This problem also presents in commercial serological tests such as Tubex and Typhidot, which demonstrated moderate sensitivity and specificity for the diagnosis of typhoid fever in Papua New Guinea, India and Bangladesh. In these studies, the sensitivity and specificity of these commercial kits varied from 51.1-60 % and 58-88.3 %, respectively (Tubex) and 56-70.0 % and 54-88 %, respectively (Typhidot). Despite these limitations, the use of the Felix-Widal test with suitable local cut-off values for positive diagnosis can be helpful in areas where access to alternative, more expensive methods is limited. PCR-based assays have also been developed for typhoid diagnosis. However, they are either unreliable or have poor sensitivity when performed on DNA extracted from blood samples. As a result, these methods are not widely used and also considered impractical in many typhoid
endemic areas. Serological assays based on novel S. Typhi-specific immunogenic proteins are under development which could potentially be converted into cheap and rapid diagnostic kits \(^\text{76,77}\). Alternative approaches based on host-specific biomarkers such as metabolomics are also promising and warrant further investigation \(^\text{78}\).

### 1.5.4 Typhoid treatment

In endemic areas, up to 90 percent of typhoid cases are managed at home with oral antimicrobials, supportive care and regular follow-up for complications \(^\text{31,79}\). For hospitalized patients with severe symptoms, effective antimicrobials, good nursing care, maintenance of appropriate nutrition and hydration, and timely recognition and treatment are crucial to prevent serious complications and deaths. Appropriate antimicrobial therapy is critical to cure typhoid and avoid complications. The selection of antimicrobial drugs depends on a number of important criteria such as availability, efficacy and cost. An ideal drug should have fast time to defervescence and clinical improvement, render blood and stool cultures sterile during and after treatment, prevent relapse, and be available through oral and intravenous routes for both adults and children with low cost and minimal adverse effects \(^\text{80}\).

Prior to the 1990s, chloramphenicol, ampicillin and trimethoprim-sulphamethoxazole were the first drugs of choice for typhoid treatment as these drugs are inexpensive, widely available and rarely associated with side effects. However, the emergence and global dissemination of plasmid-mediated multidrug resistant S. Typhi in the 1990s rendered these drugs ineffective and resulted in the widespread use of fluoroquinolones as the treatment of choice \(^\text{33}\). Data from published clinical trials
demonstrated that fluoroquinolones were the most effective drugs for typhoid treatment. They were rapidly effective even with short courses of treatment (3-7 days), clearing fever and symptoms in 3 to 5 days with a cure rate exceeding 96 percent and very low rates of post-treatment carriage (less than 2 percent) \(^{81-86}\). However, nalidixic acid resistant \(S\). Typhi and \(S\). Paratyphi A exhibiting reduced susceptibility to fluoroquinolones have since emerged and become endemic in many South and Southeast Asian countries, and are associated with increased rates of fluoroquinolone treatment failure \(^{87-90}\). For patients infected with nalidixic acid resistant organisms, treatment with the maximal recommended dose of fluoroquinolones should be given for a minimum of 10 to 14 days and patients should be followed up carefully to assess the shedding of bacteria in their stools. Third generation cephalosporins (ceftriaxone, cefixime, cefotaxime, and cefoperazone) and the macrolide, azithromycin, have been successfully used for typhoid treatment and can be used as effective alternative drugs for treating \(S\). Typhi and \(S\). Paratyphi A with reduced susceptibility to fluoroquinolones \(^{87,91-100}\). In areas where fluoroquinolones are not available or unaffordable, first line drugs (chloramphenicol, ampicillin and trimethoprim-sulphamethoxazole) remain appropriate for the treatment of typhoid if bacteria are still susceptible to these drugs. For severe typhoid, parenteral fluoroquinolones are often the drugs of choice and are given for a minimum of 10 days \(^{33}\). A high dose of intravenous dexamethasone and antimicrobials should be given promptly to adults and children with severe symptoms to reduce mortality \(^{37}\). Relapse should be treated in the same manner as primary infection. The majority of chronic carriers can be cured with a prolonged course of antimicrobials \(^{101-103}\).
1.6 The epidemiology of typhoid fever

1.6.1 Epidemiological features of typhoid fever

*S. Typhi* and *S. Paratyphi A* are human-restricted pathogens and humans are the only known natural hosts and reservoir of infection. The disease is transmitted via the fecal-oral route. The most common mode of typhoid transmission is by ingestion of water or food contaminated with human feces. Direct person-to-person transmission through contact with patients or chronic carriers who are shedding the organisms is uncommon. In endemic areas, water contaminated with human feces and food contaminated with contaminated water or by food handlers who are typhoid carriers are the main sources of infection. The incubation period and attack rate of typhoid fever are affected by the inoculum size and the vehicle of transmission. Waterborne transmission usually requires smaller inoculum, whereas foodborne transmission is associated with a larger inoculum and a higher attack rate over short periods. Bacteria causing typhoid fever can survive for weeks or months in the environment, including in seawater, sewage, pond, stream and lake water. Shellfish grown in polluted water, fruits and vegetables fertilized with sewage, milk and milk products have been documented as potential sources of infection. Often, large typhoid outbreaks occur when the source of drinking water serving large populations is contaminated or the water supply is disrupted. In developed countries, typhoid occurs sporadically and is mainly associated with travellers returning from endemic areas.
1.6.2 Typhoid carriage

Asymptomatic carriage and the shedding of bacteria in feces can occur after patients recover from acute infection. Carriage can be divided into different periods depending on the duration of shedding: convalescent (three weeks to three months), temporary (three to twelve month) and chronic (more than one year) \(^{34}\). In endemic regions, the chronic carriage rate is estimated to be 2-5 percent of the population, most of which are asymptomatic; up to 25 percent of chronic carriers have no clinical history of infection \(^{34,119}\). The relative importance of short-term and convalescent fecal carriers versus chronic carriers in the transmission dynamics of typhoid in endemic regions is largely unknown, probably due to a lack of follow-up and the absence of a robust method for detecting \(S.\) Typhi carriers. Risk factors associated with persistent carriage are not extensively studied, as this population is very challenging to identify prospectively. Previous studies have found that risk of becoming a chronic carrier following acute infection increases with age, and is higher in women and patients with cholelithiasis and cholecystitis \(^{120,121}\). Data from a murine model and electron micrographic observation of gallstones retrieved from \(S.\) Typhi human carriers have suggested that colonization of the gallbladder epithelial cells and biofilm formation on the gallstones are the primary mechanisms by which \(Salmonella\) survives and persists in the gallbladder environment \(^{122–126}\).

Chronic carriage and fecal shedding have been long considered as central dogma determining the transmission and persistence of typhoid fever, as chronic carriers intermittently shed the bacteria into the local environment and may spread the disease in the community; additionally, these people can act as reservoirs for maintaining
specific genotypes \cite{127,128}. However, epidemiological studies have shown that direct transmission between household members primarily occurs through close contact with convalescent carriers rather than chronically infected individuals \cite{129,130}. Furthermore, recent advanced molecular epidemiological investigations in endemic regions such as Vietnam, Indonesia, and Nepal have demonstrated that acute typhoid is generally caused by a wide diversity of genotypes rather than singular local genotypes \cite{131–134}. Therefore, environmental transmission pathway appears to be most important transmission route rather than direct person-to-person contact in endemic regions. Conversely, in areas where public health interventions have been successful in reducing typhoid burden, chronic carriage might play an important role in maintaining the disease in the population, probably via foodborne transmission from asymptomatic chronic carriers who are food preparers or handlers as reported during outbreaks of typhoid fever in the United States \cite{135}. As a result, the detection and treatment of chronic carriers is thought to be essential for eliminating typhoid from a population.

1.6.3 The global burden of typhoid fever

Typhoid fever remains a significant public health problem in many low and middle-income countries where there may be a lack of safe water and adequate sanitation. Global estimates suggested that 21.7 million new cases and 217 000 deaths due to typhoid fever caused by \textit{S. Typhi} occurred in the year 2000 alone, whilst \textit{S. Paratyphi} A also caused 5.4 million cases with an unknown case fatality rate \cite{118}. The majority of illness occurred among infants, children and adolescents, with south-central and south-eastern Asia exhibiting the highest disease burden (>100 cases per 100 000 per
year) 118. More recently, Buckle and colleagues estimated that 13.5-26.9 million episodes of typhoid occurred worldwide in 2010 136. In the context of informing vaccine policy, Vittal-Mogasale and colleagues re-estimated the burden of typhoid fever with a focus on low- and middle-income countries. After adjusting for blood culture sensitivity and water-related risk, they estimated that 11.9 million typhoid fever illnesses and 129 000 deaths occurred in low and middle-income countries during 2010, and also suggested that typhoid burden is higher in Africa than previously thought 137. Between 2010 and 2014, multi-country population-based standardized surveillance for invasive Salmonella infection was conducted in 13 sentinel sites across 10 countries in sub-Saharan Africa and identified a number of sites with remarkably high burden of typhoid fever, especially in children less than 15 years of age (overall adjusted incidence rate >100 per 100 000 person-years of observation) 138. Additionally, data from this study also demonstrated that many rural populations exhibited similar or even higher typhoid incidences than urban populations.

Simultaneously, typhoid fever caused by S. Paratyphi A has emerged at an unprecedented rate in many Asian countries, including Nepal, China, Pakistan and India 139. In some regions, the isolation rate of S. Paratyphi A from typhoid patients is similar or even higher than that of S. Typhi 140,141. The rapid emergence of S. Paratyphi A infections poses a significant public health concern as S. Paratyphi A display differences in epidemiology compared to S. Typhi and, thus, effective control and preventive measures for S. Typhi may not protect against S. Paratyphi A 129,142. Currently, there is no vaccine for S. Paratyphi A and the licensed typhoid vaccines do
not provide protection against *S. Paratyphi* A infections. Additionally, it has been reported that typhoid fever caused by *S. Paratyphi* A increased after typhoid vaccine implementation in Guangxi, China, which suggests that the role of *S. Paratyphi* A in the epidemiology of typhoid fever may impede the progress in typhoid management, and calls into question the impact of vaccines in areas where these two pathogens co-circulate.  

### 1.6.4 Typhoid control and prevention

As contaminated water and food are important vehicles for typhoid transmission, improvements in water, sanitation, and hygienic food preparation represent the ultimate solutions for reducing the burden of disease. Historical surveillance data suggest that rate of typhoid fever in Western Europe and North America substantially declined in parallel with the introduction of water treatment, pasteurization of dairy products and exclusion of human feces from food production. Recent reductions in typhoid fever incidence have also been reported in Latin America and some Asian countries, in parallel with water and sanitation improvements and economic transition. However, such structural improvements remain challenging in low and middle-income countries given the huge economic costs and long timelines that are often required to improve water quality and sanitation. In view of the continued burden of typhoid fever and the increasing antimicrobial resistance of the organisms, in 2008 the World Health Organization (WHO) recommended the programmatic use of licensed typhoid vaccines for endemic and epidemic disease control. There are currently two internationally licensed typhoid vaccines, the parenteral Vi-based polysaccharide vaccine and the live oral Ty21a vaccine, both of which have been proven to be safe.
and efficacious in children aged > 2 years. Many Asian countries such as Vietnam and China have successfully used the Vi polysaccharide vaccine to reduce the burden of typhoid fever, and several other countries also have national policies targeting high-risk groups such as food handlers \(^{147}\). Additionally, a mass vaccination campaign with the Vi polysaccharide vaccine was also conducted in Fiji in cyclone-affected and high-risk areas \(^{148}\). Despite the recommendation of WHO, strong evidence for the disease burden and the low cost, availability, and efficacy of typhoid vaccines, neither of these vaccines has been widely used in endemic low-resource settings. This might be due to a lack of epidemiological data in endemic settings as well as the short duration of vaccine-induced protection; further, none of these vaccines are licensed for children less than 2 years of age \(^{149}\). A number of typhoid conjugate vaccines are under development, including Vi-rEPA (US National Institutes of Health), Typbar-TCV (Bharat Biotech International Ltd) and Vi-CRM197 (Sela Behring Vaccines for Global Health) \(^{150–152}\). Indeed, Typbar-TCV has been prequalified by the WHO and is currently being evaluated in Nepal, Bangladesh and Malawi. These new generation typhoid vaccines are expected to provide higher efficacy and longer duration of protection, and can be given to infants and young children. Efforts are also being made to develop a vaccine for \(S.\ Paratyphi\ A\) as well as a bivalent vaccine that can protect against both \(S.\ Typhi\) and \(S.\ Paratyphi\ A\) infections \(^{147}\).

1.7 Antimicrobial resistance in Typhoidal \(Salmonella\)

1.7.1 Global dissemination of multidrug resistant Typhoidal \(Salmonella\)

Antimicrobial therapy is crucial for the management of typhoid fever and for reducing mortality. In the pre-antibiotic era, reported case-fatality ratios ranged from 10 to 37
percent. In 1948, chloramphenicol was first used to treat typhoid fever and became the standard drug of choice. Fifteen years after the introduction of chloramphenicol, the mortality rate decreased from 12 to 1 percent. This drug was the mainstay for typhoid treatment until mid-1970, when antimicrobial resistant organisms started to emerge and subsequently caused epidemics in Mexico, India, Vietnam and South Korea. The outbreak strains carried a transferable resistant factor (R factor) located on a self-transmissible plasmid of the HI1 incompatibility group (IncHI1), which was responsible for resistance to chloramphenicol and also to streptomycin, sulphonamides and tetracyclines. The spread of chloramphenicol resistant S. Typhi lead to the increased use of the other two first line drugs, amoxicillin and trimethoprim-sulfamethoxazole. However, toward the end of the 1980s and the 1990s, S. Typhi developed multidrug resistance (MDR) to all first line drugs of treatment, including chloramphenicol, trimethoprim-sulfamethoxazole and ampicillin. Outbreaks of multidrug resistant S. Typhi infections were subsequently reported in many countries in South Asia, Southeast Asia, Middle East and Africa. These MDR S. Typhi carried self-transmissible plasmids of the HI1 incompatibility type with typical size from 100 to 120 MDa. Recently, there have also been reports about chromosome-mediated MDR in S. Typhi isolates in Bangladesh, India, Pakistan, Iraq, and Fiji. MDR S. Typhi are currently considered to be endemic in many areas of South, Southeast Asia and Africa. Compared to S. Typhi, S. Paratyphi A have predominantly been found to be more susceptible to antimicrobials; however, MDR S. Paratyphi A isolates have also commonly reported since the 1990s in many settings in Asia. These MDR S. Paratyphi A isolates also harbored IncHI1 plasmids that share a common backbone.
with other IncHI1 plasmids found in S. Typhi and acquired similar mobile elements, conferring a multidrug resistant phenotype\(^\text{180}\).

### 1.7.2 Emergence of fluoroquinolone resistance

The global distribution and dominance of MDR S. Typhi led to the extensive use of fluoroquinolones (ciprofloxacin, ofloxacin) as alternative drugs of choice in the 1990s. Again, the rampant use of ciprofloxacin not only for typhoid but also for other infections resulted in the emergence of S. Typhi isolates that were resistant to nalidixic acid (a quinolone) and reduced susceptibility to ciprofloxacin with MICs ranging from 0.125 to 1 μg/ml, ten times higher than the usual values for fully susceptible strains. Since the early 1990s, S. Typhi isolates with reduced susceptibility to ciprofloxacin have become a major problem in many developing countries in Asia\(^\text{89,164,181–183}\). Further, patients infected with S. Typhi with reduced susceptibility to ciprofloxacin are prone to exhibiting prolonged fever clearance times and increased rates of treatment failure\(^\text{89,90,183}\). In 1997, a massive outbreak caused by such isolates occurred in Tajikistan, infecting 8000 people and causing 150 deaths\(^\text{111}\). The first highly ciprofloxacin resistant S. Typhi isolate was identified in 2005 in India with an MIC of ciprofloxacin > 32 μg/ml\(^\text{184}\). Subsequent reports of increased incidence of infections caused by fully ciprofloxacin resistant S. Typhi later emerged in South Asia and travellers from developed countries who returned from this region\(^\text{185–192}\). It is unclear if the emergence of highly ciprofloxacin resistant S. Typhi isolates was due to clonal expansion of a particular genotype or independent acquisition of resistance in multiple genotypes. Similarly, S. Paratyphi A isolates with reduced susceptibility to ciprofloxacin also emerged and subsequently disseminated throughout South Asia...
over the same period and were associated with infections among travellers returning home from travel to these endemic areas. Alarming, S. Paratyphi A have shown a tendency to be more resistant to ciprofloxacin compared to S. Typhi. Highly ciprofloxacin resistant S. Paratyphi A isolates have also been sporadically reported in India and Japan. Together with the widespread use of ciprofloxacin, there have been reports of a decline in chloramphenicol resistance and MDR S. Typhi isolates in some areas; however, the reintroduction of first line antibiotics is likely to result in the reemergence of MDR S. Typhi.

Fluoroquinolone resistance is chromosome-mediated via point mutations occurring in various chromosomal genes (gyrA, gyrB, parC and parE). These single point mutations alter the enzymes, DNA gyrase and topoisomerase IV, all of which are targets for the quinolones. A single point mutation in the quinolone resistance determining region in the gyrA gene (frequently at codon 83 or codon 87) can result in decreased susceptibility to fluoroquinolones. Double mutation in the gyrA gene (two single point mutations at codon 83 and codon 87) results in reduced susceptibility to ciprofloxacin, whilst an additional mutation in the parC gene (typically at codon 80) is required for high level resistance to ciprofloxacin. Plasmid-mediated quinolone resistance was also reported in the late 1990s and was predominantly associated with qnr proteins that can protect DNA gyrase from ciprofloxacin and also by aminoglycoside-modifying enzyme (aac(6’)-Ib-cr), which acetylates ciprofloxacin. Resistance genes associated with these plasmids, such as qnrS1, qnrB, aac(6’)-Ib-cr, have been sporadically detected in S. Typhi isolates from India and from travellers returning from South Asian countries.
1.7.3 Resistance to third generation cephalosporins and macrolides

In areas where MDR *S. Typhi* and *S. Typhi* with reduced susceptibility to ciprofloxacin are highly prevalent, azithromycin and third generation cephalosporins (ceftriaxone, cefixime, cefotaxime and cefoperazone) tend to be used as alternative drugs of treatment for uncomplicated typhoid fever. These drugs have become important for typhoid treatment as they have been proven effective during clinical trials and the prevalence of resistance to these agents is very low \(^{45,93,213}\). However, since 2008, *S. Typhi* isolates gaining resistance to third generation cephalosporins have been reported in many countries, including India, Bangladesh, Pakistan, the Philippines, Germany, Congo, and Guatemala \(^{214-220}\). There have been a number of extended-spectrum beta-lactamase (ESBL) enzymes identified among those isolates, such as SHV-12, CTX-M types, and AmpC of the ACC-1 type. In 2013, an MDR *S. Paratyphi A* isolate harboring a CTX-M-15 beta-lactamase was also reported from a Japanese traveler returning from India \(^{221}\). Recently, outbreaks of typhoid fever caused by ESBL-producing *S. Typhi* have drawn significant public and scientific attention in India and Pakistan \(^{222}\). For azithromycin, *S. Typhi* isolates frequently have an MIC ranging from 4 to 16 μg/ml and are found to respond well to a short-course of azithromycin treatment \(^{223}\). The high effectiveness of azithromycin in clearing the infection is probably due to azithromycin’s remarkable property of high intracellular concentration (50 times more than in blood) \(^{224}\). However, treatment failures even with low azithromycin MIC of *S. Typhi* and *S. Paratyphi A* as well as high MIC of azithromycin (64 μg/ml) *S. Paratyphi A* have been sporadically reported in India and
travellers returning from Pakistan and India \textsuperscript{225-227}. The mechanism for azithromycin resistance in typhoidal \textit{Salmonella} is not well described.

1.8 The convergent evolution of \textit{Salmonella Typhi} and \textit{Salmonella Paratyphi A}

While most of the \textit{Salmonella enterica} serovars can infect a broad range of host species and are usually associated with self-limiting gastrointestinal infections, Typhi and Paratyphi A are human-restricted serovars and cause life-threatening systemic disease \textsuperscript{228}. There is a high level of similarity in pathological and epidemiological characteristics between these two typhoidal serovars. For instance, Typhi and Paratyphi A display similar pathogenic features during the course of infection, clearly distinguishing them from non-typhoidal \textit{Salmonella enterica} serovars \textsuperscript{153}. This likely results in the indistinguishable clinical features and severity of infections caused by these two serovars as shown in the largest comparison up to date \textsuperscript{197}. Further, both can establish chronic infections in the human gall bladder, which could be a strategy to increase their transmissibility \textsuperscript{130}. Regarding their transmission patterns, Typhi and Paratyphi A generally infect via similar transmission routes, namely contaminated water or food, and both can also be transmitted directly from person to person \textsuperscript{229}.

In the early 1990s, the population structure of \textit{Salmonella enterica} was described based on multilocus enzyme electrophoresis (MLEE), through which Typhi and Paratyphi A were shown to belong to distinct lineages, suggesting that these two \textit{Salmonella} serovars have different evolutionary histories and may have evolved independently to become human-adapted with identical disease phenotype \textsuperscript{15,230}. Multiple complete genomes of Typhi and Paratyphi A were subsequently generated,
followed by genome-wide comparisons between these two serovars as well as to other host-generalized serovars like Typhimurium; these studies provided unprecedented insights into the evolutionary processes of these organisms. Typhi and Paratyphi A are much more closely related to each other at the DNA level than Typhimurium. This is a consequence of extensive homologous recombination within a quarter of their genomes, which results in similar gene content and much lower nucleotide divergence (0.18%) compared to the rest of their genomes (1.2%)\(^{232}\). These recombination events might play a role in the host restriction/pathogenesis of these two serovars, although this remains unproven. Further, the genomes of Typhi and Paratyphi A contain an exceptionally high level of pseudogenes (about 4% of coding sequences in each genome), much higher than the pseudogene contents of Typhimurium (0.9%) and \(E.\ coli\) (0.7%)\(^{231,233–235}\). Pseudogenes are coding sequences that are putatively inactivated by mutations (such as nonsense substitution, frameshifts, gene truncation by deletion or rearrangement), causing the loss of their gene functions. Pseudogenes are not unique to the \(Salmonella\), and are abundant in other host-adapted bacteria such as \(Shigella\ flexneri, Yersinia\ pestis, Rickettsia prowazekii,\) and \(Mycobacterium leprae\)\(^{236–239}\). Even though there are only a limited number of pseudogenes shared between Typhi and Paratyphi A, these genes are known to be involved in adaptation to their comparable niches and the interaction between \(Salmonella\) and host. For examples, loss of gene functions related to intestinal colonization and persistence (typhoidal \(Salmonella\) are less efficient in intestinal colonization and more favorable of systemic sites\(^{55}\)); pathogenicity (disruption in genes encoding for \(Salmonella\)-translocating effector proteins of \(Salmonella\) pathogenicity island I and II); chemotaxis receptor; iron metabolism and
surface-exposed proteins \cite{231,233,240}. Similar loss of protein functions also occurs in genes belonging to the same pathway, which often results in phenotypic convergence. For instance, disruption of fimbrial clusters in both Typhi and Paratyphi A can have important implications for host interaction or disruption in the hin gene but through different mechanisms, resulting in the loss of phase variation in both Typhi and Paratyphi A (both serovars are monophasic for phase 1 flagella whilst most \textit{Salmonella enterica} subspecies 1 have a switching mechanism leading to diphasic flagella) \cite{231}. Gene degradation is likely the most important evolutionary force that alters the pathogenesis and narrows the host range of both Typhi and Paratyphi A \cite{233,234,240}. There are also some specific gene clusters gained by lateral gene transfer that are present in both Typhi and Paratyphi A; however, most of their functions are unknown and unlikely to have a role in phenotypic differences from other \textit{Salmonella enterica} serovars \cite{231}.

1.9 Genomics and molecular epidemiology of typhoid fever

1.9.1 Genomics and its role in molecular epidemiological studies in Asia and Africa

In the pre-genomic era, it was very challenging to study the population structures and genetic characteristics of both \textit{S}. Typhi and \textit{S}. Paratyphi A as they showed very little genetic variation by conventional typing methods such as MLEE, PFGE and MLST \cite{15,241–243}. At the turn of the 21st century, the first complete genome of \textit{S}. Typhi multidrug resistant strain, CT18, was published, with the first \textit{S}. Paratyphi A genome strain, ATCC9150, released a few years later \cite{231,233}. The availability of these genomes provided novel insights into the genetic traits, virulence, resistant determinants and
host adaptation of these pathogens. The S. Typhi CT18 genome comprises of a chromosome (4,809,037 bp), a multidrug resistant plasmid of incompatibility type HI1 (pHCM1-218,150 bp) and a cryptic plasmid (pHCM2-106,516 bp), whilst the S. Paratyphi A ATCC9150 genome is about 200 kb smaller (4,585,299 bp) and does not harbor any plasmid. The difference in chromosome size between these two serovars is primarily reflected in the prophage content and other mobile elements; for instance, S. Typhi carry SPI-7 (a 134 kb region containing the Vi polysaccharide biosynthetic operon, sopE prophage and type IVB pili operon), which is absent from S. Paratyphi A. On the other hand, S. Paratyphi A carries three prophage regions that are not present in S. Typhi, including a sopE prophage (SPA-2-SopE)\(^{231,233}\). Subsequently, additional genomes of S. Typhi (strain Ty2) and S. Paratyphi A (strain AKU12601) became available and allowed extensive genomic comparisons among S. Typhi isolates (CT18 versus Ty2) and S. Paratyphi A isolates (ATCC9150 versus AKU12601). Comparative genomics exhibited an exceptionally high level of pseudogenes and distinctly conserved genomic backbones in both serovars\(^{235,240}\). More than 98 percent of the two Typhi genomes are shared and only 282 single nucleotide polymorphisms were identified\(^{235}\). Similarly, the two Paratyphi A genomes are collinear and highly similar, with only 188 single nucleotide polymorphisms detected\(^{240}\). The clonal natures of both Typhi and Paratyphi A have imposed significant challenges in understanding their bacterial population structures and molecular epidemiological characteristics.

In 2006, Roumagnac and co-workers were the first to describe the global population structure and evolutionary history of S. Typhi by using a mutation discovery method
to characterize SNPs within 200 gene fragments (1.85% of the Typhi genome) from a
globally representative strain collection. The selection of gene fragments was based
on the available complete genomes of strains CT18 and Ty2. Eighty-eight SNPs were
detected and used to resolve Typhi into a rooted, maximally parsimonious
phylogenetic tree defining 59 genetically distinct haplotypes (H1-H59)\textsuperscript{128}. This study
also gave an early warning of the emergence and global spread of H58 S. Typhi,
which was particularly associated with nalidixic acid resistant mutations in the \textit{gyrA}
gene\textsuperscript{128}. Latterly, Holt \textit{et al.} paved the way for using a whole-genome approach to
study the population structure and evolutionary traits of \textit{S. Typhi}\textsuperscript{244}. Their study
identified 1,787 SNPs among 17 Typhi isolates belonging to diverse haplotypes,
providing much better resolution for the Typhi phylogenetic tree, which also
established an important framework for further molecular investigations.

Between 2008 and 2012, molecular epidemiological investigations were performed in
highly endemic settings including Indonesia, Nepal, Kenya, Vietnam, Cambodia, and
India by using SNP-based approaches to characterize the local \textit{S. Typhi} population
structure and investigate the transmission patterns of typhoid fever\textsuperscript{131–134,245–247}. These studies revealed that local bacterial populations are usually highly diverse and
involve multiple co-circulating haplotypes, and showed that environmental
transmission is dominant in urban settings. The findings also highlighted the
increasing frequency of H58 \textit{S. Typhi} in many parts of the world (except for
Indonesia); in many locations, it was observed that the H58 lineage might be
replacing locally existing haplotypes. This particularly successful clone of \textit{S. Typhi} is
predominantly non-susceptible to fluoroquinolones and associated with an MDR
phenotype, suggesting that the heavy use of antimicrobials for typhoid treatment has exerted significant selective pressure for the maintenance and expansion of the H58 S. Typhi population 128.

1.9.2 The origin and global dissemination of H58 Salmonella Typhi

The rapid emergence and subsequent global dissemination of antimicrobial resistant H58 S. Typhi throughout South Asian, Southeast Asian and African countries has become an important public health threat, which urgently requires further investigation on the origin, population structure, transmission and antimicrobial resistance determinants of this particular lineage. Additionally, the global dominance of H58 S. Typhi has also revealed the limitations of SNP-based typing methods, which show a lack of resolution within H58 isolates and have thus been recognised as less effective methods for use in molecular epidemiological studies of typhoid fever. Consequently, the whole genome approach is preferred as an ultimate solution to better understand the origin and evolution of H58 Typhi, as well as to track their global transmission. Over the last decade, the advances of next generation sequencing technology have allowed large-scale bacterial genome sequencing, marking the dawn of the genomic era and revolutionizing the molecular surveillance of infectious diseases. Using an integrative approach of genomics and conventional epidemiology, Wong and colleagues performed whole genome sequencing and phylogeographical analyses of the largest ever collection of S. Typhi (1,832 isolates from 63 countries, including 832 H58 isolates collected from 1992 to 2013) to investigate the evolutionary history, population structure and global transmission of H58 S. Typhi 248. This study confirmed that the global H58 S. Typhi population is highly clonal.
compared to non-H58 populations, an indicator of recent clonal expansion. Their findings also provided strong evidence that H58 lineage emerged in South Asia and expanded dramatically from the early 1990s. There were numerous international transmission events reconstructed from phylogeographical analyses, which demonstrated the propagation of H58 S. Typhi from South Asia to many countries in Southeast Asia, Western Asia, East and South Africa and Fiji, followed by local and region-wide transmission in different endemic areas. In particular, this study underlined the silent ongoing epidemic of MDR H58 S. Typhi across many countries in Eastern and Southern Africa, probably driven by antimicrobial use considering most of the non-H58 isolates in these regions were drug susceptible. The H58 lineage is associated with an MDR phenotype and reduced susceptibility to fluoroquinolones; however, this study showed that the pattern clearly differs between geographical regions, likely due to differences in regional antimicrobial usage. H58 S. Typhi isolates from Southeast Asia were commonly multidrug resistant and possessed gyrA mutations, yielding simultaneous factors contributing to reduced susceptibility to fluoroquinolones, whereas most of the South Asian H58 isolates from recent years were non-MDR and harbored gyrA mutations. In Africa, however, the majority of H58 isolates were multidrug resistant but did not possess gyrA mutations\(^{248}\). Importantly, the MDR gene cassette (Tn2670-like) is largely found in plasmids but can also be inserted into the chromosome at different positions, which can affect the bacterial fitness while still maintain the MDR phenotype\(^{248}\).

The global expansion of H58 S. Typhi population has been reshaping the global population structure of S. Typhi and changing the epidemiology of the disease; for
instance, African countries have recently experienced a surge of MDR typhoid fever which had previously been uncommon. Wong and her colleagues’ work set milestones in understanding the molecular epidemiology of S. Typhi, providing a global population framework and valid methodology for further local investigations into the transmission of antimicrobial resistant lineages of S. Typhi in different endemic areas. Their study also demonstrated the invaluable role of genomics in understanding the S. Typhi population structure and tracking the emergence and global transmission of antimicrobial resistant organisms.

1.10 Aims of this study

The emergence of fluoroquinolone resistant S. Typhi in South Asia, particularly Nepal, has caused significant challenges in typhoid management and has been sporadically reported to be associated with treatment failure. However, the bacterial population structure, antimicrobial resistance determinants and the association between fluoroquinolone resistant phenotype/bacterial genotype with clinical outcomes of typhoid fever have not yet been well characterized. Concurrently, S. Paratyphi A infections have also increased rapidly in Nepal and little is known about the bacterial population dynamics, antimicrobial resistance and epidemiological features in this settings. Previous epidemiological surveillance studies of typhoid fever have primarily focused on urban populations, where high incidences of typhoid fever have been reported. There has been a large gap in knowledge about the burden of typhoid fever as well as risk factors, transmission patterns and antimicrobial resistance in rural settings such as described in this thesis for Siem Reap, Cambodia. One of the most long-lasting and elusive research questions related to typhoid fever is
about the role of chronic carriage in disease transmission and bacterial persistence. Genomic characterization and comparison between chronic and acute S. Typhi isolates is necessary to provide novel understanding into the relative importance of chronic carriage.

In order to fill in these current gaps in typhoid research, the aims of my study are as follows:

- To use a combined genomic and clinical approach to characterize the population structure and dynamics of S. Typhi in Nepal, identify fluoroquinolone resistant determinants and understand the clinical features of typhoid fever.
- To characterize the spatiotemporal dynamics, antimicrobial resistance patterns/determinants and phylogenetics of S. Paratyphi A in Kathmandu, Nepal.
- Combine genomics and conventional epidemiological approaches to understand the molecular and spatial epidemiology of typhoid fever in rural areas in Siem Reap, Cambodia.
- Describe the genomic traits of carrier S. Typhi isolates and the genetic relationship between carrier and acute S. Typhi to provide insights into the role of typhoid carriage in the disease transmission
Chapter 2

Methods

2.1 Study sites and settings

2.1.1 Patan Hospital in Kathmandu, Nepal

The majority of bacterial isolates and associated metadata in my study originated from typhoid studies conducted in Patan Hospital in Kathmandu, Nepal. Kathmandu, the capital city of Nepal, has a population of 1.5 million and is situated at an altitude of 1300m. The climate varies from cold winter months (December to February) to spring (March to May), the hot monsoon season (June to August) and autumn (September to November). Traditionally, the monsoon is characterized by a heavy burden of enteric infections. Patan Hospital, a 318-bed government hospital located in the Lalitpur Sub-Metropolitan City (LSMC) in the Kathmandu Valley providing both emergency and elective inpatient services. Each year, Patan Hospital has 300,000 outpatients and 16,000 inpatients, and the bed occupancy is approximately 85%. Ninety per cent of patients are from the immediate Kathmandu Valley and Lalitpur area. Enteric fever is common at the outpatient clinic at Patan Hospital, which has approximately 200,000 outpatient visits annually. The population of LSMC is generally poor, with most living in overcrowded conditions and obtaining their water from stone spouts or sunken wells. Antimicrobials are available without prescription in the community in a variety of public and private outlets and there are numerous private physician clinics where patients may seek advice and clinical diagnosis for febrile disease. There has been no widespread implementation of a typhoid vaccine in this area, yet a generic typhoid Vi vaccine is available for purchase in some health care settings.
2.1.2 Angkor Hospital for Children in Siem Reap, Cambodia

Angkor Hospital for Children (AHC) is one of two pediatric hospitals in Siem Reap City and has approximately 125,000 attendees and 4,000 admissions per year. The patients attending AHC are <16 years of age and come from a wide geographical radius and attend the hospital for various conditions. The majority of patients reside in the province of Siem Reap, which is located in northwest Cambodia and is bordered in the south by the Tonle Sap Lake, the largest freshwater lake in Southeast Asia. According to available census data, the province had a population of 896,443 people living in an area of 10,299 km² in 2008; the province is subdivided administratively into 12 districts, 100 communes (which are within districts) and 907 villages.

Cambodia has a tropical climate with a dry and wet season each year. During the wet season (April–October) the area of the Tonle Sap Lake can expand dramatically, increasing from 3,500 km² up to approximately 14,500 km², with the depth increasing from 0.5m up to 6-9m.

2.2 Bacterial identification and antimicrobial susceptibility testing

2.2.1 Blood culture

Routine diagnosis of typhoid fever was performed by blood culture. Blood (1-3 ml for children; 5-10 ml for adults) was taken from all patients for bacterial culture on enrolment. Adult blood samples were inoculated into media containing tryptone soya broth and sodium polyanethol sulphonate, up to a total volume of 50 ml. Bactec Peds Plus culture bottles (Becton Dickinson, New Jersey, USA) were used for paediatric blood samples. The bottles were incubated at 37°C and examined daily for bacterial growth up to 7 days. Positive bottles were subcultured onto blood, chocolate and
MacConkey agar and presumptive Salmonella colonies were identified using standard biochemical tests and serotype-specific antisera (Murex Biotech, Dartford, England).

2.2.2 Bile and stool culture

Bile and stool were collected for culture from all patients undergoing cholecystectomy or laparotomy surgery between June 2007 and October 2010. Bile was inoculated into equal volumes of Selenite F broth and Peptone broth and incubated at 37°C overnight. Broths were subcultured onto MacConkey agar and Xylene Lysine Deoxycholate (XLD) agar. After overnight incubation at 37°C the plates were examined for the growth of Gram-negative bacteria and colonies were identified by standard microbiological methods and identified by API20E manufactured by bioMerieux, Inc. S. Typhi and S. Paratyphi A isolates were confirmed by slide agglutination by specific antisera (Murex Biotech, Biotech, England).

2.2.3 Antimicrobial susceptibility testing

Antimicrobial susceptibility testing was performed by the modified Bauer-Kirby disc diffusion method with zone size interpretation based on Clinical and Laboratory Standards Institute guidelines (CLSI, 2012). The control strains used for all susceptibility tests were E. coli ATCC 25922, Pseudomonas aeruginosa ATCC 27853 and Staphylococcus aureus ATCC 29213. Etests were used to determine the minimum inhibitory concentrations (MICs), following the manufacturer's recommendations (bioMérieux, France). Ciprofloxacin MICs were used to categorize S. Typhi and S.
2.3 Vi agglutination assay

The two carriage isolates lacking ViaB operon (GB003 and GB428) were grown in LB agar plates with different NaCl concentrations (1mM, 85mM and 170mM). Agglutination tests were performed on glass microscope slides by mixing 10μl of suspensions of single colonies with 50μl of antisera against Vi (Murex Biotech, Biotech, England). Agglutination was read after rocking the slide for 1 minute. Two other carriage isolates (GB125 and GB169) were used as controls for this experiment.

2.4 Data sources and bacterial isolates

2.4.1 Data sources and bacterial isolates in chapter 3

The S. Typhi isolates and corresponding clinical data for chapter 3 originated from an open-label, randomized, controlled, superiority trial conducted at Patan Hospital and the Civil Services Hospital in the Kathmandu valley, Nepal between 2011 and 2014. The trial was registered at www.clinicaltrials.gov (ISRCTN63006567). Briefly, patients were randomly assigned to seven days of treatment with either oral gatifloxacin (400 mg tablets, Square Pharmaceuticals Limited, Bangladesh) at a dose of 10 mg/kg once daily or intravenous ceftriaxone (Powercef, 1000mg injection vial, Wock-hardt Ltd, India), injected over 10 min at a dose of 60 mg/kg up to a maximum of two grams (aged 2 to 13 years) or two grams (≥ 14 years) once daily. The trained community medical auxiliaries (CMAs) visited each patient assigned to treatment twice per day for at least 10 days or until the patient was asymptomatic. The CMAs
gave the drugs, and recorded drug doses, administration times, oral temperatures, symptoms, and potential adverse effects in a standard case-record form. Complete blood count, serum creatinine, liver-function parameters and serum glucose at enrolment and on day 8 of treatment were measured. Blood from all patients was subjected to bacterial culture at enrolment and on day 8 after randomization if S. Typhi or S. Paratyphi were isolated at enrolment, or if their symptoms suggested a clinical relapse.

The primary endpoint was a composite of treatment failure, defined as the occurrence of at least one of the following events: fever clearance time (FCT) (time from the first dose of a study drug until the temperature dropped to 37.5°C and remained there for at least two days) more than seven days post-treatment initiation; requirement for rescue treatment as judged by the treating physician; blood culture positivity for S. Typhi or S. Paratyphi on day eight of treatment (microbiological failure); culture-confirmed or syndromic enteric fever relapse within 28 days of initiation of treatment; and the development of any enteric fever-related complication (e.g. clinically significant bleeding, fall in the Glasgow Coma Score, perforation of the gastrointestinal tract and hospital admission) within 28 days after the initiation of treatment. Time to treatment failure was defined as the time from the first dose of treatment until the date of the earliest failure event. FCTs were calculated electronically using twice-daily recorded temperatures and treated as interval-censored outcomes. Patients without fever clearance or relapse, respectively, were censored at the time of their last follow-up visit.
There was a total of 78 *S. Typhi* isolates identified from this trial. Additionally, 58 *S. Typhi* isolates from previous studies conducted between 2008 and 2013 in Patan Hospital and genome sequences of 19 *S. Typhi* isolates from a recent international study of the H58 lineage were also included for phylogenetic analysis. Details about these *S. Typhi* isolates and their corresponding sequencing metadata are shown in appendix A.

### 2.4.2 Data sources and bacterial isolates in chapter 4

223 *S. Paratyphi A* isolates were collected from several studies conducted at Patan Hospital, Nepal between 2005 and 2014. Of which, 206 *S. Paratyphi A* isolates were collected from acutely infected people (defined as acute *S. Paratyphi A* isolate) from four randomized control trials: gatifloxacin vs cefixime (2005) \(^{257}\), gatifloxacin versus chloramphenicol (2006-2008) \(^{258}\), gatifloxacin versus ofloxacin (2008-2011) \(^{259}\), gatifloxacin versus ceftriaxone (2011-2014) \(^{260}\) and a matched case-control study (2011) \(^{129}\). 17 *S. Paratyphi A* isolates were recovered from gallbladder bile of people who underwent cholecystectomy or laparotomy surgery for symptomatic cholelithiasis from 2007 to 2010 \(^{130}\). These 17 *S. Paratyphi A* isolates were defined as carrier *S. Paratyphi A* isolates given the fact that they were isolated from the gallbladder of asymptomatic carriers and the duration of carriage was unknown. Additionally, epidemiological and clinical information was respectively collected from the above studies where available. Major reported variables included age, sex, main water sources, water treatment methods, clinical symptoms at presentation, antimicrobial resistance information and individual GPS location. Details about the *S. Paratyphi A* isolates and their associated metadata are shown in appendix B.
2.4.3 Data sources and bacterial isolates in chapter 5

The *S.* Typhi isolates and associated metadata described in chapter 5 originated from a retrospective study of invasive salmonellosis conducted at Angkor Hospital for Children (AHC) in Siem Reap City in Cambodia between January 2007 and December 2014 (appendix C). Case and control populations were identified from the electronic hospital and laboratory information system of AHC. For this investigation, the case population was defined as the population of hospital inpatients from whom *S.* Typhi was isolated from a blood culture. The control population was defined as the patient population admitted to AHC who did not have typhoid fever based on the recorded discharge diagnosis (International Classification of Disease (ICD)-10 code). Patients with a discharge diagnosis of typhoid fever but without blood culture confirmation (*n*=410) were not included in the risk factor analysis. Additionally, for the mapping and population risk factor analyses, cases that lived outside of a 100km radius from AHC were excluded. Data on age, sex, home location (commune level), admission and discharge dates for cases and controls were extracted from the electronic hospital information system. If a case or control was readmitted to the hospital with the same discharge diagnosis within a seven-day period, only the initial admission was included in the analysis.

Commune-level census data were obtained from the Cambodian National Report on General Population Census of 2008. The extracted information included details regarding demographic indicators, age structure, literacy and education, housing and household characteristics, and access to toilet facilities and drinking water. Based on this report, a commune was classified as urban if the population density exceeded
200/km², less than half of men were employed in agriculture and the total population exceeded 2,000. Monthly average precipitation was collected from Siem Reap Weather Station and MRCS (Mekong River Commission Secretariat) 261. Shuttle Radar Topography Mission (SRTM) elevation data were obtained from the CGIAR Consortium for Spatial Information (CGIAR-CSI) 262. Shapefile layers containing 2008 commune-level population census data were accessed from Open Development Cambodia, an open-access data website providing data on Cambodia and its economic and social development (http://www.opendevelopmentcambodia.net).

2.4.4 Data sources and bacterial isolates in chapter 6

Between June 2007 and October 2010, a study of microbiology and epidemiology of invasive Salmonella carriage was conducted at Patan Hospital in Kathmandu 130. Bile samples from 1377 individuals who underwent cholecystectomy for acute or chronic cholecystitis were subjected to microbiological examination. 24 S. Typhi isolates identified from the positive bile samples and their associated metadata were retrospectively collected. These S. Typhi isolates were defined as carrier S. Typhi isolates as they were isolated from gallbladder of asymptomatic carriers and the duration of carriage was unknown. Additionally, 96 S. Typhi isolates recovered from acute typhoid patients (defined as acute S. Typhi isolates) during a randomized controlled trial conducted in Patan Hospital between July 2008 and August 2011 were also included 259. Details about these S. Typhi are shown in appendix D.

2.5 Whole genome sequencing

2.5.1 DNA extraction
Bacterial isolates grown overnight on Nutrient Agar plates were subjected to DNA extraction using the Wizard Genomic DNA Purification Kit (Promega, Wisconsin, USA). This DNA extraction kit includes nuclei lysis solution, RNase solution, protein precipitation solution and DNA rehydration solution.

There are four main steps:

- Cell lysis
- RNA digestion
- Removal of proteins by salt precipitation (DNA left in solution)
- Isopropanol precipitation to concentrate and desalt the DNA

Procedure:

- Scrape the surface the overnight culture plate and add into 1 ml of sterile saline to the appropriately labelled 1.5ml microcentrifuge tube
- Centrifuge the tubes at 13,000 rpm for 2 minutes to pellet the cells
- Remove and discard the supernatants from the tubes
- Add 600 μl nuclei lysis solution to the tubes and gently pipette until the cells are fully re-suspended
- Incubate at 80°C for 5 min to lyse the cells; then cool to room temperature
- Add 3 μl of RNase solution to the cell lysates. Invert the tubes 2-5 times to mix
- Incubate at 37°C for 5 min. Cool the samples to room temperature
- Add 200 μl of protein precipitation solution to each lysate. Vortex vigorously for 20 seconds to mix the protein precipitation solution with the cell lysates
- Incubate the samples on ice for 5 min
− Centrifuge at 13,000 rpm for 3 min
− Pipette 600 μl isopropanol into the newly labeled tubes
− Transfer the supernatants containing the DNA to the appropriately labeled microcentrifuge tubes containing 600 μl of isopropanol
− Gently mix by inversion until the thread like strands of DNA form a visible mass
− Centrifuge at 13 000 rpm for 2 min
− Carefully pour off the supernatants and drain the tubes on clean absorbent paper
− Add 600 μl of room temperature 70% ethanol to each tube and gently invert the tubes several times to wash the DNA pellet
− Centrifuge at 13,000 rpm for 2 minutes and carefully aspirate the ethanol
− Drain the tubes on clean absorbent paper and allow the pellet to air-dry (will take approximately 10-15 minutes)
− Add 200 μl of DNA rehydration solution to the tubes and incubate at 65°C for 1 hour. Periodically mix the solution gently. Alternatively the DNA can rehydrate by incubating overnight at room temperature or at 4°C
− Store the DNA at -20°C until required

2.5.2 DNA quantification
Qubit dsDNA high sensitivity kit (Thermo Fisher Scientific, Waltham, MA, USA) was used to accurately quantify the DNA concentration. The kit includes concentrated assay reagent, dilution buffer, and prediluted DNA standards.
Procedure:

- Prepare thin-wall, clear, 0.5 ml PCR tubes
- Prepare the Qubit working solution by diluting Qubit HS reagent 1:200 with Qubit HS buffer in a plastic tube

- For 2 standards (positive and negative):
  - 10 μl each standard
  - 190 μl working solution

- For DNA samples
  - 2 μl each DNA sample
  - 198 μl working solution

- Incubate all tubes at room temperature for 2 minutes
- Proceed to “reading standards and samples using Qubit 3.0 fluorometer

2.5.3 DNA library preparation

DNA samples were subjected to library construction using Nextera DNA sample prep kit (Illumina, USA). Nextera DNA library prep protocol can fragment and add adapter sequences onto template DNA with a single tube Nextera reaction to generate multiplexed pair-end sequencing libraries.

2.5.3.1 Tagment genomic DNA

Nextera library preparation procedure combines DNA fragmentation, end-polishing, and adaptor-ligation steps into one, termed tagmentation.

Procedure:
• Add 20 μl of genomic DNA at 2.5 ng/μl (50 ng total) to each 0.2 ml tube
• Add 25 μl of TD Buffer to the tubes containing genomic DNA
• Add 5 μl of TDE1 to the tubes containing genomic DNA and TD Buffer
• Pipette up and down 10 times (or more) to mix thoroughly. Quick spin
• Place on the thermal cycler with a heated lid and run the program
  55°C for 5 minutes
  Hold at 10°C

2.5.3.2 Clean up Tagmented DNA

The tagmented DNA is purified from the Nextera transposome as the transposome can bind tightly to DNA ends and interfere with downstream reactions.

Procedure:

• Transfer 50 μl from each tube to a new eppendorf, add 300 μl DNA binding buffer (Zymo), mix and transfer the mixture to Zymo-Spin column in a collection tube
  • Centrifuge 13,200 rpm for 1 minute
  • Discard the flow-through
  • Wash twice as follow
    • Add 500 μl wash buffer (Zymo) to each Zymo-Spin column
    • Centrifuge 13,200 rpm for 30 seconds
    • Discard the flow-through
  • Centrifuge at 13,200 rpm for 1 minute to make sure there is no residual wash buffer
• Transfer the Zymo-spin column to a new eppendorf
- Add 25 μl of RSB (Resuspension Buffer) directly to the Zymo-Spin column
- Incubate for 2 minutes at room temperature
- Centrifuge at 13,200 rpm for 2 minutes

2.5.3.3 Amplify Tagmented DNA

Purified tagmented DNA is amplified using a 5-cycle PCR program. This step adds index 1 (i7) and index 2 (i5) and adapters (P5, P7) required for cluster generation and sequencing. Arrange index 1 (i7) and index 2 (i5) in an Index Plate as follow:

Columns 1-12: Index 1 (i7) adapter (orange caps)
Rows A-H: Index 2 (i5) adapter (white caps)

Procedure:

- Prepare a new 0.2ml tubes
- Add 5 μl each index 1 adapter down each column
- Add 5 μl each index 2 adapter across each row
- Add 15 μl NPM (Nextera PCR Master Mix)
- Add 5 μl PPC (PCR Primer Cocktail)
- Add 20 μl each purified tagmented DNA (from step 3)
- Mix by pipetting up and down or vortex briefly. Quick spin
- Place on a thermal cycle with a heated lid and run the PCR program:
  - 72°C for 3 minutes
  - 98°C for 30 seconds
  - 5 cycles of:
    - 98°C for 10 seconds
    - 63°C for 30 seconds
72°C for 3 minutes

Hold at 10°C

2.5.3.4 Clean up libraries

This step uses AMPure XP beads to purify the library DNA and perform size selection to remove short library fragments.

- Prepare fresh 80% ethanol from absolute ethanol.
- Bring the AMPure XP beads to room temperature.

Procedure:

- Vortex AMPure beads rigorously to ensure beads are homogenous.
- Add 30 μl AMPure beads to each tube containing amplified library
- Mix by pipetting up and down
- Incubate at room temperature for 5 minutes
- Place on a magnetic stand and wait until the liquid is clear. Keep the tubes on the magnetic stand for the following steps.
- Discard supernatant
- Wash twice as follows:
  - Add 200 μl fresh 80% Ethanol. Avoid disturbing the beads
  - Incubate on the magnetic stand until the supernatant is clear
  - Remove and discard the supernatant from each tube
- Remove residual Ethanol using white tip with pipette P20
- Air-dry beads for 10 minutes
- Remove the tubes from the magnetic stand
• Add 32.5 μl RSB (Resuspension Buffer) and pipette up and down to mix
• Incubate at room temperature for 5 minutes
• Place the tubes on the magnetic stand and wait until the supernatant is clear
• Transfer 30 μl supernatant to a new tube, avoid bead carryover

2.5.3.5 Library quantification

Library quantification was performed using KAPA SYBR® FAST qPCR Kit (Kapa Biosystems, Wilmington, USA). KAPA Library Quantification Kits enable accurate and reproducible qPCR-based quantification of libraries prepared for Illumina sequencing. The kit contains six DNA standards (a 10-fold dilution series of a linear, 452 bp template ranging from 0.0002 pM up to 20 pM), Library Quantification Primer Premix (10X) and KAPA SYBR® FAST qPCR Master Mix (2X).

Procedure:

• Make 1:1000 dilution of dsDNA library using Molecular Grade (MG) Water (or 10mM Tris-HCl, pH 8 + 0.05% Tween 20)
• Prepare qPCR mix as follows:
KAPA SYBR FAST qPCR Master Mix containing Primer Premix  6 μl
Diluted library DNA or DNA Standard (1-6) 4 μl
• Run qPCR protocol
  Initial activation  95°C  5 min
  35 cycles:
  Denaturation  95°C  30 sec
  Annealing/extension/data acquisition  60°C  45 sec
Melting curve analysis (optional)

- Analyze the data and calculate the library concentration
  - Confirm 90-100% reaction efficiency for standards
  - Calculate the concentration of diluted library as determined by qPCR in relation to concentrations of DNA Standards 1-6
    - Perform a size adjusted calculation for the difference in the size between average fragment length of library and the DNA Standard (452 bp)
    - Calculate the undiluted library by taking into account the dilution factor and volume used per reaction (4 μl)

2.5.3.6 Normalize and Pool libraries

DNA Libraries were normalized to 4 nM and diluted to 20 pM before loading into the flow cell for Illumina sequencing on Miseq System.

Procedure:

Normalize and pooling

- Dilute each library to 4 nM using MG water (or 10mM Tris-HCl, pH 8 + 0.05% Tween 20)
- Transfer 5 μl each normalized library to a new eppendorf
- Vortex and spin down.

These are pooled libraries, next step is to denature with 0.2 N NaOH and dilute to 20 pM

Denaturing and Diluting

- Prepare 0.2 N NaOH (to denature the pooled libraries) as follows:
• MG water: 800 μl
• 1.0 N NaOH: 200 μl
• Vortex and spin down (use only within 12 hours)
  • Thaw HT1 (hybridization buffer) and store on ice until you are ready to dilute denatured library
  • Combine the following volumes in a 0.2 ml tube
    4 nM pooled libraries (5 μl)
    0.2 N NaOH (5 μl)
  • Vortex briefly and spin down
  • Incubate at room temperature for 5 minutes
  • Denature at 95°C for 2 minutes (optional)
  • Add 990 μl prechilled HT1 to the tube containing the denatured library. Keep on ice. This step results in 1 ml of a 20 pM denatured library ready for loading into the flow cell.

2.5.4 Perform a run on Miseq

Prepared pooled libraries were sequenced using Miseq Reagent v2 (300/500 cycles). The sequencing kit contains: Reagent Cartridge, buffer HT1, PR2 Bottle and Miseq Flow Cell.

Procedure:
  • Clean the flow cell thoroughly using ELGA water prior to use
  • Remove the reagent cartridge from -20°C, thaw in a water bath at room temperature (take about an hour)
• When thawed, remove the cartridge from water bath, gently tap it to the bench and dry the base.
• Invert the cartridge ten times to check if all the reagents are thawed
• Gently tap the cartridge on the bench to reduce air bubbles
• Load 600 μl of your prepared library onto the reagent cartridge in the reservoir labeled “Load Samples” (position 17)
• Using the illumina Experiment Manager to set up sample sheet
• Using the Miseq Control Software (MCS) interface, follow the run setup steps to load the flow cell and reagents and then start the run

2.6 Single Nucleotide Polymorphism (SNP) detection and analysis

2.6.1 SNP detection and analysis for Salmonella Typhi genomes

2.6.1.1 SNP detection and annotation

All raw Illumina reads were mapped to the reference sequence of S. Typhi strain CT18 (Accession no: AL513382), plasmid pHCM1 (AL513383) and pHCM2 (AL513384) using SMALT (version 0.7.4). Candidate single nucleotide polymorphisms (SNPs) were called against the reference sequence using SAMtools and filtered with a minimal Phred quality of 30 and a quality cut-off of 0.75. The allele at each locus in each isolate was determined by reference to the consensus base in that genome, using samtools mpileup and removing low confidence alleles with consensus base quality ≤20, read depth ≤5 or a heterozygous base call. SNPs called in phage regions, repetitive sequences or recombinant regions identified previously were excluded. Gubbins was also used to identify recombinant regions from the whole
genome alignment produced by SNP-calling isolates \(^{264}\) and SNPs detected within these regions were also removed. SNPs were annotated using the parseSNPTable.py script in the RedDog pipeline (https://github.com/katholt/RedDog). Strains belonging to haplotype H58 were defined by the SNP \(glpA\)-C1047T (position 2,348,902 in \(S.\) Typhi CT18, BiP33, as previously described \(^{128,244}\)).

2.6.1.2 \textit{Salmonella} Typhi genotyping

From the detected SNPs in \textit{Salmonella} Typhi genomes, a subset of 68 SNPs was used to assign \textit{Salmonella} Typhi isolates to previously defined lineages according to an extended \(S.\) Typhi genotyping framework (Table 2.1). This SNP-based genotyping framework included four primary clusters, 16 clades and 49 subclades, developed from whole genome data of more than 1800 globally representative \(S.\) Typhi isolates \(^{265}\). Under this genotype nomenclature, the globally disseminated haplotype H58 belongs to subclade 4.3.1.
Table 2.1 Canonical SNPs for genotyping *Salmonella Typhi* into Clades and Subclades

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>In</th>
<th>Out</th>
<th>CT18 position</th>
<th>Gene</th>
<th>Product</th>
<th>Nt</th>
<th>Codon</th>
<th>ns</th>
<th>s</th>
<th>%Div</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>11</td>
<td>T</td>
<td>C</td>
<td>655112</td>
<td>STY0655</td>
<td>alkyl hydroperoxide reductase c22 protein</td>
<td>547</td>
<td>183</td>
<td>0</td>
<td>2</td>
<td>0.35</td>
</tr>
<tr>
<td>0.0.1</td>
<td>2</td>
<td>A</td>
<td>G</td>
<td>773487</td>
<td>STY0776</td>
<td>succinate dehydrogenase hydrophobic membrane anchor protein</td>
<td>333</td>
<td>111</td>
<td>0</td>
<td>1</td>
<td>0.29</td>
</tr>
<tr>
<td>0.0.2</td>
<td>2</td>
<td>C</td>
<td>T</td>
<td>1804415</td>
<td>STY1910</td>
<td>putative ATP/GTP-binding protein</td>
<td>996</td>
<td>332</td>
<td>0</td>
<td>3</td>
<td>0.27</td>
</tr>
<tr>
<td>0.0.3</td>
<td>2</td>
<td>A</td>
<td>G</td>
<td>1840727</td>
<td>STY1951</td>
<td>putative ATP-dependent helicase</td>
<td>1302</td>
<td>434</td>
<td>1</td>
<td>6</td>
<td>0.37</td>
</tr>
<tr>
<td>0.1.1</td>
<td>2</td>
<td>A</td>
<td>G</td>
<td>364068</td>
<td>STY3775</td>
<td>primosomal protein replication factor</td>
<td>825</td>
<td>275</td>
<td>0</td>
<td>10</td>
<td>0.46</td>
</tr>
<tr>
<td>0.1.2</td>
<td>1</td>
<td>T</td>
<td>C</td>
<td>270120</td>
<td>STY0255</td>
<td>acetyl-coenzyme A carboxylase carboxyl transferase subunit alpha</td>
<td>196</td>
<td>66</td>
<td>0</td>
<td>1</td>
<td>0.10</td>
</tr>
<tr>
<td>0.1.3</td>
<td>1</td>
<td>A</td>
<td>G</td>
<td>102135</td>
<td>STY0104</td>
<td>CorD protein</td>
<td>330</td>
<td>110</td>
<td>0</td>
<td>1</td>
<td>0.26</td>
</tr>
<tr>
<td>1</td>
<td>7</td>
<td>C</td>
<td>T</td>
<td>316489</td>
<td>STY0304</td>
<td>conserved hypothetical protein</td>
<td>117</td>
<td>39</td>
<td>14</td>
<td>6</td>
<td>1.49</td>
</tr>
<tr>
<td>1.1</td>
<td>6</td>
<td>T</td>
<td>C</td>
<td>4105384</td>
<td>STY4239</td>
<td>conserved hypothetical protein</td>
<td>267</td>
<td>89</td>
<td>0</td>
<td>1</td>
<td>0.17</td>
</tr>
<tr>
<td>1.1.1</td>
<td>1</td>
<td>A</td>
<td>C</td>
<td>555826</td>
<td>STY0548</td>
<td>putative ABC transporter ATP-binding protein</td>
<td>117</td>
<td>39</td>
<td>0</td>
<td>2</td>
<td>0.30</td>
</tr>
<tr>
<td>1.1.2</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>2360997</td>
<td>STY2526</td>
<td>Ais protein</td>
<td>273</td>
<td>91</td>
<td>2</td>
<td>5</td>
<td>1.15</td>
</tr>
<tr>
<td>1.1.3</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>4664137</td>
<td>STY4803</td>
<td>ornithine carbamoyltransferase</td>
<td>213</td>
<td>71</td>
<td>11</td>
<td>13</td>
<td>2.39</td>
</tr>
<tr>
<td>1.1.4</td>
<td>1</td>
<td>A</td>
<td>G</td>
<td>2166082</td>
<td>STY2335</td>
<td>uridine kinase</td>
<td>81</td>
<td>27</td>
<td>0</td>
<td>1</td>
<td>0.16</td>
</tr>
<tr>
<td>1.2</td>
<td>1</td>
<td>T</td>
<td>C</td>
<td>30192</td>
<td>STY0031</td>
<td>fimbral subunit</td>
<td>195</td>
<td>65</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>1.2.1</td>
<td>1</td>
<td>T</td>
<td>C</td>
<td>4288272</td>
<td>STY4419</td>
<td>putative lipoprotein</td>
<td>588</td>
<td>196</td>
<td>0</td>
<td>2</td>
<td>0.31</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>A</td>
<td>G</td>
<td>2737027</td>
<td>STY2867</td>
<td>hypothetical protein</td>
<td>171</td>
<td>57</td>
<td>1</td>
<td>3</td>
<td>0.72</td>
</tr>
<tr>
<td>2.0.1</td>
<td>9</td>
<td>G</td>
<td>T</td>
<td>1215983</td>
<td>STY1260</td>
<td>putative ROK-family protein</td>
<td>651</td>
<td>217</td>
<td>2</td>
<td>2</td>
<td>0.44</td>
</tr>
<tr>
<td>2.0.2</td>
<td>4</td>
<td>A</td>
<td>G</td>
<td>4132085</td>
<td>STY4268</td>
<td>glucose utilization operon repressor</td>
<td>237</td>
<td>79</td>
<td>3</td>
<td>3</td>
<td>0.60</td>
</tr>
<tr>
<td>2.1</td>
<td>1</td>
<td>A</td>
<td>G</td>
<td>146673</td>
<td>STY0144</td>
<td>GDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alananyl ligase</td>
<td>297</td>
<td>99</td>
<td>0</td>
<td>4</td>
<td>0.29</td>
</tr>
<tr>
<td>2.1.1</td>
<td>3</td>
<td>A</td>
<td>G</td>
<td>2517324</td>
<td>STY2682</td>
<td>putative oxidoreductase</td>
<td>283</td>
<td>95</td>
<td>0</td>
<td>3</td>
<td>0.38</td>
</tr>
<tr>
<td>2.1.2</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>3920009</td>
<td>STY4059</td>
<td>conserved hypothetical protein</td>
<td>768</td>
<td>256</td>
<td>2</td>
<td>5</td>
<td>0.80</td>
</tr>
<tr>
<td>2.1.3</td>
<td>3</td>
<td>A</td>
<td>G</td>
<td>3276735</td>
<td>STY3432</td>
<td>D-galactarate dehydratase</td>
<td>909</td>
<td>303</td>
<td>0</td>
<td>1</td>
<td>0.06</td>
</tr>
<tr>
<td>2.1.4</td>
<td>1</td>
<td>T</td>
<td>C</td>
<td>1173984</td>
<td>STY1220</td>
<td>flagellar P-ring protein precursor</td>
<td>360</td>
<td>120</td>
<td>2</td>
<td>4</td>
<td>0.54</td>
</tr>
<tr>
<td>2.1.5</td>
<td>2</td>
<td>A</td>
<td>G</td>
<td>2683312</td>
<td>STY2819</td>
<td>putative oxidoreductase</td>
<td>168</td>
<td>56</td>
<td>4</td>
<td>4</td>
<td>0.87</td>
</tr>
<tr>
<td>2.1.6</td>
<td>9</td>
<td>A</td>
<td>G</td>
<td>4013386</td>
<td>STY4158</td>
<td>putative DNA-binding protein</td>
<td>73</td>
<td>25</td>
<td>1</td>
<td>3</td>
<td>0.58</td>
</tr>
<tr>
<td>2.1.7</td>
<td>49</td>
<td>A</td>
<td>G</td>
<td>4094437</td>
<td>STY4227</td>
<td>nickel responsive regulator</td>
<td>222</td>
<td>74</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>2.1.8</td>
<td>6</td>
<td>A</td>
<td>G</td>
<td>827036</td>
<td>STY0829</td>
<td>biotin synthesis protein BioC</td>
<td>93</td>
<td>31</td>
<td>1</td>
<td>1</td>
<td>0.26</td>
</tr>
<tr>
<td>2.1.9</td>
<td>12</td>
<td>T</td>
<td>C</td>
<td>3476114</td>
<td>STY3622</td>
<td>uroporphyrinogen III synthase</td>
<td>195</td>
<td>65</td>
<td>3</td>
<td>3</td>
<td>0.81</td>
</tr>
<tr>
<td>2.2</td>
<td>2</td>
<td>C</td>
<td>T</td>
<td>4355243</td>
<td>STY4468</td>
<td>lysR family regulatory protein</td>
<td>156</td>
<td>52</td>
<td>0</td>
<td>3</td>
<td>0.34</td>
</tr>
<tr>
<td>2.2.1</td>
<td>11</td>
<td>A</td>
<td>G</td>
<td>2723847</td>
<td>STY2853</td>
<td>putative sigma(54) modulation protein</td>
<td>300</td>
<td>100</td>
<td>0</td>
<td>2</td>
<td>0.39</td>
</tr>
<tr>
<td>2.2.2</td>
<td>9</td>
<td>C</td>
<td>T</td>
<td>4386609</td>
<td>STY4497</td>
<td>alpha-galactosidase</td>
<td>1092</td>
<td>364</td>
<td>3</td>
<td>2</td>
<td>0.37</td>
</tr>
<tr>
<td>2.2.3</td>
<td>7</td>
<td>A</td>
<td>G</td>
<td>703762</td>
<td>STY0707</td>
<td>glutamate/aspartate transport ATP-binding protein GltL</td>
<td>255</td>
<td>85</td>
<td>0</td>
<td>4</td>
<td>0.55</td>
</tr>
<tr>
<td>2.2.4</td>
<td>4</td>
<td>A</td>
<td>G</td>
<td>431216</td>
<td>STY0419</td>
<td>pyrroline-5-carboxylate reductase</td>
<td>351</td>
<td>117</td>
<td>0</td>
<td>2</td>
<td>0.25</td>
</tr>
<tr>
<td>2.3</td>
<td>72</td>
<td>A</td>
<td>G</td>
<td>3095443</td>
<td>STY3228</td>
<td>D-erythrose 4-phosphate dehydrogenase</td>
<td>717</td>
<td>239</td>
<td>2</td>
<td>7</td>
<td>0.86</td>
</tr>
<tr>
<td>2.3.1*</td>
<td>21</td>
<td>T</td>
<td>C</td>
<td>316186</td>
<td>STY0303</td>
<td>probable lipoprotein</td>
<td>354</td>
<td>118</td>
<td>0</td>
<td>2</td>
<td>0.37</td>
</tr>
<tr>
<td>2.3.2</td>
<td>15</td>
<td>T</td>
<td>C</td>
<td>1954711</td>
<td>STY2080</td>
<td>putative cation transporter</td>
<td>468</td>
<td>156</td>
<td>1</td>
<td>2</td>
<td>0.34</td>
</tr>
<tr>
<td>2.3.3</td>
<td>9</td>
<td>T</td>
<td>C</td>
<td>2811222</td>
<td>STY2937</td>
<td>glycine betaine-binding periplasmic protein precursor</td>
<td>48</td>
<td>16</td>
<td>2</td>
<td>1</td>
<td>0.30</td>
</tr>
</tbody>
</table>

52
<table>
<thead>
<tr>
<th>Group</th>
<th>Nt</th>
<th>Codon</th>
<th>ns</th>
<th>s</th>
<th>Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.4</td>
<td>12</td>
<td>T</td>
<td>C</td>
<td>3092900</td>
<td>fructose 1,6-bisphosphate aldolase</td>
</tr>
<tr>
<td>2.3.5a</td>
<td>15</td>
<td>G</td>
<td>A</td>
<td>2723724</td>
<td>putative sigma(54) modulation protein</td>
</tr>
<tr>
<td>2.4</td>
<td>38</td>
<td>C</td>
<td>A</td>
<td>3437570</td>
<td>see-independent protein translocase protein</td>
</tr>
<tr>
<td>2.4.1</td>
<td>24</td>
<td>A</td>
<td>G</td>
<td>1780319</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>2.5</td>
<td>42</td>
<td>T</td>
<td>C</td>
<td>1535365</td>
<td>pyridine nucleotide transhydrogenase subunit-beta</td>
</tr>
<tr>
<td>2.5.1</td>
<td>21</td>
<td>T</td>
<td>C</td>
<td>1792810</td>
<td>2-dehydro-3-deoxyphosphoacetate aldolase</td>
</tr>
<tr>
<td>3.1</td>
<td>399</td>
<td>T</td>
<td>C</td>
<td>3062270</td>
<td>llysyl RNA synthetase (LysRS)</td>
</tr>
<tr>
<td>3.0.1</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>1799842</td>
<td>ribose-phosphate pyrophosphokinase</td>
</tr>
<tr>
<td>3.0.2</td>
<td>2</td>
<td>T</td>
<td>C</td>
<td>432732</td>
<td>shikimate kinase II</td>
</tr>
<tr>
<td>3.1</td>
<td>66</td>
<td>G</td>
<td>A</td>
<td>3069182</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>3.1.1</td>
<td>25</td>
<td>A</td>
<td>G</td>
<td>2752615</td>
<td>3OS ribosomal subunit protein S16</td>
</tr>
<tr>
<td>3.1.2</td>
<td>27</td>
<td>T</td>
<td>C</td>
<td>3770391</td>
<td>ATP synthase subunit B</td>
</tr>
<tr>
<td>3.2</td>
<td>58</td>
<td>A</td>
<td>G</td>
<td>2269835</td>
<td>endonuclease IV</td>
</tr>
<tr>
<td>3.2.1</td>
<td>44</td>
<td>T</td>
<td>C</td>
<td>4215341</td>
<td>phosphoribulokinase</td>
</tr>
<tr>
<td>3.2.2</td>
<td>12</td>
<td>T</td>
<td>C</td>
<td>4602946</td>
<td>probable sugar phosphotransferase</td>
</tr>
<tr>
<td>3.3</td>
<td>62</td>
<td>T</td>
<td>C</td>
<td>3368641</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>3.3.1</td>
<td>32</td>
<td>A</td>
<td>G</td>
<td>2245432</td>
<td>cytidine deaminase</td>
</tr>
<tr>
<td>3.4</td>
<td>56</td>
<td>A</td>
<td>C</td>
<td>3164162</td>
<td>hydrogenase-2 component protein</td>
</tr>
<tr>
<td>3.5</td>
<td>122</td>
<td>A</td>
<td>G</td>
<td>3923165</td>
<td>deoxyuridine 5'-triphosphate nucleotidohydrolase</td>
</tr>
<tr>
<td>3.5.1</td>
<td>1</td>
<td>T</td>
<td>C</td>
<td>1811809</td>
<td>hydrogenase-1 operon protein HyaF</td>
</tr>
<tr>
<td>3.5.2</td>
<td>5</td>
<td>T</td>
<td>G</td>
<td>3729635</td>
<td>glutamine synthetase</td>
</tr>
<tr>
<td>3.5.3a</td>
<td>6</td>
<td>T</td>
<td>C</td>
<td>3817752</td>
<td>hypothetical protein</td>
</tr>
<tr>
<td>3.5.4</td>
<td>31</td>
<td>T</td>
<td>C</td>
<td>183033</td>
<td>dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase</td>
</tr>
<tr>
<td>4</td>
<td>1089</td>
<td>T</td>
<td>C</td>
<td>1615350</td>
<td>conserved hypothetical protein</td>
</tr>
<tr>
<td>4.1</td>
<td>96</td>
<td>A</td>
<td>G</td>
<td>2342045</td>
<td>ribonucleoside-diphosphate reductase 1 beta chain</td>
</tr>
<tr>
<td>4.1.1</td>
<td>18</td>
<td>T</td>
<td>C</td>
<td>3996717</td>
<td>alpha-amylose</td>
</tr>
<tr>
<td>4.2</td>
<td>140</td>
<td>A</td>
<td>G</td>
<td>2640029</td>
<td>SseB protein</td>
</tr>
<tr>
<td>4.2.1</td>
<td>40</td>
<td>T</td>
<td>C</td>
<td>989024</td>
<td>killing factor KicB</td>
</tr>
<tr>
<td>4.2.2</td>
<td>77</td>
<td>A</td>
<td>G</td>
<td>3806278</td>
<td>chromosomal replication initiator protein</td>
</tr>
<tr>
<td>4.2.3</td>
<td>12</td>
<td>A</td>
<td>G</td>
<td>1611156</td>
<td>putative oxidoreductase</td>
</tr>
<tr>
<td>4.3.1</td>
<td>853</td>
<td>T</td>
<td>C</td>
<td>2348902</td>
<td>anaerobic glycerol-3-phosphate dehydrogenase subunit A</td>
</tr>
</tbody>
</table>

Key: Group = Primary cluster/Clade/Subclade; N = number of S. Typhi isolates in the Group; In = allele within the group; Out = allele outside of group; Nt = nucleotide position in the gene; Codon = codon position in the gene; ns = number of non-synonymous SNPs in the gene; s = number of synonymous SNPs in the gene; %Div = % of diversity in the gene; a nested in 2.3.2; b nested in 2.3.3; c BiP48; d nested in 3.5.4; e BiP33
2.6.1.3 Functional analysis

First, SNPs occurring exclusively in carrier isolates were identified. Genes containing these SNPs were subsequently grouped by their predicted function based on the *Salmonella* Typhi functional classification scheme used by the Sanger Institute (www.sanger.ac.uk). This functional classification scheme is based on the genome annotation of *S*. Typhi strain CT18 \(^{233}\). A summary of main functional classes in this scheme was shown in Table 2.2. Similar analyses were performed to classify genes containing SNPs that occurred exclusively in acute isolates to different functional classes for comparisons with that of carrier isolates.

2.6.1.4 Pairwise SNP distance

Pairwise genetic distances (defined as difference in the number of SNPs) within and between acute and carrier isolates were estimated from their SNP alignment using packages ape (v4.1) and adegenet (v2.0.1) in R (v3.3.2). Pairwise SNP distances were extracted and plotted using the function pairDistPlot in adegenet package. Subsequently, the Wilcoxon rank sum test was used for testing the difference in the average pairwise SNP distances between groups.
<table>
<thead>
<tr>
<th>Category</th>
<th>Subcategory</th>
<th>Function</th>
<th>Related Functions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.A Degradation</td>
<td>1.G.1 Biotin</td>
<td>3.C.4 Murein sacculus, peptidoglycan</td>
<td></td>
</tr>
<tr>
<td>1.B.3 Tricarboxylic acid cycle</td>
<td>1.G.2 Folic Acid</td>
<td>4.A.6 Transport Other</td>
<td></td>
</tr>
<tr>
<td>1.B.5 b Non-oxydative branch</td>
<td>1.G.3 Lipote</td>
<td>4.B Chaperones</td>
<td></td>
</tr>
<tr>
<td>1.B.7.a Aerobic Respiration</td>
<td>1.G.5 Pentothenate</td>
<td>4.D Chemotaxis and mobility</td>
<td></td>
</tr>
<tr>
<td>1.B.8 Fermentation</td>
<td>1.G.8 Thiamine</td>
<td>4.I Pathogenicity</td>
<td></td>
</tr>
<tr>
<td>1.B.9 ATP-proton motive force</td>
<td>1.G.9 Riboflavin</td>
<td>5.A IS element, Phage related</td>
<td></td>
</tr>
<tr>
<td>1.C Central intermediary metabolism</td>
<td>1.H Fatty acid biosynthesis</td>
<td>5.B Colcin-related function</td>
<td></td>
</tr>
<tr>
<td>1.C.2 Gluconeogenesis</td>
<td>2 Broad regulatory function</td>
<td>5.D Drug/Analogue sensitivity</td>
<td></td>
</tr>
<tr>
<td>1.C.3 Sugar-nucleotide biosynthesis, conversions</td>
<td>3.A.10 Polysaccharides - (cytoplasmic)</td>
<td>5.F Adaptions and atypical conditions</td>
<td></td>
</tr>
<tr>
<td>1.C.4 Amino sugars</td>
<td>3.A.11 Phospholipids</td>
<td>5.H.a Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1.C.5 Sulphur Metabolism</td>
<td>3.A.2 Ribosomal proteins - synthesis, modification</td>
<td>5.H.b Hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>1.D Amino Acid Biosynthesis</td>
<td>3.A.3 Ribosomes - maturation and modification</td>
<td>5.I Unknown</td>
<td></td>
</tr>
<tr>
<td>1.D.1 Glutamate Family</td>
<td>3.A.5 Amino acyl tRNA synthesis; tRNA modification</td>
<td>SPI-1</td>
<td></td>
</tr>
<tr>
<td>1.D.2 Aspartate Family</td>
<td>3.A.7 DNA - replication, repair, restriction./modification</td>
<td>SPI-10</td>
<td></td>
</tr>
<tr>
<td>1.D.4 Aromatic Amino Acid</td>
<td>3.A.9 RNA synthesis, modification</td>
<td>SPI-3</td>
<td></td>
</tr>
<tr>
<td>1.D.5 Histidine</td>
<td>3.B.1 Degradation of RNA</td>
<td>SPI-4</td>
<td></td>
</tr>
<tr>
<td>1.E Polyamine synthesis</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
<td>SPI-6</td>
<td></td>
</tr>
<tr>
<td>1.F.1 Purine ribonucleotide biosynthesis</td>
<td>3.B.4 Degradations of polysaccharides</td>
<td>SPI-7</td>
<td></td>
</tr>
<tr>
<td>1.F.2 Pyrimidine ribonucleotide 2'-Deoxyribonucleotide metabolism</td>
<td>3.C Cell envelope</td>
<td>SPI-8</td>
<td></td>
</tr>
<tr>
<td>1.F.3 3'C'-Deoxyribonucleotide metabolism</td>
<td>3.C.Membranes lipoprotein</td>
<td>SPI-9</td>
<td></td>
</tr>
<tr>
<td>1.F.4 Salvage of nucleosides and nucleotides</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.F.5 Miscellaneous</td>
<td>3.C.3 Surface structure</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.6.2 SNP detection and analysis for *Salmonella* Paratyphi A genomes

Raw Illumina reads were mapped to the reference sequence of *S.* Paratyphi A strain AKU_12601 (accession no: FM200053) and plasmid pAKU_12601 (accession no: AM412236) using SMALT version 0.7.4 (http://www.sanger.ac.uk/resources/software/smalt/). Candidate single nucleotide polymorphisms (SNPs) were identified against the reference sequence using SAMtools and filtered with a minimal Phred quality of 30 and a quality cut-off of 0.75. The allele at each locus in each isolate was determined by reference to the consensus base in that genome, using *samtools mpileup* and removing low confidence alleles with consensus base quality ≤20, read depth ≤5 or a heterozygous base call. SNPs called in prophage regions, repetitive sequences (IS elements, transposases, etc.) identified previously were excluded. Subsequently, Gubbins software was used to identify potential recombinant regions from whole genome alignment produced by SNP-calling isolates. Detected regions were manually checked and SNPs found in these regions were also removed.

2.7 Phylogenetic analysis

2.7.1 Phylogenetic analysis for chapter 3

A maximum likelihood (ML) phylogeny was estimated using a 1440 SNP alignment of the 78 RCT isolates in RAxML (version 7.8.6) with the generalized time-reversible substitution model (GTR) and a gamma distribution, with support for the phylogeny assessed via 1000 bootstrap replicates. The alignment was then compared to a global *S.* Typhi sequence database, with a particular focus on identifying sequences with a mutational profile suggestive of shared ancestry with a divergent H58 clade identified
in the previous phylogeny. A secondary ML phylogenetic tree was then inferred from the SNP alignment (1642 SNPs) of the 136 Nepalese _S. Typhi_ along with 19 recently described _S. Typhi_ H58 with the aforementioned mutational profile, using the same parameters as above.

2.7.2 Phylogenetic analysis for chapter 4

The best-fit evolutionary model for the SNP alignment of 223 _S. Paratyphi A_ isolates was identified based on the Bayesian Information Criterion in jModelTest implemented in IQ-TREE software and maximum likelihood phylogenetic trees were subsequently reconstructed under the selected model (TVM). Clade support for this maximum likelihood tree was assessed via 1000 bootstrap pseudo-analyses. Phylogenetic subgrouping of major lineages in the phylogenetic tree was performed on the basis of monophyletic groups (at least three isolates) with moderately to highly supported bootstrap values (>60%). In order to investigate the population structure of Nepalese _S. Paratyphi A_ in the global context, a secondary maximum likelihood phylogenetic tree was inferred from a separate alignment of 3958 SNPs of the 223 Nepalese _S. Paratyphi A_ along with 111 _S. Paratyphi A_ isolates from a global sequence database described previously. Bootstrap support for this maximum likelihood tree was estimated from 100 pseudo-replicates.

2.7.3 Phylogenetic analysis for chapter 5

A maximum likelihood phylogenetic tree was constructed from a 188 chromosomal SNP alignment of H58 isolates with RAxML (version 7.8.6) using the generalized time-reversible model (GTR) and a gamma distribution to model site-specific rate
variation (GTR+Γ nucleotide substitution model in RAxML). Support for the ML phylogeny was assessed via 1,000 bootstrap pseudo-analyses of the alignment data. Phylogenetic subgrouping was defined based on monophyletic groups (lineages) with bootstrap values indicative of strong support (≥85%). To investigate the short-term divergence within the bacterial population and the transmission within the local population, a minimum spanning tree was reconstructed from the SNP alignment of lineage III and lineage IV identified in the ML tree (accounting for 95% of isolates) using the goeBURST algorithm in Phyloviz software (version 1.1)\(^{268}\). This algorithm identified seven sublineages based on similarity among allelic profiles and frequency of isolation within the population. Sequences with identical SNP profiles and isolated at the highest frequency within each sublineage were assigned as founder genotypes (viewed as the central nodes within each of the sublineages), with descendant genotypes (represented by terminal nodes surrounding the founder genotype) assigned based on similarity to founder SNP profiles. These descendent genotypes can differ from the parental genotype by a single or multiple SNPs.

### 2.7.4 Phylogenetic analysis for chapter 6

A maximum likelihood phylogenetic tree was reconstructed from the SNP alignment of 120 \(S.\) Typhi isolates, plus a \(S.\) Paratyphi A isolate included as an outgroup to root the tree, using RAxML version 8.2.8 with the generalized time-reversible model and a Gamma distribution to model the site-specific rate variation (GTR+Γ). Support for the maximum likelihood tree was assessed via bootstrap analysis with 1000 pseudoreplicates. To investigate the phylogenetic structure of Nepalese H58 \(S.\) Typhi in the global context, a secondary ML phylogenetic tree was inferred from a separate
SNP alignment of 78 Nepalese H58 S. Typhi along with 836 globally representative H58 S. Typhi described previously. Support for this ML tree was assessed via 100 bootstrap replicates.

2.8 Resistance gene and plasmid analysis

Investigation of antimicrobial resistance genes and plasmids of S. Typhi and S. Paratyphi A isolates was performed using a local assembly approach with ARIBA (Antimicrobial Resistance Identifier by Assembly). Resfinder and Plasmidfinder were used as databases of antimicrobial resistant genes and plasmid replicons, respectively. Briefly, reference sequences in the database are clustered by similarity (minimum 90% sequence identity) and paired sequence reads are mapped to the reference sequences to generate a set of reads for each cluster. The reads for each cluster are assembled independently and the closest reference sequence to the resulting contigs is identified. The reads for the cluster are subsequently mapped back to the assembled contig to identify variants. ARIBA not only reports the acquired antimicrobial resistance genes, but also the quality of assemblies and any variants detected between the sequencing reads and the reference sequences including known resistance SNPs.

2.9 Pan-genome analysis

Raw reads from each isolate were de novo assembled using the short-read assembler Velvet with parameters optimized by Velvet Optimizer. Contigs that were less than 300 bp long were excluded and the assembled contigs were annotated using Prokka. The pan-genome was subsequently reconstructed using Roary.
large-scale prokaryote pan-genome analysis) with a threshold set to 95% sequence similarity at the amino-acid level. Unique and shared gene content between groups of isolates was identified based on set operations on the pan-genome.

2.10 Spatiotemporal cluster analysis

2.10.1 Cluster analysis for chapter 4

Spatial and spatiotemporal clustering analyses were performed using SaTScan v9.4. A Bernoulli model was used to examine the spatial clusters of each identified genotype, using all other genotypes as the background distribution of *S. Paratyphi A* cases in Lalitpur, Kathmandu. The upper limit for cluster detection was specified as 25% of the study population over 25% of the study duration (for spatiotemporal clustering). The significance of the detected clusters was assessed by a likelihood ratio test, with a *p*-value obtained by 999 Monte Carlo simulations generated under the null hypothesis of a random spatial and spatiotemporal distribution. All maps were created in ArcGIS 10.2 (ESRI, Redlands, CA, USA).

2.10.2 Spatiotemporal clustering detection for chapter 5

Spatiotemporal clustering analysis of *S. Typhi* in Siem Reap, Cambodia was performed using Moran’s I and SaTScan methodologies. First, Moran’s I test was used to evaluate global autocorrelation amongst communes that reported at least one case (n=78) of typhoid fever in GeoDa software (v1.6.7, https://geodacenter.asu.edu/). This test statistic provides an evaluation of whether the rates across the area of interest are spatially random (Moran’s I=0), over-dispersed (Moran’s I<0) or clustered (Moran’s I>0). Next, Kulldorff’s scan statistic in SaTScan (v9.1.1,
http://www.satscan.org/) was used to identify the location of clusters of communes with high rates of typhoid fever over space and time. A cylindrical window was used to scan the area for clusters, with the size of the circle corresponding to the spatial scan and the height of the cylinder corresponding to time. The significance of the detected clusters was assessed by a likelihood ratio test, with a p-value obtained by 999 Monte Carlo simulations generated under the null hypothesis of random spatiotemporal distribution. In this analysis, scan windows were used to fit discrete Poisson models. For the sublineage-specific analyses, all case communes were included and those without cases of a specific sublineage were classified as having 0 cases. The upper limit for cluster detection was specified as 25% of the study population over each year. All maps were created in ArcGIS 10.2 (ESRI, Redlands, CA, USA).

### 2.11 Statistical analysis

#### 2.11.1 Statistical analysis for chapter 3

Comparison of baseline characteristics patient groups, stratified by the H58 status or susceptibility category of their corresponding S. Typhi isolates was performed using the Kruskal Wallis test for continuous variables and Fisher’s exact test for categorical variables. Time to treatment failure was analysed using Firth’s penalized maximum likelihood bias reduction method for Cox regression as a solution for the non-convergence of likelihood function in the case of zero event counts in subgroups. For comparisons between treatment arms, H58 status, or ciprofloxacin susceptibility group, the model included treatment arm, H58 status, or susceptibility group as a single covariate. Confidence intervals (CI) and p-values were calculated by profile-
penalized likelihood. FCT was analyzed as an interval-censored outcome, i.e. as the time interval from the last febrile temperature assessment until the first afebrile assessment, using parametric Weibull accelerated failure time models. Median and inter-quartile range (IQR) FCT calculations for subgroups were based on models for each subgroup separately. Acceleration factors were based on models that included treatment arm as the only covariate. The non-parametric maximum likelihood estimator (NPMLE) was used to visualize the distribution of FCT between groups. Heterogeneity between subgroups was tested with models that included an interaction between treatment arm and the sub-grouping variable. All analyses were performed using R software version 3.2.2.

2.11.2 Statistical analysis for chapter 5

Rates of hospitalized typhoid fever were calculated at the commune level using the population under the age of 15 years from 2008. Multivariable negative binomial regression was used to identify commune-level risk factors associated with the rate of cases per 1,000 population under the age of 15 years. Interaction between commune level factors was evaluated using the likelihood ratio test. Variables included in the evaluation of the final model included those with significant associations ($p<0.10$) in the univariate analysis and $a\ priori$ sanitation and water source variables. Variables that did not add significantly to the fit of the final model (determined by the likelihood ratio test) were not included. All analyses were performed in STATA (v13, College Station, TX, USA) and plots were created in R v3.2.2 (R Foundation for Statistical Computing, Vienna, Austria, https://cran.r-project.org/) using ggplot2.
Emergence of a novel ciprofloxacin-resistant subclade of H58

Salmonella Typhi associated with fluoroquinolone treatment failure in Nepal

3.1 Introduction

Antimicrobial treatment is essential to effectively manage typhoid fever and avoid serious complications. However, antimicrobial resistance in S. Typhi has now become a major global health problem, limiting the treatment options and increasing the treatment failure rates. Over the last 30 years, the extensive use of antimicrobials for typhoid treatment has successively driven the emergence and global dissemination of resistant organisms. First line drugs of choice such as chloramphenicol, amoxicillin and trimethoprim-sulfamethoxazole have been rendered ineffective due to the emergence of MDR S. Typhi since the 1980s. Fluoroquinolones were subsequently used in the 1990s and became standard drugs of choice after officially recommended by WHO for typhoid treatment in 2003. However, the effectiveness of this group of compounds has been diminished since S. Typhi isolates with reduced susceptibility to fluoroquinolones started to emerge. Reduced susceptibility to fluoroquinolones in S. Typhi isolates is induced by point mutation(s) in quinolone resistance determining region in the gyrA gene and such organisms appear to have a fitness advantage, even in the absence of antimicrobial exposure. It is now known that the widespread of S. Typhi isolates with reduced susceptibility to
fluoroquinolones has been partly facilitated by the dissemination of a specific MDR lineage (H58) across Asia and Africa\textsuperscript{248}. H58 lineage has been replacing other genotypes rapidly and reshaping the global population structure of \textit{S. Typhi}.

Previous studies have shown that patients infected with \textit{S. Typhi} isolates with elevated Minimum Inhibitory Concentrations (MIC) against ciprofloxacin and ofloxacin are more likely to have prolonged fever clearance time (FCTs) and more frequent treatment failure\textsuperscript{89,90,283,285}. Whilst older fluoroquinolones have showed a reduced efficacy in curing the disease, the fourth-generation fluoroquinolone, gatifloxacin, has been demonstrated as highly efficacious for uncomplicated typhoid, even in patients infected with \textit{S. Typhi} isolates displaying reduced susceptibility to ciprofloxacin (MIC $\geq$0.125 $\mu$g/ml)\textsuperscript{257,259,286,287}. However, during a randomised controlled trial (RCT) comparing gatifloxacin and ceftriaxone conducted recently in Nepal, there was an increased number of treatment failures associated with \textit{S. Typhi} strains with high level of ciprofloxacin and gatifloxacin resistance (ciprofloxacin MIC $>$32$\mu$g/ml, gatifloxacin MIC $>$1$\mu$g/ml), prompting the data safety and monitoring board to stop the trial\textsuperscript{288}. In this study, I aimed to investigate the molecular epidemiology of the \textit{S. Typhi} isolates from this trial and understand how bacterial genotype may affect the treatment outcome. All \textit{S. Typhi} isolates were genome-sequenced and stratified by genotype and the association between genotype and clinical presentation and outcome was assessed. The findings from my research were published in eLife journal in 2016 (appendix G)\textsuperscript{289}.
3.2 Results

3.2.1 Salmonella Typhi whole genome sequencing

I performed whole genome sequencing (WGS) on 78 S. Typhi isolates from patients in both RCT treatment arms (gatifloxacin and ceftriaxone) (Appendix A). The resulting phylogeny indicated that the majority of isolates (65/78; 83.3%) fell within the H58 lineage, while the remaining 13 (16.7%) represented eight different lineages (Figure 3.1). All but four of the H58 strains contained the common DNA gyrase (gyrA) mutation in codon 83 (S83F), which confers reduced susceptibility against fluoroquinolones (FQs) (ciprofloxacin MIC; 0.125- 0.5 μg/ml) \(^{290}\). Nested within the S83F H58 group but separated from the rest of the group by a branch defined by 30 SNPs, was a H58 subclade comprised of 12 isolates containing the S83F gyrA mutation, a mutation in gyrA at codon 87 (D87N), and an additional mutation in the topoisomerase gene, parC (S80I) (H58 triple mutant). Notably, these H58 triple mutants shared high MICs against ciprofloxacin (≥24 μg/ml). Further, an additional two non-H58 RCT isolates with ciprofloxacin MIC≥24 μg/ml had the S83F gyrA mutation, an alternative mutation at codon 87 (D87V), the S80I parC mutation, and an A364V mutation in parE (Figure 3.1, Appendix A).
Figure 3.1 The phylogenetic structure of 78 Nepali *Salmonella* Typhi isolated during a gatifloxacin versus ceftriaxone randomised controlled trial

Maximum likelihood phylogeny based on core-genome SNPs of 78 *Salmonella* Typhi isolates with the corresponding metadata, including the presence of mutations (dark grey) in *gyrA* (S83F, D87V and D87N), *parC* (S80I) and *parC* (A364V) and susceptibility to ciprofloxacin (susceptible, light blue; intermediate, mid-blue and non-susceptible, dark blue) by Minimum Inhibitory Concentration (MIC). Missing metadata is indicated by black boxes. Red lines linking to metadata show isolates belonging to the *Salmonella* Typhi H58 lineage (with H58 triple mutants highlighted), other lineages (non-H58) are shown with black lines. The scale bar indicates the number of substitutions per variable site (see methods). Asterisks indicate ≥85% bootstrap support at nodes of interest.
3.2.2 Clinical presentation of *Salmonella* Typhi infections

I stratified clinical data from the RCT by H58 status of the corresponding *S*. Typhi isolates (H58; N=65, non-H58; N=13) and compared baseline characteristics between these groups. I found no significant differences in demographics and no association between disease severity at presentation between those infected with an H58 *S*. Typhi isolate or a non-H58 isolate (Table 3.1). Next, I compared the baseline characteristics of patients stratified by ciprofloxacin susceptibility (susceptible, intermediate and resistant), and found no differences in disease severity or demographics on presentation; the only exception being FQ resistant *S*. Typhi were more frequently isolated from adults (Table 3.2). A significantly lower proportion of H58 *S*. Typhi (4/65; 6.2%) were susceptible to FQs compared to non-H58 isolates (6/13; 46%) ($p=0.001$) (Table 3.3) and, overall, H58 isolates had significantly higher MICs against the majority of tested antimicrobials than non-H58 isolates (Table 3.3).
Table 3.1 Baseline characteristics by *Salmonella* Typhi lineage

<table>
<thead>
<tr>
<th>Baseline characteristic</th>
<th>n</th>
<th>Non-H58 (N=13)</th>
<th>n</th>
<th>H58 (N=65)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) – median (IQR)</td>
<td>13</td>
<td>18.0 (13.0,21.0)</td>
<td>65</td>
<td>18.0 (13.0,22.0)</td>
<td>0.75</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>13</td>
<td>12 (92.3%)</td>
<td>65</td>
<td>46 (70.8%)</td>
<td>0.17</td>
</tr>
<tr>
<td>Temperature (°C) - median (IQR)</td>
<td>13</td>
<td>39.0 (38.3,39.4)</td>
<td>62</td>
<td>39.0 (38.3,39.4)</td>
<td>0.77</td>
</tr>
<tr>
<td>Days of illness before enrolment - median (IQR)</td>
<td>13</td>
<td>5.0 (4.0,7.0)</td>
<td>65</td>
<td>5.0 (4.0,7.0)</td>
<td>0.39</td>
</tr>
<tr>
<td>Antimicrobials in last two weeks</td>
<td>13</td>
<td>1 (7.7%)</td>
<td>65</td>
<td>9 (13.8%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Previous history of typhoid</td>
<td>13</td>
<td>1 (7.7%)</td>
<td>65</td>
<td>5 (7.7%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Family history of typhoid</td>
<td>13</td>
<td>1 (7.7%)</td>
<td>65</td>
<td>8 (12.3%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Typhoid vaccination</td>
<td>13</td>
<td>1 (7.7%)</td>
<td>65</td>
<td>0 (0%)</td>
<td>0.17</td>
</tr>
<tr>
<td>Fever</td>
<td>13</td>
<td>13 (100%)</td>
<td>64</td>
<td>64 (100%)</td>
<td>NA</td>
</tr>
<tr>
<td>Cough</td>
<td>12</td>
<td>4 (33.3%)</td>
<td>62</td>
<td>15 (24.2%)</td>
<td>0.49</td>
</tr>
<tr>
<td>Constipation</td>
<td>12</td>
<td>1 (8.3%)</td>
<td>63</td>
<td>4 (6.3%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Headache</td>
<td>13</td>
<td>12 (92.3%)</td>
<td>64</td>
<td>59 (92.2%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>12</td>
<td>5 (41.7%)</td>
<td>63</td>
<td>24 (38.7%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Vomiting</td>
<td>12</td>
<td>4 (33.3%)</td>
<td>63</td>
<td>16 (25.4%)</td>
<td>0.72</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>12</td>
<td>2 (16.7%)</td>
<td>62</td>
<td>18 (29.0%)</td>
<td>0.49</td>
</tr>
<tr>
<td>Anorexia</td>
<td>12</td>
<td>9 (75.0%)</td>
<td>64</td>
<td>53 (82.8%)</td>
<td>0.68</td>
</tr>
<tr>
<td>Nausea</td>
<td>12</td>
<td>9 (75.0%)</td>
<td>61</td>
<td>38 (62.3%)</td>
<td>0.52</td>
</tr>
<tr>
<td>Leucocyte count (×10⁹/L) - median (IQR)</td>
<td>13</td>
<td>5.9 (4.5,6.6)</td>
<td>65</td>
<td>6.2 (4.8,7.3)</td>
<td>0.54</td>
</tr>
<tr>
<td>Neutrophils (%) - median (IQR)</td>
<td>13</td>
<td>65.0 (60.7,3.0)</td>
<td>65</td>
<td>70 (65.0,75.0)</td>
<td>0.17</td>
</tr>
<tr>
<td>Lymphocytes (%) - median (IQR)</td>
<td>13</td>
<td>33.0 (22.0,37.0)</td>
<td>65</td>
<td>28.0 (22.0,34.0)</td>
<td>0.32</td>
</tr>
<tr>
<td>Haematocrit (%) - median (IQR)</td>
<td>12</td>
<td>38.7 (37.1,41.2)</td>
<td>65</td>
<td>38.0 (35.0,41.4)</td>
<td>0.39</td>
</tr>
<tr>
<td>Platelet count (×10⁹/L) - median (IQR)</td>
<td>13</td>
<td>177.0 (165.0,189.0)</td>
<td>65</td>
<td>158.0 (140.2,210)</td>
<td>0.24</td>
</tr>
<tr>
<td>AST (U/L) - median (IQR)</td>
<td>12</td>
<td>52.5 (38.2,64.8)</td>
<td>62</td>
<td>59.0 (44.0,83.8)</td>
<td>0.21</td>
</tr>
<tr>
<td>ALT (U/L) - median (IQR)</td>
<td>13</td>
<td>46.0 (37.0,57.0)</td>
<td>63</td>
<td>50 (35.5,66.5)</td>
<td>0.18</td>
</tr>
</tbody>
</table>

N refer to the number of patients in each group
n refers to the number of observations with non-missing data for the respective characteristic
*Comparisons between the two were done using Fisher’s exact test for categorical variables and the Wilcoxon rank sum test for continuous variables
Table 3.2 Baseline characteristics by *Salmonella* Typhi ciprofloxacin susceptibility

<table>
<thead>
<tr>
<th>Baseline characteristic</th>
<th>n</th>
<th>Susceptible (N=10)</th>
<th>n</th>
<th>Intermediate (N=52)</th>
<th>n</th>
<th>Resistant (N=16)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years) – median (IQR)</td>
<td>10</td>
<td>14.0 (9.2,17.0)</td>
<td>52</td>
<td>18.0 (13.0,22.0)</td>
<td>16</td>
<td>20 (17.8,21.2)</td>
<td>0.05</td>
</tr>
<tr>
<td>Sex (male)</td>
<td>10</td>
<td>8 (80%)</td>
<td>52</td>
<td>39 (75.0%)</td>
<td>16</td>
<td>11 (68.8%)</td>
<td>0.86</td>
</tr>
<tr>
<td>Temperature (°C) - median (IQR)</td>
<td>10</td>
<td>39.0 (38.6,39.4)</td>
<td>49</td>
<td>39.0 (38.3,39.4)</td>
<td>16</td>
<td>38.8 (38.3,39.3)</td>
<td>0.42</td>
</tr>
<tr>
<td>Days of illness before enrolment - median (IQR)</td>
<td>10</td>
<td>4.0 (4.0,6.0)</td>
<td>52</td>
<td>5.0 (4.0,7.0)</td>
<td>16</td>
<td>5.0 (4.0,7.0)</td>
<td>0.88</td>
</tr>
<tr>
<td>Antimicrobials in last two weeks</td>
<td>10</td>
<td>1 (10%)</td>
<td>52</td>
<td>7 (13.5%)</td>
<td>16</td>
<td>2 (12.5%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Previous history of typhoid</td>
<td>10</td>
<td>0 (0%)</td>
<td>52</td>
<td>6 (11.5%)</td>
<td>16</td>
<td>0 (0%)</td>
<td>0.34</td>
</tr>
<tr>
<td>Family history of typhoid</td>
<td>10</td>
<td>1 (10%)</td>
<td>52</td>
<td>6 (11.5%)</td>
<td>16</td>
<td>2 (12.5%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Typhoid vaccination</td>
<td>10</td>
<td>1 (10%)</td>
<td>52</td>
<td>0 (0%)</td>
<td>16</td>
<td>0 (0%)</td>
<td>0.13</td>
</tr>
<tr>
<td>Fever</td>
<td>10</td>
<td>10 (100%)</td>
<td>51</td>
<td>51 (100%)</td>
<td>16</td>
<td>16 (100%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Cough</td>
<td>10</td>
<td>2 (20%)</td>
<td>49</td>
<td>13 (26.5%)</td>
<td>15</td>
<td>4 (26.7%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Constipation</td>
<td>10</td>
<td>0 (0%)</td>
<td>50</td>
<td>4 (8.0%)</td>
<td>15</td>
<td>1 (6.7%)</td>
<td>1.00</td>
</tr>
<tr>
<td>Headache</td>
<td>10</td>
<td>10 (100%)</td>
<td>51</td>
<td>45 (88.2%)</td>
<td>16</td>
<td>16 (100%)</td>
<td>0.34</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>10</td>
<td>3 (30%)</td>
<td>50</td>
<td>21 (42.0%)</td>
<td>14</td>
<td>5 (35.7%)</td>
<td>0.82</td>
</tr>
<tr>
<td>Vomiting</td>
<td>10</td>
<td>2 (20%)</td>
<td>50</td>
<td>14 (28.0%)</td>
<td>15</td>
<td>4 (26.7%)</td>
<td>0.93</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>10</td>
<td>1 (10%)</td>
<td>50</td>
<td>16 (32.0%)</td>
<td>14</td>
<td>3 (21.4%)</td>
<td>0.38</td>
</tr>
<tr>
<td>Anorexia</td>
<td>10</td>
<td>8 (80%)</td>
<td>51</td>
<td>41 (80.4%)</td>
<td>15</td>
<td>13 (86.7%)</td>
<td>0.91</td>
</tr>
<tr>
<td>Nausea</td>
<td>10</td>
<td>6 (60%)</td>
<td>48</td>
<td>31 (64.6%)</td>
<td>15</td>
<td>10 (66.7%)</td>
<td>1.00</td>
</tr>
</tbody>
</table>

N refers to the number of patients in each group
n refers to the number of observations with non-missing data for the respective characteristic
* Comparison between the three groups were done using Fisher’s exact test for categorical variables and the Kruskal-Wallis test for continuous variables
Table 3.3 Comparison of antimicrobial susceptibility by *Salmonella* Typhi lineage

<table>
<thead>
<tr>
<th>E test</th>
<th>Non-H58 (N=13)</th>
<th>H58 (N=65)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MIC50</td>
<td>MIC90</td>
<td>GM (range)</td>
</tr>
<tr>
<td>Amoxicillin</td>
<td>0.5</td>
<td>1</td>
<td>0.77 (0.38–38)</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>3</td>
<td>4</td>
<td>2.7 (1.5–8)</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>0.06</td>
<td>0.06</td>
<td>0.06 (0.05–0.13)</td>
</tr>
<tr>
<td>Gatifloxacin</td>
<td>0.13</td>
<td>0.25</td>
<td>0.06 (0.01–2)</td>
</tr>
<tr>
<td>Nalidixic acid</td>
<td>&gt;256</td>
<td>&gt;256</td>
<td>21.6 (1–256)</td>
</tr>
<tr>
<td>Ofloxacin</td>
<td>0.25</td>
<td>0.75</td>
<td>0.24 (0.03–32)</td>
</tr>
<tr>
<td>Trimethoprim sulphate</td>
<td>0.02</td>
<td>0.05</td>
<td>0.03 (0.02–0.05)</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>0.13</td>
<td>0.75</td>
<td>0.11 (0.01–32)</td>
</tr>
</tbody>
</table>

**Ciprofloxacin susceptibility group**

<table>
<thead>
<tr>
<th></th>
<th>Non-H58 (N=13)</th>
<th>H58 (N=65)</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Susceptible</td>
<td>6 (46.2%)</td>
<td>4 (6.2%)</td>
<td></td>
</tr>
<tr>
<td>- Intermediate</td>
<td>4 (30.8%)</td>
<td>48 (73.8%)</td>
<td></td>
</tr>
<tr>
<td>- Resistant</td>
<td>3 (23.1%)</td>
<td>13 (20.0%)</td>
<td></td>
</tr>
</tbody>
</table>

*Comparisons between *Salmonella* Typhi lineage for MICs and ciprofloxacin susceptibility groups were based on the Wilcoxon rank sum test and Fisher’s exact test, respectively.

MIC: minimum inhibitory concentration, measured in µg/ml

#p value for comparison of susceptible vs. intermediate/resistant combined between groups by Fisher’s exact test is 0.001.

GM: geometric mean, the upper range of the values was determined by multiplying the MIC by 2 if the result was >X (for example, >256 = 256*2 = 512).
3.2.3 Treatment failure and fever clearance times

The primary endpoint of the RCT in which these data were generated was a composite for treatment failure (method section 2.4.1). Treatment failure with H58 S. Typhi was significantly less common in the ceftriaxone group (3/31; 9.7%) than the gatifloxacin group (15/34; 44.1%) (Hazard Ratio (HR) of time to failure 0.19, 95%CI 0.05-0.56, \( p=0.002 \)) (Table 3.4). Conversely, there was no significant difference in treatment failure between those infected with non-H58 isolates treated with gatifloxacin (0/6; 0%) or ceftriaxone (2/7; 28.6%) (\( p=0.32 \)). Similarly, time to fever clearance differed significantly between the two treatment groups in H58 infections, with median FCTs of 5.03 days (interquartile range (IQR): 3.18-7.21) in the gatifloxacin group and 3.07 days (IQR: 1.89-4.52) in the ceftriaxone group (\( p<0.0006 \)). Again, this trend was not mirrored in the non-H58 S. Typhi infections, with FCTs of 2.87 (IQR: 2.08-3.7) and 3.12 (IQR: 2.2-4.12) days for gatifloxacin and ceftriaxone, respectively (\( p=0.61 \)) (Table 3.4). Moreover, in the gatifloxacin arm, H58 S. Typhi tended to be associated with a higher risk of treatment failure (\( p=0.06 \)) (Figure 3.2.A, Table 3.5) and a lengthier fever clearance time (\( p=0.013 \)) (Figure 3.2.C, Table 3.4).

As I identified two non-H58 isolates that were also FQ-resistant (Figure 3.1), I additionally stratified outcome for the gatifloxacin arm (N=40 patients) by FQ susceptibility of the infecting organism. Those infected with FQ-resistant S. Typhi failed gatifloxacin treatment more frequently (8/10; 80%) than those infected with an intermediately resistant organism (7/25; 28%) or a susceptible organism (0/5; 0%) (\( p=0.007 \)) (Figure 3.2.B, Tables 3.6 and 3.7). Furthermore, in the gatifloxacin arm,
those infected with FQ-resistant organisms had significantly higher median FCTs than those infected with S. Typhi with alternative FQ susceptibility profiles (median FCTs (days): susceptible, 2.96 (IQR: 2.13-3.85), intermediate, 4.01 (IQR: 2.76-5.37) and resistant 8.2 (IQR: 5.99-10.5), respectively ($p$<0.0001)) (Figure 3.2.D, Table 3.6). Comparatively, the median FCT for those infected with an FQ resistant organism but randomised to ceftriaxone was 3.83 days (IQR: 2.96-4.7) ($p$<0.0001 for the between-treatment comparison) (Table 3.6).
Table 3.4 Summary of time to treatment failure and fever clearance time by *Salmonella* Typhi lineage

<table>
<thead>
<tr>
<th>Time to treatment failure</th>
<th>Gatifloxacin (events/N)</th>
<th>Ceftriaxone (events/N)</th>
<th>Hazard ratio of time to failure (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H58</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- H58</td>
<td>15/34</td>
<td>3/31</td>
<td>0.19 (0.05, 0.56); p=0.002</td>
<td>0.020</td>
</tr>
<tr>
<td>- Non-H58</td>
<td>0/6</td>
<td>2/7</td>
<td>3.87 (0.31, 534.24); p=0.32</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fever clearance time</th>
<th>Gatifloxacin median (IQR) days</th>
<th>Ceftriaxone median (IQR) days</th>
<th>Acceleration factor (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H58</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- H58</td>
<td>5.03 (3.18, 7.21)</td>
<td>3.07 (1.89, 4.52)</td>
<td>1.59 (1.22, 2.09); p=0.0006</td>
<td>0.07</td>
</tr>
<tr>
<td>- Non-H58</td>
<td>2.87 (2.08, 3.7)</td>
<td>3.12 (2.2, 4.12)</td>
<td>0.90 (0.59, 1.36); p=0.61</td>
<td></td>
</tr>
</tbody>
</table>

*Likelihood ratio test p=0.06 and 0.40 for comparison of time to treatment failure between H58 vs. non-H58 groups in gatifloxacin arm only and in all patients, respectively.

¥p=0.013 and p=0.029 for comparison of interval censored time to fever clearance between H58 vs. non-H58 groups in gatifloxacin arm only and in all patients, respectively.
### Table 3.5 Treatment failure in detail by *Salmonella* Typhi lineage in the gatifloxacin treatment group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Non-H58 (N=6)</th>
<th>H58 (N=34)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment failure</td>
<td>0 (0%)</td>
<td>15 (44.1%)</td>
<td>0.06</td>
</tr>
<tr>
<td>- Fever past seven days</td>
<td>0 (0%)</td>
<td>7 (21.2%)</td>
<td></td>
</tr>
<tr>
<td>- Rescue treatment required</td>
<td>0 (0%)</td>
<td>9 (27.3%)</td>
<td></td>
</tr>
<tr>
<td>- Microbiological failure</td>
<td>0 (0%)</td>
<td>2 (6.5%)</td>
<td></td>
</tr>
<tr>
<td>- Relapse within 28 days</td>
<td>0 (0%)</td>
<td>4 (14.8%)</td>
<td></td>
</tr>
<tr>
<td>- Confirmed relapse within 28 days</td>
<td>0 (0%)</td>
<td>4 (14.8%)</td>
<td></td>
</tr>
<tr>
<td>Any relapse during 6 month follow-up</td>
<td>1 (16.7%)</td>
<td>4 (11.8%)</td>
<td>0.62</td>
</tr>
<tr>
<td>Confirmed relapse during 6 month</td>
<td>1 (16.7%)</td>
<td>4 (11.8%)</td>
<td></td>
</tr>
</tbody>
</table>

*Time to treatment failure and time to relapse were analyzed using Firth’s penalized maximum likelihood bias reduction method for Cox regression and p values were calculated from likelihood ratio tests*
Table 3.6 Summary of time to treatment failure and fever clearance time by ciprofloxacin susceptibility

<table>
<thead>
<tr>
<th>Ciprofloxacin susceptibility group</th>
<th>Time to treatment failure (events/N)</th>
<th>Ceftriaxone (events/N)</th>
<th>Hazard ratio of time to failure (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>†</td>
<td></td>
<td></td>
<td></td>
<td>0.08</td>
</tr>
<tr>
<td>- Susceptible</td>
<td>0/5</td>
<td>1/5</td>
<td>2.40 (0.13, 350.21); p=0.57</td>
<td></td>
</tr>
<tr>
<td>- Intermediate</td>
<td>7/25</td>
<td>2/27</td>
<td>0.27 (0.05, 0.99); p=0.049</td>
<td></td>
</tr>
<tr>
<td>- Resistant</td>
<td>8/10</td>
<td>2/6</td>
<td>0.27 (0.05, 1.01); p=0.052</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ciprofloxacin susceptibility group</th>
<th>Fever clearance time (median (IQR) days)</th>
<th>Ceftriaxone median (IQR) days</th>
<th>Acceleration factor (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>‡ p&lt;0.0001 for comparison of interval censored time to fever clearance between MIC groups in gatifloxacin arm only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Susceptible</td>
<td>2.96 (2.13, 3.85)</td>
<td>4.78 (4.01, 5.5)</td>
<td>0.71 (0.49, 1.02); p=0.07</td>
<td></td>
</tr>
<tr>
<td>- Intermediate</td>
<td>4.01 (2.76, 5.37)</td>
<td>2.63 (1.52, 4.05)</td>
<td>1.31 (0.97, 1.76); p=0.07</td>
<td></td>
</tr>
<tr>
<td>- Resistant</td>
<td>8.2 (5.99, 10.5)</td>
<td>3.83 (2.96, 4.7)</td>
<td>2.23 (1.57, 3.17); p&lt;0.0001</td>
<td></td>
</tr>
</tbody>
</table>

\[^\text{†}\text{Likelihood ratio test } p=0.007\text{ for comparison of time to treatment failure between MIC groups in gatifloxacin arm only}\
\[^\text{‡}p<0.0001\text{ for comparison of interval censored time to fever clearance between MIC groups in gatifloxacin arm only}\]
Table 3.7 Treatment failure in detail by ciprofloxacin susceptibility in the gatifloxacin treatment group

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Susceptible (N=5)</th>
<th>Intermediate (N=25)</th>
<th>Resistant (N=10)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment failure</td>
<td>0 (0%)</td>
<td>7 (28.0%)</td>
<td>8 (80%)</td>
<td>0.007</td>
</tr>
<tr>
<td>- Fever past seven days</td>
<td>0 (0%)</td>
<td>2 (8.0%)</td>
<td>5 (55.6%)</td>
<td></td>
</tr>
<tr>
<td>- Rescue treatment required</td>
<td>0 (0%)</td>
<td>3 (12.0%)</td>
<td>6 (66.7%)</td>
<td></td>
</tr>
<tr>
<td>- Microbiological failure</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>2 (22.2%)</td>
<td></td>
</tr>
<tr>
<td>- Relapse within 28 days</td>
<td>0 (0%)</td>
<td>2 (9.5%)</td>
<td>2 (28.6%)</td>
<td></td>
</tr>
<tr>
<td>- Confirmed relapse within 28 days</td>
<td>0 (0%)</td>
<td>2 (9.5%)</td>
<td>2 (28.6%)</td>
<td></td>
</tr>
<tr>
<td>Any relapse during 6 month follow-up</td>
<td>1 (20%)</td>
<td>2 (8.0%)</td>
<td>2 (20%)</td>
<td>0.31</td>
</tr>
<tr>
<td>Confirmed relapse during 6 month</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>follow-up</td>
<td>1 (20%)</td>
<td>2 (8.0%)</td>
<td>2 (20%)</td>
<td></td>
</tr>
</tbody>
</table>

*Time to treatment failure and time to relapse were analyzed using Firth’s penalized maximum likelihood bias reduction method for Cox regression and p values were calculated from likelihood ratio tests
Figure 3.2 The association of *Salmonella* Typhi lineage and ciprofloxacin susceptibility with treatment failure and fever clearance time in patients randomised to gatifloxacin

A) Kaplan-Meier curve for time to treatment failure by H58 and non-H58 *Salmonella* Typhi. B) Kaplan-Meier curve for time to treatment failure by *Salmonella* Typhi susceptibility group (susceptible, intermediate, resistant) against ciprofloxacin. C) Non-parametric maximum likelihood estimators for interval-censored fever clearance time (see methods) by H58 and non-H58 *Salmonella* Typhi. D) Non-parametric maximum likelihood estimators for interval-censored fever clearance time by *Salmonella* Typhi susceptibility group (susceptible, intermediate, resistant against ciprofloxacin).
3.2.4 The emergence of fluoroquinolone-resistant *Salmonella Typhi*

Given the unusual long branch leading to the H58 triple mutant subclade on the phylogenetic tree (Figure 3.1), I hypothesised that the H58 triple mutants represented a contemporary importation into Nepal. To explore this, I compared the genomes of the 78 RCT *S. Typhi* isolates with those from 58 supplementary *S. Typhi* isolates from previous studies conducted between 2008 and 2013 in this setting (Figure 3.3, Appendix A). I found that the majority of the local H58 isolates (84/121; 69.4%) were closely related; these organisms represented “endemic” Nepali H58 clade containing a single S83F *gyrA* mutation. Additionally, I identified a further five Nepali organism isolated in 2013 that belonged to the H58 triple mutant group, and had an MIC ≥24 μg/ml against ciprofloxacin. Incorporating additional genome sequences from a recent international study of the H58 lineage, I found that all the Nepali H58 triple mutants were closely related (5 SNPs to nearest neighbour) to H58 triple mutants isolated in India between 2008 and 2012 (Figure 3.3).
Figure 3.3 The phylogenetic structure of fluoroquinolone resistant *Salmonella* Typhi H58 in a regional context

Maximum likelihood phylogeny based on core-genome SNPs of 136 (78 from the RCT) *Salmonella* Typhi isolates from Nepal and neighbouring India (Appendix A). Main tree shows the overall phylogenetic structure and the presence of specific combinations of mutations in *gyrA* (S83F, D87V and D87N), *parC* (S80I) and *parE* (A364V). The inset shows a magnified view of the fluoroquinolone resistant *Salmonella* Typhi H58 triple mutants from Nepal and their close association with similarly fluoroquinolone resistant *Salmonella* Typhi H58 triple mutants from India. The scale bar on the primary tree indicates the number of substitutions per variable site, while that in the inset indicates genetic distance in number of SNPs. Asterisks indicate ≥85% bootstrap support at nodes of interest
3.3 Discussion

My study shows that a new FQ resistant subclade of H58 S. Typhi has been introduced into Nepal and is associated with a lack of FQ efficacy. This subclade was associated with longer FCTs and treatment failure in patients treated with the FQ, gatifloxacin. For the first time I can conclusively show how enteric fever patients respond to FQ treatment when infected with a specific subclade of H58, thereby linking organism genotype with a treatment phenotype. Given the international significance of FQs for the treatment of enteric fever and other bacterial infections, my findings have major global health implications for the use of this group of antimicrobials.

A single S. Typhi organism with the same combination of triple mutations was isolated in Nepal in 2011. This isolate was also associated with treatment failure, although, this organism was not genome sequenced and was assumed to be an isolated case. More significantly, several closely related bacteria were genome sequenced during an international study of H58 S. Typhi. These organisms had the same combination of triple FQ resistance mutations as those described here; my analysis shows they belong to the same subclade of H58. These isolates had equivalently high MICs against ciprofloxacin and were isolated in India between 2008 and 2012. However, there were no associated outcome data for these strains and other reports from India have been limited. My data strongly suggests this lineage was recently introduced into Nepal from India (or nearby) and has since entered in an endemic
transmission cycle in Kathmandu. Given the large extent of human movement between India and Nepal, I propose this is a likely route of introduction.

Appropriate antimicrobial therapy is critical for enteric fever, as effective drugs curtail symptoms and prevent life threatening complications. My data has major repercussions for enteric fever treatment, and I advocate that FQs should no longer be used for empirical enteric fever therapy on the Indian subcontinent, as I predict these organisms are now widespread and may be associated with poor outcomes. Notably, one of the arms in the RCT used the newer generation FQ, gatifloxacin, which binds to a different location on the DNA gyrase than the older FQs and is not as susceptible to the common resistance mutations\textsuperscript{292}. The isolates in this study were not generally resistant to gatifloxacin according to the current CLSI guidelines\textsuperscript{293}; I suggest that these guidelines be modified to reflect these new clinical data. I additionally propose that S. Typhi genotyping, mapping and susceptibility testing is performed routinely and rapidly in reference laboratories outside South Asia to monitor the international spread of this lineage and to ensure the provision of alternative efficacious therapies to returning travellers\textsuperscript{294,295}. In cases of infection with these FQ-resistant isolates, I suggest that ceftriaxone and azithromycin are used as alternatives, and do not currently recommend a return to the use of first-line drugs without contemporary data on treatment outcome. Whilst none of the isolates in this study were MDR, a rapid recrudescence of MDR strains may occur if we return to older first-line alternatives.

This study has limitations. First, the clinical data was collected from one study in a single location, thus limiting utility outside this setting. Second, the overall sample
size (and the gatifloxacin group subsampling) of those with culture-positive *S*. Typhi-associated enteric fever was relatively small. Despite these limitations, I was able to show a highly significant association between disease outcome and susceptibility profile of the infecting organism. Further, by using WGS I was able to pinpoint causative mutations, identify the subclade responsible for treatment failure and relate these strains to other isolates circulating outside Nepal. The methodologies presented here, in which clinical outcome data is combined with genome sequences and antimicrobial susceptibility data, should become the gold standard for informing empiric treatment for enteric fever and understanding the role of bacterial genotype and resistance profile on disease outcome for other bacterial infections.

In conclusion, my data show a significant association between *S*. Typhi genotype, antimicrobial susceptibility and disease outcome for those treated with gatifloxacin in a cohort of Nepali enteric fever patients. A FQ-resistant variant of Typhi H58 has emerged in Nepal and is associated with the clinical failure of FQs. My data suggest these isolates are likely widespread in the subcontinent and FQs should not be recommended for empirical enteric fever therapy in this setting.
Chapter 4

The phylogenetics and spatiotemporal dynamics of *Salmonella Paratyphi* A in Kathmandu, Nepal

4.1 Introduction

Even though typhoid fever has been studied for more than one hundred years, it still causes significant disease burden in South and Southeast Asian countries, specifically in locations with limited access to clean water and improved sanitation. In the last decade, there have been major advances in typhoid research, which have provided a better understanding of the disease epidemiology and risk factors associated with typhoid fever as well as underlining the emergence and global dissemination of antimicrobial resistant organisms \(^{138,248}\). These studies have had a major impact on the disease management and public health interventions. However, recent advances and most typhoid research have been focused mainly on *S. Typhi*, as this pathogen is considered to be more important than *S. Paratyphi* A, particularly with respect to disease magnitude and severity \(^{296}\). This gap in knowledge and a lack of public attention have rendered *S. Paratyphi* A a neglected pathogen for many years. Since the turn of the 21\(^{st}\) century, *S. Paratyphi* A has emerged at an unprecedented rate in many Asian countries including India, Pakistan, Nepal, Bangladesh, China, and Indonesia, becoming a increasingly important causative agent of typhoid (enteric) fever \(^{139}\). Clinical and epidemiological investigations have revealed that typhoid fever caused by *S. Paratyphi* A is clinically indistinguishable from *S. Typhi*; furthermore, *S.
Paratyphi A also displays dissimilar epidemiological features to S. Typhi infection, with infection more likely to arise outside the household\textsuperscript{297,142}. These studies highlight an underestimation of the clinical relevance of S. Paratyphi A infection and questions if the current control and preventive measures for typhoid would also protect against this pathogen. Currently, there is no vaccine against S. Paratyphi A and the licensed typhoid (S. Typhi) vaccines do not provide cross protection against S. Paratyphi A and are likely to be less effective for controlling typhoid fever in areas where these pathogens are co-circulating.

In comparison to S. Typhi, there have been growing evidence suggesting that S. Paratyphi A is more likely to develop resistance to nalidixic acid, an indicator of reduced susceptibility to fluoroquinolones (the first line drug recommended by WHO) and they are generally less susceptible to various classes of antimicrobial frequently used as alternative treatments\textsuperscript{298}. S. Paratyphi A cases associated with organisms that are fully resistant to ciprofloxacin (MIC of 8 \textmu g/ml), ceftriazone (MIC >16 \textmu g/ml) and azithromycin (MIC >64 \textmu g/ml) are also not uncommon\textsuperscript{299,221,225}. In endemic areas where there is a lack of safe drinking water and inadequate food hygiene, case detection and antimicrobial therapy become crucial for the management of typhoid fever. However, limited access to laboratory facilities, low sensitivity and specificity of current diagnostic methods and similar clinical manifestations caused by S. Paratyphi A and S. Typhi have imposed significant challenges for the diagnostic accuracy as well as the appropriate antimicrobial therapy\textsuperscript{300}. As a result, the actual burden of S. Paratyphi A infection remains unknown in many settings and empiric antimicrobial treatment may be problematic. At present, available epidemiological
data of *S. Paratyphi A* is very limited and remains unclear; moreover, conventional epidemiological approaches only describe the demographic and epidemiological features of reported cases without any characterization of the bacterial population structure and dynamics. Understanding the structure and changing dynamics of the pathogen population in time and space, as well as the driving forces governing these in an epidemiological context, is essential for providing compelling evidence and novel insights into the source of infection and the patterns of typhoid transmission \(^{301}\).

The fact that *S. Paratyphi A* isolates exhibit little genetic diversity makes whole-genome sequencing technology a fundamental approach for understanding the population of this organism. However, this approach remains a big challenge in many endemic settings.

Kathmandu, Nepal, is a highly endemic location for typhoid fever, and *S. Paratyphi A* infections have increased in recent years. During a period of sustained blood culture surveillance for typhoid fever at Patan hospital in Kathmandu between 1993 and 2003 the proportion of typhoid fever caused by *S. Paratyphi A* increased annually with the highest isolation rate up to 34% of all cases \(^{298}\). From 2005-2009, the isolation rate of *S. Paratyphi A* remained high and accounted for 31.5% of total confirmed typhoid cases; moreover, for some months, the number of *S. Paratyphi A* and *S. Typhi* infections was comparable \(^{302}\). Despite the overall decline of typhoid fever during this study period, *S. Paratyphi A* emerged rapidly and coincided with ciprofloxacin non-susceptibility. A previous case-control study demonstrated that *S. Paratyphi A* infection rate increased with age and was higher among those consuming street food,
suggesting that direct contact and consumption of contaminated food may play a significant role in the transmission of S. Paratyphi A in this region 129.

Approximately 1-5% of typhoid patients become chronic carriers and these people can act as reservoirs for the persistence and transmission of typhoid fever, especially if they are food handlers. Chronic carriage of S. Paratyphi A has been reported, and the prevalence was 1.6% among those undergoing cholecystectomy in Nepal 130. This is comparable to the chronic carrier rate of S. Typhi, even though the number of acutely infected S. Paratyphi A patients was lower during the same period 130. Little is known about genetic relatedness between acute and chronic S. Paratyphi A isolates as well as the relative role of chronic carriage in the transmission cycle of this pathogen. Additionally, all previous epidemiological investigations of S. Paratyphi A infections in Nepal have not included a detailed characterization of bacterial population and, as a result, have overlooked disease transmission dynamics. Here, I have, for the first time, performed a detailed molecular epidemiological study, in which I utilized whole genome sequencing to assess the phylogenetics and populations structure of S. Paratyphi A and combined this highly resolved genetic information with individual GPS location to interrogate the spatiotemporal dynamics of S. Paratyphi A infections in Nepal. I additionally reconstructed a Nepalese S. Paratyphi A phylogeny in a global context to further investigate into the dynamics of this bacterial population. This was also the first study characterizing the genetic relationship between carrier and acute S. Paratyphi A isolates to provide new insights into the role of typhoid carriage in this setting.
4.2 Results

4.2.1 Baseline characteristics

Between 2005 and 2014, 223 *S. Paratyphi* A isolates were collected and whole-genome sequenced, of which 92% (206/223) were from acutely infected patients and the remainder were from the gallbladder of carriers (8%, 17/223). The median age of acute *S. Paratyphi* A patients was 19 years (IQR: 13-24 years), while the median age of asymptotically infected individuals was 37 years (IQR: 32-44 years) \((p<0.001, \text{Mann-Whitney U test})\). Over the observed time period, the median age of acutely infected patients did not vary substantially \((p=0.120, \text{Mann-Whitney U test})\). 71\% (146/206) of the acute *S. Paratyphi* A patients were male, this contrasted with asymptotically infected individuals, of which only 25\% (4/16) were male \((p<0.001, \text{Fisher’s exact test})\). There were no major differences in sex distribution between patients enrolled in acute studies \((p=0.251, \text{Fisher’s exact test})\). The spatial distribution demonstrated that most *S. Paratyphi* A patients lived in the central Kathmandu Valley, specifically within Lalitpur; the carriers were more likely to live further away from the centre of the city (Figure 4.1). This is likely due to a lack of cholecystectomy facilities available outside of central Kathmandu.

A clinical history was collected for all acutely infected patients. The majority of patients reported a history of headache (93%) and anorexia (67%). Diarrhoea (13%) and constipation (12%) were rare. There were also no major differences between years were present for any symptom. Of 162 *S. Paratyphi* A patients with recorded water use information, only about half (52\%, 85/162) reported using the municipal supply water as their main source. Other water sources included stone spout (14\%,
23/162), sunken well (10%, 17/162), bottled water (12%, 19/162). Over a third (35%, 57/161) of patients reported using untreated water for drinking, with an additional 35% (57/161) reporting the use of a filter and 14% (23/161) reporting boiling as their primary treatment method. Additionally, a total of 52% (12/23) of patients reporting stone spout use did not treat the water prior to drinking, compared to 24% (20/85) of those using municipal supply water.
Figure 4.1 Locations of *Salmonella* Paratyphi A isolates in Kathmandu

Map shows the locations of *S. Paratyphi A* in Kathmandu between 2005 and 2014 with carrier isolates shown in red and acute isolates in blue. Note: 4 additional carrier isolates are not shown on the map due to the fact that they are 70km from central Kathmandu.
4.2.2 Antimicrobial resistance

MICs of all S. Paratyphi A isolates were determined as described in section 2.2.3. Overall, S. Paratyphi A isolates were susceptible to all first-line antimicrobials including chloramphenicol, amoxicillin, and cotrimoxazole. However, 93% (208/223) of these isolates displayed resistance to nalidixic acid. Additionally, of those tested, 8% (14/178) were resistant to ciprofloxacin (MIC ≥1 μg/ml) and 39% (69/178) were resistant to ofloxacin (MIC ≥ 2 μg/ml) according to the updated CLSI breakpoints. None were resistant to ceftriaxone and gatifloxacin. As shown in Figure 4.2, MICs against ciprofloxacin, ofloxacin and gatifloxacin did not change significantly over time. In contrast, MICs to other first-line antimicrobials remained largely constant over the study period. MICs to azithromycin declined between 2006-2010 (Pearson correlation coefficient = -0.387, p<0.001). Additionally, the median MIC to azithromycin was significantly higher in organisms from acutely infected patients compared to carriers (p=0.004, Mann-Whitney U test). Nalidixic acid resistance was also more common in S. Paratyphi A isolates from acute patients (96%, 198/206) than the asymptomatic individuals (71%, 12/17) (Chi-square test, p<0.001). There were no significant differences in median MICs to ciprofloxacin and ofloxacin between these two groups. An in silico resistome analysis also found that none of S. Paratyphi A isolates in Nepal carried any plasmid or gene cassette associated with antimicrobial resistance.
Figure 4.2 Minimal inhibitory concentrations (MICs) of S. Paratyphi A isolates to various antimicrobials over time
4.2.3 The population structure and dynamics of *S. Paratyphi A* isolates in Nepal

To investigate the population structure and dynamics of *S. Paratyphi A* in Nepal, I reconstructed the phylogeny of all Nepalese *S. Paratyphi A* isolates along with 111 genomes from a global collection. These data showed that the majority of Nepalese isolates (94.2%, 210/223) fell within global lineage A, which contained organisms circulating mostly within South Asian countries such as India, Nepal, and Pakistan (Figure 4.3). A small proportion of Nepalese *S. Paratyphi A* (5.4%, 12/223) also belonged to global lineage C. Unlike lineage A, lineage C has successfully spread globally and contains organisms from a wide range of geographical areas including South Asia (India, Nepal, Pakistan, Sri Lanka), Southeast Asia (Cambodia, Vietnam, Thailand, Indonesia), Middle East (Turkey), East Asia (China) and Africa (Ghana, Chad, Senegal, Guinea, Mali, Morocco, Egypt). Only 1/223 (0.4%) *S. Paratyphi A* isolate in Nepal grouped within global lineage F, which included mostly historic isolates.

Noticeably, between 2005 and 2011 I observed the phenomenon of clonal expansion of nalidixic resistant Nepalese *S. Paratyphi A* within global lineage A, which was designated as sub-lineage A1. This sub-lineage constituted 64% (135/210) of Nepalese isolates belonging to lineage A; 98.5% (133/135) had a mutation in codon 83 of *gyrA*, changing serine to phenylalanine. These organisms were resistance to nalidixic acid. From the global collection, there were one Indian and one Nepalese isolates recovered in 1999 and 2000, respectively also within sub-lineage A1; however, these two isolates were susceptible to nalidixic acid. As a result, this resistance associated mutation in *gyrA* likely developed and subsequently become
fixed in the bacterial population between 2000 and 2005. All acute S. Paratyphi A organisms isolated after 2005 had this mutation, suggesting a strong selective advantage for nalidixic acid resistance. Additionally, in comparison to the reference genome AKU_12601 (lineage C), isolates from sub-lineage A1 did not carry the SPA-2 sopE phage (SSPA2377 to SSPA2423) and the prophage region SPA-3 (SSPA2424 to SSPA2446). Alternatively, these organisms had acquired a novel intact prophage inserted in the chromosome between CDSs SSPA3930 and SSPA3931. This prophage was 33.8 kb in size and mostly closely related to phage P88 (NC_026014) (Figure 4.4). This novel prophage carried approximately 50 proteins, including a variant of the sopE gene, which exhibited 220/240, 222/240 and 221/240 amino acid identity to the sopE gene present in SPA-2 phage of the S. Paratyphi A reference strain AUK_12601 (FM200053); prophage of S. Typhimurium SL1344 (NC_016810.1) and SPI-7 in S. Typhi CT18 (AL513382), respectively. SopE is an effector protein secreted via a type III secretion system (SPI-1) and plays an important role in bacterial invasion into non-phagocytic cells by activating the actin cytoskeleton rearrangements and stimulating membrane ruffling. SopE-expressing S. Typhimurium has been shown to increase virulence and organisms carrying this additional gene have been associates with several epidemics. It is unclear whether this novel sopE prophage modulates the virulence of these S. Paratyphi A. In addition to gaining resistance to nalidixic acid and a novel sopE prophage, there were also a number of non-synonymous mutations and gene disruptions associated with sub-lineage A1 (Table 4.1).
Figure 4.3 Phylogenetics of Nepalese S. Paratyphi A isolates in a global context

Maximum likelihood phylogenetic tree of 223 Nepalese S. Paratyphi A isolates together with 111 global S. Paratyphi A isolates and their corresponding metadata. The tree consisted of six major lineages (A-F) corresponding to different colored branches. The terminal nodes exhibit the country of isolates with isolates originating from Nepal and neighboring countries highlighted. The bars show the year of isolation, two major sublineages A1 (red) and A2 (blue) within lineage A, the division of six major lineages and the presence of gyrA mutations. The scale bar indicates the number of substitutions per variable site.
Figure 4.4 Novel sopE prophage of sub-lineage A1
Table 4.1 Nonsynonymous mutations and indels associated with sub-lineage A1

<table>
<thead>
<tr>
<th>Position in FM200053</th>
<th>CDS</th>
<th>Product</th>
<th>Mutation</th>
<th>Ref</th>
<th>Alt</th>
</tr>
</thead>
<tbody>
<tr>
<td>253654</td>
<td>SSPA0212</td>
<td>2,3,4,5-tetrahydroxypyridine-2-carboxylate N-succinyltransferase</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>860533</td>
<td>SSPA0723</td>
<td>dTDP-glucose 4,6-dehydratase</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>2753956</td>
<td>SSPA2470</td>
<td>4-aminobutyrate aminotransferase</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>1346343</td>
<td>SSPA1187</td>
<td>putative NADP-dependent oxidoreductase</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>1690227</td>
<td>SSPA1517</td>
<td>ABC transporter ATP-binding subunit</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>1881695</td>
<td>SSPA1689</td>
<td>30S ribosomal protein S1</td>
<td>Nonsynonymous</td>
<td>T</td>
<td>A</td>
</tr>
<tr>
<td>2899096</td>
<td>SSPA2613</td>
<td>sulfite reductase (NADPH) flavoprotein beta subunit</td>
<td>Nonsynonymous</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>3135049</td>
<td>SSPA2833</td>
<td>conserved hypothetical protein</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>4441670</td>
<td>SSPA3972</td>
<td>putative inner membrane protein</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>73777</td>
<td>SSPA0061</td>
<td>CitG protein</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>482924</td>
<td>SSPA0385</td>
<td>ethanolamine ammonia-lyase light chain</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>2311942</td>
<td>SSPA2069</td>
<td>DNA polymerase III epsilon subunit</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>2600568</td>
<td>SSPA2336</td>
<td>putative sigma-54-dependent transcriptional regulator</td>
<td>Nonsynonymous</td>
<td>G</td>
<td>A</td>
</tr>
<tr>
<td>2794835</td>
<td>SSPA2511</td>
<td>putative lipoprotein</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>4023577</td>
<td>SSPA3609</td>
<td>anaerobic C4-dicarboxylate transporter</td>
<td>Nonsynonymous</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>4332842</td>
<td>SSPA3866</td>
<td>putative amino acid permease</td>
<td>Nonsynonymous</td>
<td>C</td>
<td>A</td>
</tr>
<tr>
<td>679384</td>
<td>SSPA0561</td>
<td>outer membrane protein</td>
<td>Nonsynonymous</td>
<td>A</td>
<td>C</td>
</tr>
<tr>
<td>4569233</td>
<td>SSPA4083</td>
<td>putative two-component response regulator (creB)</td>
<td>Nonsense</td>
<td>C</td>
<td>T</td>
</tr>
<tr>
<td>3631814</td>
<td>SSPA3280</td>
<td>xylose operon regulatory protein</td>
<td>Insertion</td>
<td>GT</td>
<td>GGT</td>
</tr>
<tr>
<td>2809585</td>
<td>SSPA2524</td>
<td>formate hydrogenlyase subunit 2</td>
<td>Deletion</td>
<td>CGGGCGGCGGC</td>
<td>C</td>
</tr>
</tbody>
</table>

Nonsynonymous mutations and indels in sub-lineage A1 were identified relative to the inferred most recent common ancestor of sub-lineage A1 and sub-lineage A2.
The genomics data suggested that, after clonal expansion, sub-lineage A1 gradually began to disappear from the population after 2011 and was replaced by the distantly related sub-lineage, A2 (Figure 4.3). The main branch of sub-lineage A2 differed by 96 SNPs from sub-lineage A1 and contained 67/210 (32%) of the Nepalese isolates belonging to lineage A. Isolates belonging to sub-lineage A2 was first detected in 2008 and initiated a clonal expansion, which replaced all other lineages from 2011. Sub-lineage A2 became the most dominant genotype accounting for 83-96% of isolates annually between 2012 and 2014. The median pairwise SNP distance of the Nepalese isolates within sub-lineage A2 was only 2 SNPs (IQR: 1-4 SNPs), which was lower than the median pairwise SNP distance of isolates belonging to sub-lineage A1 (5 SNPs, IQR: 4-7 SNPs), indicating a very recent clonal expansion of sub-lineage A2. All Nepalese isolates in sub-lineage A2 also had a Ser83Phe mutation in gyrA and were resistant to nalidixic acid. In comparison to the reference strain AKU_12601, isolates within sub-lineage A2 carried a comparable complement of prophage regions and there was no apparent difference in gene content between the first A2 isolate in 2008 and the remaining isolates in later years of this sub-lineage. As shown in Table 4.2, there were various non-synonymous mutations and gene disruptions associated with sub-lineage A2. Gene disruptions included SSPA0470, which encodes for a conserved hypothetical across Salmonella enterica and SSPA2905 (tdcD), encoding for propionate kinase. The Tdc operon (tdcABCDEFG) consists of seven genes which encode enzymes essential for the degradation of short-chain fatty acids, which serve as an alternative carbon and energy source in the absence of a preferred nutrient in human gut. Additionally, a number of non-synonymous mutations were located in genes involved in bacterial virulence, LPS and
gene regulation, such as secreted protein SifA, an effector protein of the SPI-2 Type 3 secretion system, which plays an important role in Salmonella virulence \(^{304}\), WaaP (SSPA3336) involved in lipopolysaccharide core biosynthesis, signal peptidase I (LepAB), sensor protein KdpD in the two-component KdpD/KdpE regulatory system \(^{305}\). Details regarding all nonsynonymous mutations and indels associated with sub-lineage A2 are shown in Table 4.2.
<table>
<thead>
<tr>
<th>Position in FM200053</th>
<th>CDS</th>
<th>Product</th>
<th>Mutation</th>
<th>Ref</th>
<th>Alt</th>
</tr>
</thead>
<tbody>
<tr>
<td>2176873</td>
<td>SSPA1948</td>
<td>rare lipoprotein A precursor</td>
<td>Nonsynonymous A G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2412607</td>
<td>SSPA2166</td>
<td>proline-specific permease ProY</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>475620</td>
<td>SSPA0379</td>
<td>putative aldehyde dehydrogenase</td>
<td>Nonsynonymous T C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1671505</td>
<td>SSPA1500</td>
<td>putative MutT-family protein</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3035334</td>
<td>SSPA2730</td>
<td>conserved hypothetical protein</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978639</td>
<td>SSPA1777</td>
<td>hypothetical ABC transporter ATP-binding protein</td>
<td>Nonsynonymous A G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3156617</td>
<td>SSPA2852</td>
<td>possible lipoprotein</td>
<td>Nonsynonymous T A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4160748</td>
<td>SSPA3728</td>
<td>phosphoribosylaminomimidazolecarboxamide formyltransferase and IMP cyclohydrolase (bifunctional enzyme)</td>
<td>Nonsynonymous C A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>320372</td>
<td>SSPA0268</td>
<td>signal peptidase I</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2118532</td>
<td>SSPA1899</td>
<td>membrane-associated protein kinase KdpD</td>
<td>Nonsynonymous G C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2118529</td>
<td>SSPA1899</td>
<td>membrane-associated protein kinase KdpD</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1035960</td>
<td>SSPA0902</td>
<td>aspartyl-tRNA synthetase</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1798103</td>
<td>SSPA1620</td>
<td>putative 4-hydroxyphenylacetate permease</td>
<td>Nonsynonymous T C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2936045</td>
<td>SSPA2640</td>
<td>L-serine dehydratase 2 (L-serine deaminase 2)</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>825360</td>
<td>SSPA0698</td>
<td>conserved hypothetical protein</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3321069</td>
<td>SSPA3005</td>
<td>oxaloacetate decarboxylase beta chain</td>
<td>Nonsynonymous C G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3204223</td>
<td>SSPA2895</td>
<td>putative membrane protein</td>
<td>Nonsynonymous T C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>999519</td>
<td>SSPA0865</td>
<td>putative glucose-6-phosphate dehydrogenase</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1684077</td>
<td>SSPA1511</td>
<td>putative virulence determinant</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3109130</td>
<td>SSPA2806</td>
<td>possible transferase</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1133549</td>
<td>SSPA1000</td>
<td>alanine racemase</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1106787</td>
<td>SSPA0973</td>
<td>PTS system, mannose-specific IIAB component</td>
<td>Nonsynonymous T C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3289612</td>
<td>SSPA2977</td>
<td>nitrogen regulatory IIA protein</td>
<td>Nonsynonymous A G</td>
<td></td>
<td></td>
</tr>
<tr>
<td>279714</td>
<td>SSPA0232</td>
<td>putative secreted chitinase</td>
<td>Nonsynonymous A T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1216127</td>
<td>SSPA1073</td>
<td>anthranilate synthase component I</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3210225</td>
<td>SSPA2900a</td>
<td>conserved hypothetical transport protein (pseudogene)</td>
<td>Nonsynonymous G A</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3700432</td>
<td>SSPA3336</td>
<td>lipopolysaccharide core biosynthesis protein</td>
<td>Nonsynonymous C T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>578579</td>
<td>SSPA0470</td>
<td>conserved hypothetical protein</td>
<td>Nonsynonymous G T</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3215599</td>
<td>SSPA2905</td>
<td>propionate kinase tdcD</td>
<td>Deletion AT T</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nonsynonymous mutations and indels in sub-lineage A2 were identified relative to the inferred most recent common ancestor of sub-lineage A1 and sub-lineage A2.
4.2.4 Genetic relatedness between acute and carriage S. Paratyphi A isolates in Nepal

The majority of S. Paratyphi A carriage isolates belonged to sub-lineage A1 (53%, 9/17) and clustered within the acute isolates, of which 8/9 were resistant to nalidixic acid and had the same mutation Ser83Phe in gyrA. These carriage isolates probably resulted from recent asymptomatic infections for the following reasons. First, as mentioned above, the resistance mutation Ser83Phe in gyrA likely occurred and subsequently become fixed in the bacterial population after 2005. Second, these carriage isolates were collected between 2007 and 2009. There was one nalidixic acid susceptible isolate from a carrier (GB624) that was collected in 2009 and clustered within the global nalidixic acid susceptible isolates. The duration of asymptomatic carriage by this isolate is at least 5 years considering the establishment of nalidixic resistant sub-lineage A1 from 2005. The median pairwise SNP distance between acute isolates within this sub-lineage A1 was 5 SNPs (IQR: 3-7 SNPs), which was significantly lower than the median pairwise SNP distance between carrier isolates (7 SNPs, IQR: 5-7 SNPs) (Wilcoxon rank sum test, \( p = 0.003 \)). My data demonstrated that the genetic diversity within the carriage isolates was higher compared to acute isolates in sub-lineage A1 considering the number of acute isolates was much higher.

Additionally, there was one carrier isolate collected in 2009 (GB672) that clustered within sub-lineage A2. This isolate was resistant to nalidixic acid (mutation Ser83Phe) and was only 1 SNP different from the first acute isolate collected in 2008, therefore this organism likely originated from recent asymptomatic infection. Three other carriage isolates (GB193, GB640, and GB726) belonged to global lineage A,
but they were distantly related to Nepalese isolates and clustered tightly with Indian isolates. The remaining 4 carrier isolates belonged to global lineage C (3) and lineage F (1).

Overall, my data demonstrated that the carriage isolates belonged to a wide range of genotypes; the most common genotype was sub-lineage A1. Additionally, the genotypic distribution of carrier isolates reflected the general genetic structure in the bacteria population. I found sufficient evidence to suggest that half of carrier isolates likely originated from recently asymptomatic carriage. Further, the fact that carriage isolates within sub-lineage A1 displayed higher diversity than acute isolates suggests that the carriage isolates may have been exposed to distinctive evolutionary process to induce adaptation for survival and replication in the gall bladder. Generally, the bacterial population of S. Paratyphi A in Kathmandu was highly diverse and dynamic with evidence of clonal expansion followed by clonal replacement. Sub-lineage A2 replaced all other pre-existing lineages since 2011 suggesting that carrier isolates seemed to have limited role in maintaining the bacterial diversity or facilitating the transmission in this population.

Most carriage and acute Nepalese isolates within lineage C clustered in proximity to other South Asia global isolates. These isolates were mostly detected in 2005 and gradually disappeared from this population in Kathmandu after 2011, this may be associated with the expansion and replacement of other successful genotypes. The only carriage strain in lineage F was distantly related and highly distinctive from the rest of the Nepalese isolates.
4.2.5 Spatial and spatiotemporal distribution of S. Paratyphi A genotypes

4.2.5.1 Genotypic subgrouping

To investigate the short-term evolution and dynamics of bacterial transmission in this setting, I firstly reconstructed the phylogeny for Nepalese isolates using a maximum-likelihood method and further subdivided sub-lineage A1 and A2 into different monophyletic clusters (designated as A1.1-11 and A2.1) based on moderate-to-high supported bootstrap values (60-99%). Subsequently, a SNP alignment of organisms belonging to sub-lineage A1 and A2 was extracted and used to build a minimal spanning tree using the goeBURST algorithm. Subgroups identified using this method were identical to the monophyletic clusters within the maximum-likelihood phylogeny; the minimal spanning tree was displayed for better visualization (Figure 4.5). I then combined this genotype information with individual GPS location data to investigate the spatiotemporal distribution and specific demographic and clinical characteristics associated with the common genotypes.
Figure 4.5 Minimal spanning tree showing different clonal clusters within sub-lineages A1 and A2

Minimum spanning tree showing various clonal clusters within sub-lineage A1 (A1.1-A1.11) and sub-lineage A2 (A2.1). The different clonal clusters are color-coded for reference. The number on each of the branches indicates the number of SNPs between each cluster.
I identified a total of 16 different genotypes of *S. Paratyphi* A. Between 2005 and 2009, sub-clade A1 was dominant and accounted for 91-96% (2005-2008) and 67% (2008-2009) of all *S. Paratyphi* A, respectively. During this period, this sub-lineage diversified and formed 11 distinct clonal clusters (A1.1-A1.11). There were substantial changes in the dominant cluster over time as well as the formation and extinction of minor clusters (Figure 4.6). For example, cluster A1.2 was the most common genotype in 2005, accounting for 27.9% of all isolates collected this year; however, this genotype subsequently disappeared and was replaced by another common cluster (A1.11) in the following years. Overall, the diversity of sub-lineage A1 decreased with time, declining from 9 different clusters in 2005 to only 3 in 2009. Between 2011-2014 there were dramatic changes in the bacterial population structure with the sudden increase of sub-lineage A2 in 2011 (62.7%, 32/51), which was concurrent with the disappearance of clonal clusters within sub-lineage A1. During this period (2011-2014), sub-lineage A2 predominated and a clonal cluster (A2.1) was represented at an almost constant rate annually (22-33%). Lineage C was constantly detected over the whole study period but represented only a small proportion of organisms.

After comparing some characteristics between sub-lineage A1 and A2, I found that the median age of patients within sub-lineage A1 was significantly older (21 years (IQR: 15-25 years) versus 18 years (IQR: 11-23 years) \( p=0.0498 \), Kruskal-Wallis test) (Table 4.3). Clinical symptoms at presentation between these two sub-lineages were comparable except that organisms within sub-lineage A1 were more likely to be associated with abdominal pain \( p=0.044 \), Chi-squared test). Water use also differed
dramatically, with people infected with sub-lineage A2 more likely to report drinking bottled water (21%, 14/66) compared to those infected with sub-lineage A1 (5.7%, 5/88) \( (p=0.0037, \text{ Chi-squared test}) \). Additionally, a higher proportion of patients infected with sub-lineage A2 (60%, 40/66) organisms reported using municipal supply as the main water source compared to 47.7% (42/88) to those infected with sub-lineage A1 organisms \( (p=0.113, \text{ Chi-squared test}) \) and were also more likely to filter water (42.4%, 28/66) compared to sub-lineage A1 (28.7%, 25/87) \( (p=0.078, \text{ Chi-squared test}) \). Lastly, those infected with sub-lineage A2 organisms were significantly less likely to use well water as the main water source (3% compared to 15.9%, \( p=0.015, \text{ Fisher’s exact test} \)). There was no difference in the antimicrobial resistance profiles between sub-lineage A1 and A2.
Figure 4.6 Annual distribution of *S.* Paratyphi A genotypes in Nepal

The graph shows the annual percentage of each *S.* Paratyphi A genotype in Nepal between 2005 and 2014. Different genotypes are color-coded for reference.
Table 4.3 Characteristics of two main sub-lineages A1 and A2 of \textit{S. Paratyphi A} in Nepal

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Sub-lineage</th>
<th>p_value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (median(IQR))</td>
<td>A1: 21 (15-25)</td>
<td>A2: 18 (11-23)</td>
</tr>
<tr>
<td>Male sex</td>
<td>67% (89/135)</td>
<td>78% (52/67)</td>
</tr>
<tr>
<td>Anorexia</td>
<td>67.7 (86/127)</td>
<td>66.7 (44/66)</td>
</tr>
<tr>
<td>Nausea</td>
<td>37.8 (48/127)</td>
<td>36.4 (24/66)</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>37 (47/127)</td>
<td>22.7 (15/66)</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>12.6 (16/127)</td>
<td>12.1 (8/66)</td>
</tr>
<tr>
<td>Constipation</td>
<td>13.4 (17/127)</td>
<td>10.6 (7/66)</td>
</tr>
<tr>
<td>Headache</td>
<td>95.3 (121/127)</td>
<td>87.9 (58/66)</td>
</tr>
<tr>
<td>Water source</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Municipal supply</td>
<td>47.7 (42/88)</td>
<td>60.6 (40/66)</td>
</tr>
<tr>
<td>Bottled</td>
<td>5.7 (5/88)</td>
<td>21.2 (14/66)</td>
</tr>
<tr>
<td>Stone spout</td>
<td>14.8 (13/88)</td>
<td>10.6 (7/66)</td>
</tr>
<tr>
<td>Tanker</td>
<td>1.1 (1/88)</td>
<td>1.5 (1/66)</td>
</tr>
<tr>
<td>Well</td>
<td>15.9 (14/88)</td>
<td>3.0 (2/66)</td>
</tr>
<tr>
<td>Mix</td>
<td>14.8 (13/88)</td>
<td>3.0 (2/66)</td>
</tr>
<tr>
<td>Water treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boil</td>
<td>13.8 (12/87)</td>
<td>12.1 (8/66)</td>
</tr>
<tr>
<td>Chlorine</td>
<td>6.9 (6/87)</td>
<td>0 (0/66)</td>
</tr>
<tr>
<td>Filter</td>
<td>28.7 (25/87)</td>
<td>42.4 (28/66)</td>
</tr>
<tr>
<td>Mix</td>
<td>10.3 (9/87)</td>
<td>7.6 (5/66)</td>
</tr>
<tr>
<td>Other</td>
<td>4.6 (4/87)</td>
<td>0 (0/66)</td>
</tr>
<tr>
<td>Untreated</td>
<td>35.6 (31/87)</td>
<td>37.9 (25/66)</td>
</tr>
</tbody>
</table>
4.2.5.2 Spatial and spatiotemporal mapping of S. Paratyphi A genotypes

Overall, there were a number of genotypes that exhibited spatial clustering. For the organisms within sub-lineage A1, genotype A1.2 (11/17 cases) clustered within a radius of 747m (RR = 8.6, \( p = 0.004 \)), genotype A1.3 (9/10 cases) also exhibited significant clustering within a radius of 696m (RR = 37.7, \( p < 0.0001 \)). Genotypes A1.6 (3/3 cases) and A1.10 (3/5 cases) were associated with smaller clusters containing almost all identified patients with radii <150m (RR = \( \infty \), \( p < 0.001 \) and RR = 110.0, \( p = 0.002 \), respectively). For sub-lineage A2, I detected a cluster of genotype A2.1 (10/21 cases) in the south of Lalitpur with a radius of 769m (RR=12.6, \( p < 0.001 \)). I also investigated the remaining isolates within sub-lineage A2 and identified one cluster in the North of Lalitpur (14/46 cases) with a radius of 1.2 km (RR=3.76, \( p = 0.003 \)) and another smaller cluster (6/46 cases) with a radius of 157m (RR=5.2, \( p = 0.01 \)) nested within the cluster of genotype A1.3. The location and the magnitude of significant clusters of different genotypes over the entire period of study is shown in Figure 4.7.

I also found a number of genotypes that clustered spatiotemporally over the study period. Spatiotemporal clusters were far more common in the beginning of study period compared to later years. For example, I identified 4 distinct and significant spatiotemporal clusters of genotypes A1.2, A1.6, A1.8 and A1.10 in 2005; noticeably all clusters were concentrated between June and September. Similarly in 2006 there were two significant clusters of genotypes A1.3 and A1.11, which were also found during the same period between May and August. Sub-lineage A2 started to replace other genotypes and expand in the bacterial population from 2011 with one significant spatiotemporal cluster detected (genotype A2.1). This cluster incorporated the
majority of patients presenting between May and September in 2011. In later years, sub-lineage A2 disseminate throughout the geographical study area, but I identified a significant spatiotemporal cluster of sub-lineage A2 spanning 2011 to 2013 in the north of Lalitpur, suggesting this area was a hot spot of S. Paratyphi A infection in these years.

The combination of evidence from both spatial and spatiotemporal clustering analyses suggests that localized clusters of a variety of genotypes within sub-lineage A1 were common early on in the study period (2005-2006) and declined in favor of setting-wide dominance of sub-lineage A2 by 2011. Genotypes with overlapping spatial and spatiotemporal clusters (namely, A1.2, A1.3, A1.6, A1.0, A2.1) showed the strongest evidence of concentrated, localized outbreaks synced in time and space. Genotypes with only spatiotemporal clustering (A1.4 and A1.11) were diffuse over space, but more concentrated temporally. Sub-lineage A2 clustered spatiotemporally most obviously in 2011 when it started to replace the majority of other genotypes and subsequently dispersed over the study area.
Figure 4.7 Clustering of S. Paratyphi A genotypes over space between 2005 and 2014

Clusters are shown in different colours, as indicated by the legend in the lower right. Significant clusters are highlighted by circles of various radii. The red star shows the location of Patan Hospital.
4.3 Discussion

*S. Paratyphi A* has become an increasingly important agent of typhoid fever, particularly with respect to disease prevalence and antimicrobial resistance. However, research on this pathogen is lagging behind that of *S. Typhi* and even other *Salmonella* serovars as a result of lack of public attention and a challenge of robust methodology. Previous epidemiological investigations of *S. Paratyphi A* lack a detailed characterization of the pathogen population and therefore do not provide any information about the disease dynamics. Here, I utilized a combination of bacterial genomics and more conventional epidemiology to provide new insights into the structure and dynamics of *S. Paratyphi A* in Nepal, which is essential for understanding the pattern of disease transmission in this endemic setting. I found that the *S. Paratyphi A* population in Nepal was diverse and highly dynamic and I observed evidence of a clonal expansion in early of study period followed by the replacement and expansion of a distinct clone in later years. *S. Paratyphi A* spread very rapidly throughout the geographical area during the clonal expansions.

From the global phylogeny, there were at least nine occasions where isolates from Nepal clustered with other South Asian isolates, mostly from India and Pakistan, suggesting that organism transfer between these neighbouring countries was common. Sub-lineages A1 and A2, which expanded in this location, shared common ancestors with Indian isolates. Therefore, it is unclear whether these expansions were due to successful clones locally or as a result of organism importation from neighbouring countries.
None of these Nepalese *S. Paratyphi* A in this study was resistant to any first-line drugs. Therefore, I propose that a mutation in *gyrA* and the acquisition of a novel *sopE* prophage were factors associated with the expansion of sub-lineage A1, which may indicate a strong selective pressure for fluoroquinolones in this setting. Previous surveillance between 1993-2003 in Nepal found that multi-drug resistant *S. Paratyphi* A were diminishing in this location population, this was occurring concurrently with decreased susceptibility against to fluoroquinolones. Such phenomenon was also observed in Nepal for *S. Typhi* and is concerning given that fluoroquinolones are widely used to treat typhoid fever and can be purchased without prescription. The replacement and rapid expansion of sub-lineage A2 throughout the study area from 2011 warrants further investigations as such events are rarely observed for *S. Paratyphi* A. These data imply that organisms within sub-lineage A2 have entered into a population that was either naïve or had lower immunity. *S. Paratyphi* A has been endemic in Nepal for many years and sub-lineage A1 was historically the dominant lineage, therefore this genotype shift may be associated with a subtle change in antigenicity of the organisms.

Defining the population structure of *S. Paratyphi* A in Nepal allowed me to further classify these organisms into different genotypes for investigating their short-term spatiotemporal distribution, which could not be performed by conventional approaches. The spatiotemporal clustering analysis demonstrated that a substantial number of genotypes clustered by space and time, suggesting a high level of acute, person-to-person transmission of *S. Paratyphi* A in Kathmandu. Previous case-control studies in Indonesia and Nepal also found that person-to-person transmission plays an important role in *S. Paratyphi* A infection with risk factors also including flooding and
contaminated street food, whereas contaminated drinking water were more associated with *S. Typhi* infection. Noticeably, all localized *S. Paratyphi A* outbreaks occurred during monsoon months between May and September, which corresponds with the seasonal distribution of typhoid fever in this setting. Previous study also demonstrated a high level of faecal contamination of the water supply, which correlated with increased rainfall. Given that only half of patients in this study access to the municipal water, the contamination of drinking water and water used in food preparation during that rainy season was likely the most important source of infection.

Outbreaks of typhoid fever caused by *S. Paratyphi A* are not uncommon. Very recently, an outbreak of *S. Paratyphi A* infection caused by an identical organisms was reported in Cambodia, which spread to several provinces and sickened dozens of European, American, Japanese travellers returning home. Even though the source of this outbreak remains unknown, contaminated food is the most reasonable cause of this large-scale infection. Such outbreaks may go unobserved in Nepal, therefore improvements in food safety should be considered as an important primary control measure for *S. Paratyphi A* infection. While improvements in safe water, food safety and proper sanitation seems to be the most effective way to manage typhoid fever, it is unlikely that it would be an imminent solution for Nepal. As a result, there is an urgent need of *S. Paratyphi A* vaccine, or more ideally, a bivalent vaccine which can prevent both *S. Typhi* and *S. Paratyphi A* infections.

The relative role of typhoid carriage in the disease transmission, as well as maintaining genetic diversity and generating novel genotypes, has been extensively
questioned. By whole-genome sequencing 17 carriage isolates and relating these genomic data with acute isolates, I have, for the first time, provided new insights into the genetic relatedness between acute and carriage S. Paratyphi A isolates and then assessed the potential role of carriage in this endemic setting. I found a high genetic diversity within the carriage isolates from this location; the genetic diversity suggested that the majority of these isolates (59%, 10/17) were likely associated with recent asymptomatic/acute infections. However, the exact duration of carriage of these individuals is unknown as the majority did not have recent history of typhoid. The remainders of the carriage isolates were generally distantly related to the acute isolates. However, the genetic diversity within the Nepalese S. Paratyphi A reduced substantially after 2011 with almost all isolates falling within sub-lineage A2. Furthermore, the carriage isolates appear to have accumulated a set of differing mutations during gall bladder carriage and have become genetically distinct from contemporary acute isolates. My limited data suggests that S. Paratyphi A carriers did not play an important role in maintaining the genetic diversity nor in the disease transmission in this endemic setting.

In conclusion, this chapter represented a detailed molecular epidemiological investigation of S. Paratyphi A in Nepal. The structure and dynamics of the pathogen population was described and distinguishing by both clonal expansion and clonal replacement. I found a high rate of reduced fluoroquinolone susceptibility in S. Paratyphi A in this location, which likely contributed to the successful expansion of these organisms. The rapid dissemination within local population and diversification of dominant clones resulted in short-term spatiotemporal clusters of various genotypes, indicating the importance of human-to-human transmission for S.
Paratyphi A in this setting. My data also demonstrated that in a highly endemic setting, S. Paratyphi A carriage likely plays a limited contribution to the disease transmission and maintenance of the bacterial diversity.
Chapter 5

The molecular and spatial epidemiology of typhoid fever in rural Cambodia

5.1 Introduction

Typhoid fever is mainly contracted through the consumption of contaminated water or food. Therefore, the disease is largely preventable through the provision of safe water, food safety and adequate sanitation. However, such interventions are still challenging in many endemic locations given the huge economic costs and long timelines. Licensed typhoid vaccines have been proved to be an effective short-term strategy in reducing typhoid burden and are currently recommended by the World Health Organization (WHO) in areas with high disease burden and increasing antimicrobial resistance. Despite this recommendation, the programmatic use of typhoid vaccines is very limited, largely due to insufficient data on the epidemiology of the disease and prevalence of antimicrobial resistant organisms for evidence generation and policy making. Additionally, identification of high-risk areas and vulnerable populations for targeted vaccination appears as a big challenge in resource-limited settings where usually lack of systemic disease surveillance. In such circumstances, case detection and appropriate antimicrobial therapy become crucial for typhoid management in endemic locations. However, the effectiveness of these strategies is diminished due to a lack of accurate and inexpensive rapid diagnostic tests for typhoid as well as the emergence of antimicrobial resistance organisms.
Most studies on typhoid have focused on urban slum populations, where high incidence rates have been reported\textsuperscript{133,311,312}. However, the epidemiological features of typhoid fever in different environmental settings remain unclear, especially in rural areas with poor access to healthcare. Investigation on the transmission patterns and risk factors of typhoid fever in both urban and rural areas is critical for effective disease control and prevention\textsuperscript{309}. Over the last decade, advances of molecular genotyping methods such as Single Nucleotide Polymorphisms (SNPs) typing and whole genome sequencing (WGS) have successfully provided unprecedented insights into the bacterial population structure and evolutionary relationships between isolates\textsuperscript{128,244}. SNP-based typing and WGS have played a significant role in studying the molecular epidemiology of typhoid fever, discovering a diverse range of \textit{S. Typhi} haplotypes that are commonly found co-circulating in different geographical settings\textsuperscript{131–134,246}. These studies have outlined the importance of environment transmission within the localized human populations. Such molecular investigations are also critical for understanding the bacterial population dynamics and tracking the emergence of antimicrobial resistant organisms. The current population structure of \textit{S. Typhi} has been primarily driven by series of clonal expansions and scattering of a specific haplotype (H58) in Asia and Africa. This globally dominant H58 haplotype mostly exhibits reduced susceptibility to fluoroquinolones and is commonly multidrug resistant against first-line agents (ampicillin, chloramphenicol, and trimethoprim-sulphamethoxazole)\textsuperscript{248}.

Typhoid fever is endemic in Cambodia but epidemiological data regarding to the mortality, morbidity and risk factors for the disease are limited. Between 2006 and 2008, a community-based study was conducted near Phnom Penh (the capital city)
and multidrug resistant *S. Typhi* with reduced susceptibility to ciprofloxacin was routinely detected from patients presenting with acute fever\(^{313}\). Additionally, a hospital-based surveillance of pediatric bloodstream infections in Siem Reap (Northwest of Cambodia) from 2007 and 2011 reported that *S. Typhi* was the most common pathogen causing pediatric bloodstream infection and H58 *S. Typhi* exhibiting reduced susceptibility to ciprofloxacin was dominant (accounting for 96% of *S. Typhi* isolates)\(^{247}\). In this investigation, I utilized whole genome sequencing to further characterize the H58 *S. Typhi* population in Siem Reap, Cambodia and combined the resulting phylogenetic information with additional epidemiological approaches to investigate the spatiotemporal distribution of *S. Typhi* and identify population-level risk factors associated with typhoid fever in this location. The data presented in this study was published in PLoS Neglected Tropical Diseases in June 2016 (Appendix G).

### 5.2 Results

#### 5.2.1 Baseline characteristics

Between 2007 and 2014, there were 284 microbiologically confirmed cases of typhoid fever caused by *S. Typhi* at Angkor Hospital for Children (AHC) in Siem Reap (method section 2.4.3). *S. Paratyphi A* was uncommon, with only three cases in 2008 followed by an isolated outbreak in 2013-2014 (38 cases). A total of 262/284 (93%) of the confirmed *S. Typhi* cases lived within a 100 km radius of AHC and spanned 78 communes; these 78 communes were selected for a spatial investigation and to determine population level risk factors for typhoid fever. During this same period there were 19,877 admissions with an ICD-10 discharge diagnosis other than typhoid fever originating from the same geographic area. The baseline characteristics
of all communes and those with at least one case of typhoid fever are shown in Table 5.1.

Of the 262 cases of typhoid fever living within a 100 km radius of AHC, the median age was 8.2 years (interquartile range (IQR): 5.1-11.5 years). Additionally, 62/262 (24%) of the cases were less than five years of age and 142/262 (54%) were female. As shown in Figure 5.1a, the absolute number of confirmed cases of typhoid fever increased dramatically (from 12 cases per year to 71 cases per year) between 2009 and 2012, but then declined in 2013 and 2014 (28 and 45 cases in 2013 and 2014; respectively); data from my non-confirmed typhoid cases also reflected this trend. Over this same time period (2009 to 2014) the number of patients attending AHC for other conditions (control population) mirrored the distribution of the cases (Figure 5.1b). There was seasonal variation in the number of typhoid cases, with the majority of the cases (178/262; 68%) occurring during the early monsoon months (April, May, June and July) (Figures 5.1c & 5.1d). In late monsoon months (August to October), the number of cases declined to less than two cases per month and generally remained below this threshold in the dry season (November to March) (Figures 5.1c & 5.1d).
Table 5.1 Baseline characteristics of all communes and those with at least one case of typhoid fever

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All communes</th>
<th>Typhoid communes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>IQR</td>
</tr>
<tr>
<td>n=243</td>
<td>n=78</td>
<td></td>
</tr>
<tr>
<td>Population density/km²</td>
<td>105.7</td>
<td>53-210</td>
</tr>
<tr>
<td>Elevation, m</td>
<td>17</td>
<td>12-28</td>
</tr>
<tr>
<td>Distance to lake, km</td>
<td>45.3</td>
<td>24-63</td>
</tr>
<tr>
<td>Average household size</td>
<td>4.8</td>
<td>4.6-5.0</td>
</tr>
<tr>
<td>Percent of population &lt;15 yr</td>
<td>36.4%</td>
<td>34-39%</td>
</tr>
<tr>
<td>Median age of population, yr</td>
<td>19.5</td>
<td>18-21</td>
</tr>
<tr>
<td>Adult literacy</td>
<td>72.8%</td>
<td>59-82%</td>
</tr>
<tr>
<td>Female adult literacy</td>
<td>65.3%</td>
<td>50-75%</td>
</tr>
<tr>
<td>Total attending school</td>
<td>28.6%</td>
<td>26-31%</td>
</tr>
<tr>
<td>Female attending school</td>
<td>26.8%</td>
<td>24-29%</td>
</tr>
<tr>
<td>Female education &gt;25 years /1,000 population</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary not completed</td>
<td>85.6</td>
<td>63-101</td>
</tr>
<tr>
<td>Primary/Lower secondary</td>
<td>27.8</td>
<td>16-55</td>
</tr>
<tr>
<td>Secondary or above</td>
<td>0.51</td>
<td>0.1-1.6</td>
</tr>
<tr>
<td>Toilet, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>83.1%</td>
<td>63-92%</td>
</tr>
<tr>
<td>Sewage</td>
<td>5.3%</td>
<td>2-14%</td>
</tr>
<tr>
<td>Septic tank</td>
<td>3.9%</td>
<td>1-16%</td>
</tr>
<tr>
<td>Pit latrine</td>
<td>2.0%</td>
<td>1-5%</td>
</tr>
<tr>
<td>Drinking water, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piped</td>
<td>1.5%</td>
<td>1-4%</td>
</tr>
<tr>
<td>Tube/pipe well</td>
<td>10.2%</td>
<td>3-28%</td>
</tr>
<tr>
<td>Dug well</td>
<td>26.9%</td>
<td>11-56%</td>
</tr>
<tr>
<td>Spring/river</td>
<td>23.9%</td>
<td>4-54%</td>
</tr>
<tr>
<td>Drinking water location, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within premises</td>
<td>19.3%</td>
<td>10-35%</td>
</tr>
<tr>
<td>Near premises</td>
<td>31.1%</td>
<td>22-40%</td>
</tr>
<tr>
<td>Away premises</td>
<td>41.8%</td>
<td>23-56%</td>
</tr>
</tbody>
</table>

IQR: interquartile range.
Figure 5.1 The annual and seasonal distribution of typhoid fever cases at Angkor Hospital for Children in Cambodia

a) The annual number of culture confirmed (solid line) and non-confirmed (broken line) typhoid cases at AHC from 2007 to 2014. b) The annual number of total admissions at AHC from 2007 to 2014. c) The mean monthly count of typhoid cases aggregated from 2007 to 2014. d) The average monthly rainfall (mm) per month over the study period.
5.2.2 Spatiotemporal clustering of typhoid fever cases

The majority of \textit{S.} Typhi cases (241/284; 85\%) originated from communes located within Siem Reap province (Figure 5.2). The median population density in communes with at least once case of typhoid fever was 119 people/km$^2$ (IQR: 60-212), and 70/78 (90\%) of communes with a typhoid fever case were classified as rural. Compared to typhoid cases, the non-typhoid fever population controls came from a larger area (243 communes), the median population density of which was lower at 106 people/km$^2$ (IQR: 53-210); however, a similar proportion of these communes (220/243; 91\%) was also classified as rural (Figure 5.2).

The estimated median commune level minimum incidence of reported cases of typhoid fever over the study period was 0.62/1,000 children aged <15 years (IQR: 0.37-1.02; range: 0.5-11.36). The reported incidence varied significantly across the 78 communes. Kampong Kleang commune (Soutr Nikom district, Siem Reap) showed the highest incidence of typhoid fever over the study period with 11.36 cases of typhoid fever /1,000 population of children aged <15 years (Figure 5.2c). This area is renowned for its floating villages and is situated on the edge of Tonle Sap Lake, approximately 35 km southeast of Siem Reap City. The second highest incidence was identified in Kaoh Chiveang commune (Aek Phnum district, Battambang, 33 km southwest of Siem Reap City) with 4.1 cases/1000 people aged <15 years over the study period (Figure 5.2c). Both of these areas experience heavy flooding when the Tonle Sap Lake expands during the rainy season.

Overall, there was some evidence of positive spatial autocorrelation (case clustering) across the 78 communes that had at least one case of typhoid fever between 2007 and
2014 (Moran’s I=0.11, \( p<0.056 \)). The magnitude of this autocorrelation varied over time and was the most significant in 2013 (Moran’s I=0.19, \( p<0.019 \)) but was non-significant in other years. I was able to identify three significant spatiotemporal clusters associated with high rates of typhoid fever. The first occurred in 2008 toward the west of the study area and had a radius of 23.8 km; this cluster had 1.27 predicted cases and 10 observed cases (relative risk [RR]=8.17, \( p=0.002 \)). The second cluster occurred in 2012 in the central northern area and had a radius of 10.8 km, with 1.67 predicted cases and 12 observed cases (RR=7.47, \( p<0.001 \)). The final cluster occurred in 2013 in the southeastern area and had a radius of 15.5 km, with 0.88 predicted cases and 14 observed cases (RR=16.8, \( p<0.0001 \)) (Figure 5.2d).
Figure 5.2 The spatial distribution of typhoid fever cases in Siem Reap province, Cambodia

a) North oriented map of Cambodia, the black cross shows the location of AHC. b) Map showing the population density (people/km², color-coding in key) of the 78 communes within the typhoid study area. AHC is shown by the black cross, the black border denotes Siem Reap province and the left and right asterisks are mark the locations of the communes with highest incidence of typhoid fever; Kaoh Chiveang and Kampong Kleang, respectively. c) Map of the study area showing the rate of reported typhoid cases per 1,000 population under the age of 15 years (color-coding in key). d) Map of the study area showing significant spatiotemporal clusters of typhoid during the study period, the size of the grey circles corresponds to the radius of the cluster and the years of the clusters are denoted.
5.2.3 The population structure of Salmonella Typhi in Siem Reap province, Cambodia

The resulting WGS data demonstrated that 97% (203/209) of the sequenced Cambodian isolates could be attributed to haplotype H58. The majority (199/203, 98%) of the H58 isolates exhibited intermediate susceptibility against fluoroquinolones (0.12-0.5 μg/mL) via the common amino acid substitution of serine to phenylalanine at codon 83 (S83F) in the DNA gyrase protein encoded by gyrA. There was a strong association between haplotype H58 and an IncHI1 plasmid, which confers an MDR phenotype, with 89% (180/203) of the H58 isolates harboring the common IncHI1 plasmid and the corresponding antimicrobial resistance phenotype. For the six non-H58 isolates, no mutations were observed in the gyrA gene, while two (33%) carried the same IncHI1 plasmid as found in the H58 isolates. I identified 188 SNPs across the H58 population and, from a SNP-based phylogeny, identified the circulation of at least four lineages of H58 circulating in the selected area of Cambodia between 2007 and 2012 (Fig 5.3a). These lineages, designated here as I-IV, differed from each other by as little as three to five SNPs and were phylogenetically well-supported (bootstrap values ≥ 87%). The majority of the H58 isolates fell into lineage IV (152/203, 75%) and lineage III (41/203, 20%).
Figure 5.3 The phylogenetic structure of the H58 lineage of Cambodian *Salmonella Typhi*

a) Maximum likelihood phylogenetic tree of the 203 H58 isolates identified during this project (scale bar denotes SNP differences). The sub-lineages are shown on the major branches. Isolates exhibiting a multi-drug resistance (MDR) phenotype are indicated by black nodes. The tree is midpoint-rooted for the purpose of clarity. Bootstrap values >85% are indicated by an asterisk.

b) Minimum spanning tree subdividing H58 lineage III and IV into the various sublineages (IIIa, IIIb, IIIc, IV, IVa, IVb, IVc). The various sublineages are color-coded for reference and the number of each variant is indicated by the cluster size. The number on each of the branches signifies the number of SNPs between each cluster.
5.2.4 The spatiotemporal distribution of *Salmonella* Typhi genotypes

To investigate short-term evolutionary traits within the identified lineages, I constructed a SNP-based minimum spanning tree (Figure 5.3b). Using these data, I was able to investigate the local population dynamics and detected several clonal clusters emerging from lineage III (IIIa-IIIc) and lineage IV (IVa-IVc). My data show a complex temporal distribution of *S*. Typhi H58 sublineages circulating in this location between 2007 and 2012 (Figure 5.4a). The distribution of these various strains was highly dynamic, with strain replacements, potential extinctions and the specific microevolution and expansion of H58-IVc (Figure 5.4a). In 2011 and 2012, H58-IVc became the dominant genotype, accounting for 44% (18/42) and 85% (61/72) of all *S*. Typhi isolates in these years, respectively.

I next aimed to identify spatiotemporal clustering of the various *S*. Typhi H58 sublineages, and found that IIIc, IV, IVb and IVc all displayed significant evidence of clustering over space and time. Notably, the locations of these clusters were generally different between sublineages, signifying some degree of geographical variation of the circulating *S*. Typhi strains. For example, I identified significant clustering of H58-IIIc in the western part of the study area in 2011 (*p*<0.001, RR: 26.7, radius: 36km) (Figure 5.4b) and clustering of the emergent H58-IVc strain in both 2011 (Kampong Khleang commune, *p*<0.001, RR: 39.4, radius: <1km) and in two locations in 2012 (smaller cluster, *p*=0.017, RR: 5.17, radius: 6.2km; larger cluster, *p*<0.001, RR: 5.87, radius: 33.9km).
Figure 5.4 The spatiotemporal distribution of the various *Salmonella* Typhi lineages/sublineages in Siem Reap province, Cambodia

a) Bar chart shows the annual distribution of the various *S. Typhi* lineages/sublineages from 2007 to 2012; sublineages are color-coded as in Figure 3b. b) Maps showing significant spatiotemporal clusters identified for sublineages IIIc, IV, IVb and IVc. The timing of each cluster is shown by the year in black text and the dotted circle represents the radius of the detected cluster. Background colors represent the rate of each sublineage per 1,000 population aged under 15 years. The incidence rates vary between sublineages, ranging from 0 to a maximum of 0.8 (IIIc), 3.12 (IV), 2.56 (IVb) and 5.84 (IVc) 5.84 cases/1,000 population aged under 15 years.
5.2.5 Population risk factors for typhoid fever

I additionally investigated associations between rates of typhoid in children and demographic and sanitation variables at the commune level. I found a number of significant risk factors (e.g. low female education level and collection of drinking water near the household premises) and protective factors (e.g. higher population density, elevation, distance from lake and attendance at school) associated with the rate of typhoid hospitalizations in the univariate analysis (Table 5.2). However, after controlling for confounders, I found that the distance of the centroid of the commune to the perimeter of the lake was strongly and significantly associated with rate of typhoid cases (10km increase in distance from the lake, incidence rate ratio (IRR): 0.38, 95%CI 0.26-0.55, \( p < 0.001 \)) (Table 5.2). Furthermore, the relative numbers of households within the commune connected to public sewage services and households using a sunken well were also strongly protective, however these associations were reversed through interaction with increasing number of households using wells and distance from the lake, respectively (Table 5.2). Finally, a high number of households reporting drinking water retrieval from ‘within the household premises’ were also associated with a significant protective effect (log households/1,000 households, IRR: 0.65, 95%CI: 0.49-0.86, \( p = 0.003 \)).
Table 5.2 Regression results of highlighting factors associated with typhoid cases

<table>
<thead>
<tr>
<th>Commune characteristic</th>
<th>Univariable</th>
<th>Multivariable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR (95%CI)</td>
<td>p</td>
</tr>
<tr>
<td>Population density^</td>
<td>0.81 (0.70-0.95)</td>
<td>0.008</td>
</tr>
<tr>
<td>Elevation, 10m</td>
<td>0.89 (0.81-0.99)</td>
<td>0.026</td>
</tr>
<tr>
<td>Distance to lake, 10km</td>
<td>0.81 (0.74-0.89)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average household size</td>
<td>1.54 (0.66-3.57)</td>
<td>0.317</td>
</tr>
<tr>
<td>Total attending school/1,000^</td>
<td>0.11 (0.04-0.33)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female education &gt;25 years /1000 population^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary not completed</td>
<td>2.59 (1.53-4.38)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Primary/Lower secondary</td>
<td>0.94 (0.70-1.25)</td>
<td>0.654</td>
</tr>
<tr>
<td>Secondary or above</td>
<td>0.96 (0.80-1.17)</td>
<td>0.714</td>
</tr>
<tr>
<td>Toilets per 1000 people^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1.24 (0.84-1.79)</td>
<td>0.246</td>
</tr>
<tr>
<td>Sewage</td>
<td>0.96 (0.81-1.15)</td>
<td>0.676</td>
</tr>
<tr>
<td>x households with wells</td>
<td>1.19 (1.07-1.32)</td>
<td>0.001</td>
</tr>
<tr>
<td>Septic tank</td>
<td>0.89 (0.78-1.02)</td>
<td>0.090</td>
</tr>
<tr>
<td>Pit latrine</td>
<td>0.95 (0.79-1.15)</td>
<td>0.621</td>
</tr>
<tr>
<td>Drinking water, hh/1000 hh^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piped</td>
<td>0.87 (0.74-1.03)</td>
<td>0.104</td>
</tr>
<tr>
<td>Tube/pipe well</td>
<td>0.82 (0.73-0.92)</td>
<td>0.001</td>
</tr>
<tr>
<td>Dug well</td>
<td>0.83 (0.64-0.83)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>x distance to lake</td>
<td>1.16 (1.09-1.25)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Spring/river</td>
<td>1.15 (1.05-1.25)</td>
<td>0.003</td>
</tr>
<tr>
<td>Drinking water location, hh/1000 hh^</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within premises</td>
<td>0.71 (0.55-0.93)</td>
<td>0.013</td>
</tr>
<tr>
<td>Near premises</td>
<td>3.38 (2.24-5.10)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Away premises</td>
<td>0.88 (0.73-1.06)</td>
<td>0.177</td>
</tr>
</tbody>
</table>

^log of the variable was included ; hh: household
5.3 Discussion

In this study I combined conventional epidemiological methods, current genome sequencing tools and geospatial mapping to add insight into the epidemiology of typhoid fever in atric patients attending a single healthcare facility in central Cambodia. The majority of recent typhoid fever studies originate from urban locations in low-income countries. This study provides a new perspective into this important community-acquired infection from a predominantly rural setting. The primary finding of this study is that there is a considerable and widespread burden of pediatric typhoid fever in rural Cambodia, thus questioning the dogma that typhoid fever is predominantly geographically restricted to urban populations with poor sanitation systems. My data are consistent with findings from a recent study conducted across sub-Saharan Africa. The Typhoid Surveillance in Africa Programme (TSAP) found a large burden of typhoid fever in younger children and almost equivalent population incidences between urban and rural settings. This distribution was most apparent in West Africa (Burkina Faso and Ghana) and was similarly restricted to children aged less than 15 years. Therefore, I infer that the epidemiology of typhoid fever in Cambodia may be more similar to contemporary observations from sub-Saharan Africa, as opposed to the urban distribution that has commonly been observed across much of Asia.

The impending availability of Vi-conjugate vaccine raises the question of who should be given this vaccine and when it should be given to obtain maximum benefit in the control of typhoid fever. This issue is complicated by a lack of population-based incidence data and a poor understanding of the burden of disease in school and preschool aged...
children, for whom the conjugated form of the Vi polysaccharide vaccine would be particularly beneficial. My data indicate a substantial burden of typhoid fever in school and preschool aged children in this area, with a hospital-based incidence (i.e. a minimum population incidence) of 11.36 cases of typhoid fever /1,000 population in children aged <15 years over the study period. The overall burden of typhoid fever in this population is likely to be greater than I have estimated due to poor sensitivity of blood culture and restriction of the study to a single healthcare centre. Siem Reap province could be a suitable location in which to trial, or even introduce, the next generation typhoid vaccines in Cambodia that have been tested elsewhere. Further, I suggest that immunizing school-aged children in the period prior to the wet season may provide the most economic and prudent approach for vaccine introduction.

Between 2007 and 2012, I observed a sharp increase in the number of typhoid cases concurrent with an increasing geographic expansion. I also observed that typhoid fever in this population followed a seasonal pattern, suggesting an association with rainfall and potentially with localized flooding and the contamination of water sources. The population-based risk factors support these hypotheses, as living further away from Tonle Sap Lake and access to water within the household were highly protective. Additionally, I found that two communes located next to the lake (Kaoh Chiveang and Kampong Kleang) had the highest incidence of typhoid fever and had large clusters of cases in 2008 and 2013. This case clustering in specific locations warrants further investigation at the household level to understand specific sanitation-associated risk factors and likely exposures to S. Typhi in this setting. It appears that access to lake water in some of
these communes, such as Kaoh Chiveang, is vital for the household water supply and I hypothesize that the lake water is more prone to localised fecal contamination at specific times throughout the year.

Using targeted SNP-specific PCR, I have previously shown that MDR H58 S. Typhi strains dominate in this population. My WGS investigation confirmed these findings and identified additional diversification in this population. I was able to separate these H58 strains into seven (IIIa, IIIb, IIIc, IV, IVa, IVb, IVc) major sublineages. These discrete groups varied in size and were segregated by only limited numbers of SNPs. I did observe some evidence of expansion of sublineage IVc between 2009 and 2012; this correlated with several spatiotemporal clusters suggesting small disease outbreaks. I currently cannot explain the expansion of this group and my strain selection for sequencing was limited by the availability of strains isolated only up to 2012. Despite some clustering of closely related strains, the overall temporal and spatial distribution of strains was random, with a range of S. Typhi H58 sublineages circulating throughout the study period, which is similar to patterns described in urban settings in Asia.

This study has some limitations. First the data originated from patients attending a single healthcare facility, without the added support of healthcare utilization data. This approach, while cost-effective, induces bias in the spatial and risk factor analyses. Furthermore, while the associations identified in the regression analysis are plausible and provide direction for future investigations, they should be viewed with caution. The population level census data does not allow examination of exposures at an individual or
household-level and provides only broad epidemiological evidence. However, the association with distance to the lake and water and sanitation variables suggests these factors should be examined more rigorously in the future with respect to the dynamics of typhoid fever outbreaks. Similarly, the identification and location of the spatiotemporal clusters should be interpreted with some degree of caution. Communes without cases were not included in the cluster analyses due to a lack of data as to whether these regions truly lacked typhoid cases. A dataset with more complete spatial information on presence and absence of typhoid would permit a more reliable analysis.

In summary, I found a large burden of typhoid fever in children in rural Cambodia. My conventional population-based risk factor analysis identified access to water in the household and increasing distance from Tonle Sap Lake as protective against typhoid fever in communes. Spatial mapping and WGS provided additional resolution to investigate these findings and confirmed that proximity to that lake was associated with discrete disease clusters. I confirmed the dominance of MDR H58 S. Typhi in this location and found a substantial amount of diversification within this lineage. My data provide a platform for additional studies in the Cambodian population and suggest that this is a suitable location in which to introduce Vi conjugate vaccines for school children.
Chapter 6

Genetic traits of *Salmonella* Typhi gallbladder carriage isolates and their role in disease transmission in Kathmandu, Nepal

6.1 Introduction

Typhoid carriage has been recognized as an important public health problem for over a century. Mary Mallon (better known as Typhoid Mary), a cook in New York, and Mr N, a milker in England, were associated with the first true epidemiological investigations of infections and typhoid outbreaks caused by asymptomatic carriers in the early 1900s. Typhoid carriage is typically thought to occur in patients who fail to fully clear the organisms after recovering from their acute phase of illness. These people become asymptomatic carriers and can shed the bacteria for up to three months. A subset of these individuals (~2-5 percent) may develop into a chronic carrier state, whereby they intermittently shed the bacteria via stool and urine for more than a year. Chronic carriage may provide an ecological niche that facilitates the transmission and persistence of the bacteria in human populations, potentially posing a major public health threat. As a result, to prevent disease transmission and eventually eradicate the causative pathogens, it is essential to detect and provide treatment for the chronic carriers. However, this remains a huge challenge due to several impeding factors exist such as a lack of epidemiological data on the long-term follow-up of typhoid patients, intermittent fecal shedding, sup-optimal diagnostic methods with poor sensitivity to detect *Salmonella*
carriage, and a lack of scientific data on effectiveness of antimicrobial treatment for chronic carriage. The situation is further exacerbated given the fact that infections can be symptomless and up to 25 percent of carriers have no history of acute typhoid.

Some understanding of typhoid carriage has only been gleaned in recent decades. Evidence from mouse models of *Salmonella* infection and clinical investigations have demonstrated that the gallbladder is the primary permissive niche for long-term bacterial persistence. Various epidemiological investigations have shown an association between typhoid carriers and individuals with gallstones in their gallbladder. *Salmonella* can form dense biofilms on the gallstone surface and specifically bind to cholesterol gallstones, which is considered as a hallmark of typhoid carriage. Moreover, recent studies have suggested that biofilms form on the gallbladder epithelium and it has been suggested that intracellular invasion of epithelial cells lining the gallbladder may be another possible mechanism for persistence. Despite these advances, there are still unresolved questions regarding the development of a chronic carrier state such as host responses to long-term colonization of *S. Typhi*, bacterial adaptive mechanisms for surviving within the gallbladder environment, the role of chronic carriers in disease transmission in endemic areas, regulatory mechanisms of biofilm formation, and the interaction between *S. Typhi* and the human gallbladder. Genomic characterization of individual carriage isolates as well as genomic comparison between acute and carrier isolates have been performed to investigate the genetic signatures associated with typhoid carriage. However, these
studies provided limited information because of the limited number of isolates, making the population structure and genetic background of the acute and chronic isolates difficult to resolve. In this chapter, I performed the genome sequencing and analyses of 24 isolates from typhoid carriers and 96 isolates from acute patients collected during the time same period in Kathmandu, Nepal to describe the population structure and characterize the genetic relatedness between acute and carrier S. Typhi isolates. This investigation also aimed to decipher potential genetic traits associated with carrier isolates and estimate the role of carriage in the disease transmission in an area that is highly endemic for typhoid fever.

6.2 Results

6.2.1 The phylogenetic structure of Nepalese acute and carrier Salmonella Typhi isolates between 2007 and 2010

My genomic and phylogenetic analyses of 120 S. Typhi isolates (24 from typhoid carriers and 96 from acute patients) demonstrated that genotype 4.3.1 (H58) was the most dominant genotype isolated during this study period, constituting 65.6% (63/96) of all acute isolates and 62.5% (15/24) of all carriage isolates. The second most common genotype was 3.3.0 (H1), accounting for 14.6% (14/96) and 12.5% (3/24) of all acute and carrier isolates, respectively. Additionally, there was high genetic diversity within this bacterial population with various minor genotypes co-circulating, including the subclades 3.2.2; 3.0.1; 2.2.2; 2.2.1; clades 4.1; 3.1; 2.2 and primary cluster 2 (Figure 6.1). Of these minor genotypes, the carriage isolates fell within subclade 3.2.2 (8.3%; 2/24), 2.2.2 (4.2%; 1/24), clade 2.2 (8.3%; 2/24) and primary cluster 2 (4.2%; 1/24). Overall, the
carriage isolates were not significantly more likely to be H58 in comparison to non-H58 organisms \((p=0.774, \text{Chi-squared test})\). These data suggest that *S. Typhi* carriage isolates were not restricted to any particular bacterial genotype; instead, the genotype distribution among carrier isolates mirrored the general genetic structure of the sampled bacterial population.

I hypothesized that carriage isolates may be more resistant to antimicrobials, therefore I firstly selected to investigate the presence of fluoroquinolone-resistance mutations. The genome sequence data showed that mutations conferring non-susceptibility to fluoroquinolones were predominately found within the H58 organisms \((64/78, 82.1\%)\), rather than the non-H58 organisms \((1/42, 2.4\%)\). The majority of the mutations occurred at codon 83 of the DNA gyrase gene *gyrA*, changing serine to phenylalanine (S83F) in the protein. Comparing *gyrA* mutations between the acute and carriage isolates, respectively, within the H58 organisms, I found that 76.2\% (48/63) and 60\% (9/15) contained the S83F mutation; 7.9\% (5/63) and 13.3\% (2/15) had a S83Y mutation and 15.9\% (10/63) and 26.7\% (4/15) had no mutation. There was no significant difference \((p=0.327, \text{Chi-squared test})\) in the presence of fluoroquinolone resistance-associated mutations between acute and carrier isolates in this group.
Figure 6.1 The phylogenetic structure of carriage and acute *S. Typhi* isolates collected between 2007 and 2010

Mid-point rooted maximum likelihood tree based on core-genome SNPs of 120 *Salmonella* Typhi isolates including 24 from typhoid carriers and 96 from acute patients with the corresponding metadata: genotype, gyrA mutations and a heat map of functional distribution of nonsynonymous SNPs that were specific to chronic and acute isolates. The heat map is based on the number of nonsynonymous SNPs specific to chronic (represented as red bars) and acute isolates (represented as blue bars) in each functional class, with an emphasis on functions related to bacterial regulation, surface structures, pathogenicity and virulence. Chronic and acute isolates are shown as red and dark circles at the terminal nodes, respectively. Terminal branches leading to chronic isolates are highlighted in red.
6.2.2 Dissecting the genetic traits of *Salmonella Typhi* carriage isolates

There was a total of 2186 SNPs (in comparison with reference genome CT18) identified across this collection of 120 genomes from acute and carriage *S. Typhi* carriage isolates. 1038/2186 (47.5%) of these SNPs were nonsynonymous (44.7% (977/2186) missense and 2.8% (61/2186) were nonsense); 32.4% (708/2186) were synonymous and 20% (440/2186) were in the intergenic regions and pseudogenes. The mean $\text{dN/dS}$ for acute isolates was slightly less than the mean $\text{dN/dS}$ for carrier isolates (0.478 vs 0.495, respectively), although both of these figures suggest that most of the mutations were neutral. In order to better understand the adaptive point mutations associated with typhoid carriage, all nonsynonymous SNPs (NSs) occurring exclusively within carrier *S. Typhi* genomes were identified. Genes containing these SNPs were further grouped by their predicted functions according to the *S. Typhi* functional classification scheme generated by the Sanger Institute. A comparable analysis was performed for all NSs in the acute *S. Typhi* isolates. There were a total of 228 carriage-specific NSs (212 missense and 16 nonsense mutations, appendix E) and 469 acute-specific NSs (437 missense and 32 nonsense mutations, appendix F). There was no significant difference ($p=0.924$, Chi-square test) in the proportion of nonsense mutations out of total specific NSs in the acute versus the carriage isolates across the whole phylogenetic structure. However, for genotype 4.3.1 (H58), the proportion of nonsense mutations out of total specific NSs for carrier isolates was significantly higher than that of acute isolates (10/60 compared to 2/67, Fisher exact test, $p=0.009$). These data suggested that gene degradation by nonsense mutations was more common in carriage isolates in comparison to the acute isolates within genotype 4.3.1. The inactivated genes among the carriage isolates included genes...
involved in the synthesis of peptidoglycan (*pbpC*), vitamin B12 receptor (*btuB*), general stress response regulator (*rpoS*), laterally acquired protein in SPI-7 (STY4562), membrane transport protein (STY3932), central metabolism (STY0230, *ggt*), hypothetical proteins (STY0929, STY4178) and osmotically inducible lipoprotein E precursor (*osmE*).

Overall, the NSs associated with acute versus carrier *S. Typhi* isolates could be grouped into 78 functions. The highest prevalence of NSs was found in gene functions related to hypothetical proteins, membranes lipoproteins, unknown function, transport/binding proteins, SPI-7, general regulatory function, surface polysaccharides and antigens, carbohydrate degradation, and DNA replication/modification (Figure 6.2). There was no statistically significant difference found in the prevalence of these NSs across all gene functions between acute and carrier isolates. However, the data showed that the proportion of NSs in the *viaB* operon was significantly higher in carriage isolates in comparison to the acute isolates (9/228 compared to 7/469, Chi squared test, *p*=0.04). This finding was also reflected for *S. Typhi* isolates belonging to genotype 4.3.1, with carriage isolates having more specific NSs in the *viaB* operon than the acute isolates (5/60 compared to 1/67, Fisher’s exact test, *p*=0.08). Additionally, there were two carriage isolates (GB428 and GB003) that had lost the Vi capsular polysaccharide due to the deletion of the entire SPI-7 region (134kb). Their loss of Vi expression was confirmed by agglutination test using Vi antisera under different osmolarity conditions (method section 2.3). The *viaB* operon is a ~14 kb region located within SPI-7 in the *S. Typhi* chromosome containing five genes responsible for Vi polysaccharide synthesis.
(tviABCDE) and a further five genes (vexABCDE) involved in transportation of the Vi capsule to the cell surface. The Vi capsular polysaccharide is an important virulence factor enhancing bacterial resistance to complement and phagocytic killing. Anti-Vi antibody titers have been found to be abnormally high in the blood of typhoid carriers and have been used as a serological marker for the detection of typhoid carriage. Here, I show that multiple S. Typhi carriage isolates had lost the ability to express Vi.
Figure 6.2 Top ten functional classes with highest prevalence of acute-specific nonsynonymous SNPs versus carrier-specific nonsynonymous SNPs
6.2.3 Positive selection associated with typhoid carriage

Next, I investigated signatures of positive selection by trying to identify similar genetic changes within different carrier isolates. Among the carrier-specific NSs identified, there were a number of different mutations occurring in the same gene or the same biological pathways in at least two phylogenetically unlinked carrier isolates. For example, within the \textit{viaB} operon as mentioned above, there were two NSs at codon 137 and 462 in the \textit{tviE} gene (isolates GB580 and GB026) and six NSs in codons 166, 504, 506, 508, 665, 752 in the \textit{tviD} gene (isolates GB005, GB026, GB076, GB125 and GB281). Both genes are known to be involved in the polymerization and translocation of the Vi capsule \textsuperscript{338}. Convergent NSs also occurred the \textit{rpoS} gene of isolates GB125 (nonsense mutation at codon 247) and GB705 (NSs at codon 94 and 250). The \textit{rpoS} gene encodes the sigma factor sigma-38, a central regulatory protein of the general stress responses (temperature, pH, osmolarity, redox state, antimicrobial peptide) and nutrient starvation. RpoS has been showed to down-regulate Vi expression and RpoS-negative strains of \textit{S. Typhi} overexpressed Vi polysaccharide at low and medium osmolarities, which can affect their invasiveness and macrophage resistance \textsuperscript{339,340}. A further example was NSs at codon 59 and 230 in the \textit{degS} gene (serine protease) (isolates GB005 and GB169). \textit{DegS} is a component of DegS-DegU two-component system, which is involved in expression of several degradative enzymes for salt stress responses and growth-limiting conditions in Gram-positive bacteria, suggesting it might have an important role for bacterial survival in the high salt concentration within the gall bladder. Additionally, three isolates (GB005, GB026, and GB705) had different NSs in codons 335, 406, 946, respectively, in STY1242 (\textit{ptsG} - glucose-specific PTS system IIIBC component). The PtsG enzyme is a
component of the glucose-specific phosphotransferase system, plays a role in phosphorylation and translocation of glucose across the bacterial membrane, and is induced in carbon-limited conditions \(^{341}\). None of the acute isolates had mutations in this gene. There were also several other genes containing NSs in more than two carriage isolates, such as STY0429 (\textit{SbcC} - exonuclease), STY0661 (\textit{dmsC} - molybdopterin containing oxidoreductase membrane anchor subunit), STY1447 (putative ribulose-5-phosphate 3-epimerase) and STY2760 (\textit{ratA} - putative exported protein).

With respect to convergent mutations within the same biological pathways, there were a number of carriage-specific NSs involved in LPS O-antigen synthesis and modification. For example, a NS in the \textit{rfc} gene (regulator of O-antigen polymerization) in isolate GB441; a NS in the STY2629 gene (LPS modification acyltransferase) of isolate GB335; two NSs in the \textit{rfbE} (CDP-tyvelose-2-epimerase) and \textit{rfaG} genes (lipopolysaccharide core biosynthesis protein) in isolate GB281 and three NSs in the \textit{rfbK} (phosphomannomutase), \textit{manB} (phosphomannomutase) and \textit{rfaD} genes (ADP-L-Glycero-D-mannoheptose-6-epimease) in isolate GB026. \textit{RfbK} and \textit{manB} are both related to GDP-mannose synthesis for the LPS and \textit{rfaD} is an enzyme that catalyzes the conversion of ADP-D-glycerol-D-mannoheptose to ADP-L-glycerol-D-mannoheptose, a precursor for the synthesis of inner-core LPS. The enrichment of NSs related to LPS structure suggests that LPS has an important role for the long-term colonization of these carrier isolates.
6.2.4 Estimating the role of typhoid carriage in disease transmission in Kathmandu, Nepal

Although the majority of acute and carriage isolates generally clustered together within the same genotypes across the phylogenetic tree of Nepalese S. Typhi, many carriage isolates were associated with atypically long terminal branches, suggesting they exhibited a greater degree of genetic variation from the acute isolates (Figure 6.1). One rational explanation for this phenomenon was that many of the S. Typhi isolates from typhoid carriers have been colonizing the gallbladder for a prolonged period and have undergone distinct mutational accumulation for adaptation, whereby they have gradually become distantly related to contemporary acute isolates. Additionally, there were very few cases where contemporary acute isolates were directly linked to carriage isolates within the phylogenetic structure. Taken together, these results potentially indicate that these typhoid carriers had limited contribution to the transmission of disease in Kathmandu.

Alternatively, there may be other explanations for the presence of long terminal branches in the phylogenetic tree of carriage isolates. For example, carriers may have been colonized by imported variants with a different genetic background, or more simply, the S. Typhi isolates from acute patients may have been under-sampled or not representative of the diversity circulating in the environment. I therefore aimed to identify the most probable molecular mechanisms contributing to the variation in branch lengths. The number of S. Typhi isolates was not evenly distributed among different genotypes, therefore, for this sub-analysis I selected the acute and carrier isolates belonging to only genotype 4.3.1 (H58), which is also the most successful and globally dominant genotype.
First, I reconstructed a global phylogeny of 4.3.1 S. Typhi isolates including 78 isolates from this study (63 from acute patients and 15 from typhoid carriers) and a global collection of 798 published previously isolates. The phylogenetic reconstruction showed that the majority of Nepalese 4.3.1 isolates belonged to a dominant group comprising of 9/15 carrier isolates and 45/63 acute isolates (Figure 6.3). For the remaining isolates, two carrier isolates (GB003 and GB044) clustered together with Indian isolates; the other four carrier isolates (GB076, GB441, GB266, and GB387) were distantly related with respect to the rest of the 4.3.1 isolates and consistently associated with long terminal branch lengths; the acute isolates were found to cluster with Indian (n=6) and Southeast Asian isolates (n=6) and other minor groups within the Nepalese 4.3.1 population. This finding suggested that there has been a clonal expansion of 4.3.1 S. Typhi in Kathmandu, with some evidence of inter-country typhoid transmission. However, it was difficult to determine the direction of transmission without accompanying epidemiological information. Furthermore, there were little data to support that long terminal branches associated with carrier isolates were driven by colonization of imported strains.
Figure 6.3 Phylogenetic structure of acute and carrier 4.3.1 *Salmonella* Typhi isolates from Nepal in the global context

Maximum likelihood phylogenetic tree of Nepalese H58 *S. Typhi* isolates (63 from acute patients and 15 from typhoid carriers) in the global context. H58 *S. Typhi* isolates from this study are highlighted in red circles at the terminal nodes. Terminal branches associated with carriage isolates are shown in red color. The outer ring exhibits the location of the isolates from Nepal and its neighboring countries as well as other regions in the world.
I next extracted and compared the branch lengths from the most recent common ancestor (tMRCA) for acute versus carriage isolates belonging to genotype 4.3.1. My data showed that the mean branch length from the tMRCA for the acute isolates was 0.00447 substitutions/site (range: 0.00196-0.0083) equivalent to 12 SNPs (range: 5-21 SNPs). The mean branch length from the tMRCA for the carriage isolates was 0.0052 substitutions/site (range: 0.00276-0.01108) equivalent to 14 SNPs (range: 7-28 SNPs). Therefore, although some carriage isolates within genotype 4.3.1 were associated with longer branch lengths in comparison to acute isolates, this difference was not significant. One explanation for this may be a variable duration of S. Typhi carriage.

Lastly, I estimated and compared the pairwise genetic distances within the acute isolates versus the pairwise genetic distances within carriage isolates. The rationale for this was that the number of acute isolates (n=63) was higher than chronic isolates (n=15), therefore the genetic diversity within acute isolates should be higher if there was no selection and the mutation rates were comparable. However, the median pairwise SNP distance within the acute isolates was 13 SNPs (IQR: 8-19 SNPs), which was significantly lower than the median pairwise SNP distance within the carriage isolates: 21 SNPs (IQR: 12-24) (Wilcoxon rank sum test, \( p=2.8\times10^{-9} \)) (Figure 6.4). Additionally, the median pairwise SNP difference between the acute and the carriage isolates was 17 SNPs, suggesting some degree of overlap in SNP content between these two groups. These findings demonstrate that there was more genetic diversity among carriage isolates than acute isolates, which may result from an accumulation of adaptive point mutations in
response to alternative selective pressures within the gallbladder.

These data suggest that carriage isolates likely follow different evolutionary pathways from acute isolates. During carriage *S. Typhi* consistently accumulates potentially adaptive point mutations, which may assist long-term survival within human gallbladder. My data proposes that the longer carrier isolates reside within the gallbladder, the more genetically distinct they are from acute isolates. Furthermore, it suggests that typhoid carriage may not be an important source of disease transmission in Kathmandu.
Figure 6.4 Distribution of pairwise SNP distances within and between acute and carrier isolates
6.3 Discussion

Understanding the molecular mechanisms associated with chronic typhoid carriage represents a challenging research area given that this population is difficult to identify prospectively. Additionally, there has been a lack of robust genetic methodology to investigate these questions due to the genetic conservation of *S. Typhi*. The environmental factors driving the evolution of *S. Typhi* within the gallbladder are poorly understood and little is known about the adaptive mechanisms that promote long-term survival. Previous genomic comparisons between *S. Typhi* isolates recovered from acute patients and typhoid carriers have failed to take into consideration the heterogeneity of the bacterial population, which may have a large effect on determining genetic variation between acute and carrier isolates. An ideal approach for genomic comparison should include a characterization of the bacterial population structure in the context of typhoid endemicity. In this study, my data demonstrated that typhoid carriage was induced by a diverse range of bacterial genotypes. Further, the development of a carrier state was neither restricted to any particular genotype nor to a fluoroquinolone-resistant phenotype. My study thus rejects the notions that typhoid carriage may be associated with one particular genotype or that a shift towards fluoroquinolone therapy has reduced carriage. In fact, the genetic distribution of carriage isolates largely reflected the overall genetic structure in the bacterial population, with the most dominant genotype being 4.3.1. Genotype 4.3.1 (H58) started to emerge and disseminate across South Asia in the early 1990s. Until now, 4.3.1 has been replacing other older preexisting genotypes to become the dominant genotype circulating in South Asia, Southeast Asia and Africa. According to my data, it is likely that the continued spread of 4.3.1 *S. Typhi* will
dominate over other minor genotypes in Kathmandu. This may also increase the detection of 4.3.1 *S. Typhi* in typhoid carriers over time, compromising the potential role that typhoid carriers play in maintaining bacterial diversity.

Assessing the role of chronic carriage in disease transmission represents one of the most important public health issues in typhoid control. Typhoid carriers have been widely considered as an important source of infection; however, their exact contribution to newly infected cases in endemic areas like Kathmandu remains questionable. Previous molecular epidemiological studies in endemic regions in Nepal, Vietnam, and Indonesia have shown the abundance of environmental transmission in these areas, with a wide diversity of co-circulating bacterial genotypes identified among acute typhoid patients. These studies have also implied a minimal contribution of person-to-person transmission to new typhoid cases in endemic settings. Here, a phylogenetic reconstruction for non-recombinant non-repetitive core genomes demonstrated a high level of genetic diversity of *S. Typhi* genotypes circulating in Kathmandu. More importantly, the phylogenetic structure of the acute and chronic isolates collected during the same period in a single location provided a unique opportunity to investigate the phylogenetic relationship between these isolates and assess the role of carriage in disease transmission. There were very few examples in the phylogenetic tree where the carriage isolates clustered in close proximity or directly gave rise to acute isolates. Further, many carriage isolates were associated with atypically long terminal branches. While there might be several factors that may lead to this phenomenon, my investigations suggested that the accumulation of adaptive point mutations in the carriage isolates was the most
likely molecular mechanism. Considering acute and carrier isolates are circulating in different ecological niches, long-term exposure to different selective pressures exerted by these environments may result in a difference in their accumulated adaptive mutations over time. This also means that the longer the duration of carriage in the gallbladder, the more genetically distantly related they are in comparison to contemporary acute isolates. This speculation is supported by the fact that genotype 4.3.1 carrier isolates exhibited a significantly higher level of genetic diversity in comparison to the acute isolates. Taken together, my analysis strongly advocates that S. Typhi have undergone a distinctive evolutionary pathway during the carrier state and play a limited role in disease transmission in Kathmandu.

I additionally aimed to investigate the genetic signatures associated with typhoid carriage with a focus on genetic function, particularly at the single nucleotide level. As typhoid carriage is not restricted to any bacterial genotype, gene acquisition is unlikely to contribute to the development of a carrier state. By identifying NS mutations occurring specifically in carrier isolates and classifying them into predicted functional classes for comparisons with those of acute isolates, I found that gene degradation by nonsense mutations was significantly higher in carriage than acute isolates within genotype 4.3.1. The phenotypic effect of gene inactivation on the bacterial phenotype and carriage is unknown. However, it is an intriguing phenomenon and worthy of further investigation, as gene inactivation has been shown to be an important evolutionary mechanism in the adaptation of S. Typhi. There was also evidence regarding the enrichment of NS mutations related to the Vi polysaccharide capsule in the carrier isolates. Vi antigen is
immunogenic and anti-Vi antibody gradually wanes in acute typhoid patients after recovery, but is persistent in the blood of chronic carriers.\textsuperscript{337,342} Despite the fact that anti-Vi antibody is not a reliable serological marker for the detection of typhoid carriers in endemic areas, data from sero-surveillance studies for chronic carriage in these areas have commonly reported a high prevalence of elevated levels of anti-Vi antibodies in healthy individuals, which may be associated with both carriers and repeatedly infected persons.\textsuperscript{343,344} Immunofluorescent staining of biofilms produced by \textit{S. Typhi} on the surface of human gallstones has shown the abundance of Vi capsule in the biofilm extracellular matrix, suggesting that \textit{S. Typhi} consistently express Vi antigen during the carrier state.\textsuperscript{329} The increased frequency of nonsynonymous mutations in the viaB operon (\textit{tviB}, \textit{tviD} and \textit{tviE} gene) of carrier isolates in this study together with their high level of anti-Vi antibodies in the blood (unpublished data) suggest that \textit{S. Typhi} residing in the gallbladder are under consistent selection pressure imposed by the human immune response.

Identifying genes that may be under positive selection among carriage isolates is crucial for understanding the evolutionary forces and bacterial adaptation to the gallbladder environment during the carrier state. Signatures of positive selection were detected in a number of genes containing differing carriage-specific NS mutations in at least two phylogenetically unlinked carriage isolates. Many of these genes were associated with gene regulation under stress conditions and the expression of virulence genes. For example, the global regulatory gene \textit{rpoS} is not only responsible for general stress responses and nutrient starvation but also regulates genes involved in biofilm formation,
the colonization of Peyer’s patches, persistence of S. Typhimurium in the spleen, and
virulence and Vi polysaccharide synthesis of S. Typhi. Furthermore, the degS gene
is involved in salt stress responses and growth-limiting conditions in Gram-positive
bacteria; STY1242 (ptsG - glucose-specific PTS system IIBC component) is activated
under the stress of carbon starvation. These findings suggest that S. Typhi is exposed to a
range of differing stresses within the human gallbladder. Furthermore, the genes
responsible for LPS biosynthesis and modification also displayed a marked accumulation
of NS mutations in the carriage isolates. LPS is a major component of the outer
membrane of all Gram-negative bacteria and represents one of the main factors
contributing to the resistance for high concentrations of bile salt in the gallbladder. LPS is also a key structural component of the biofilm extracellular matrix which forms on
human gallstones. The disruption of genes involved in LPS biosynthesis of S.
Typhimurium may have a negative effect on the production of biofilms and the
attachment of bacteria on contact surfaces. The enrichment of NS mutations in genes
involved in LPS biosynthesis and modification can lead to structural changes of the LPS
and thereby might enhance bile resistance or affect biofilm formation. There were also
several other genes potentially under positive selection which could have unknown
effects on carriage, including STY0429 (SbcC - exonuclease), STY0661 (dmsC -
molybdopterin containing oxidoreductase membrane anchor subunit), STY1447 (putative
ribulose-5-phosphate 3-epimerase), and STY2760 (ratA - putative exported protein).

This study has some limitations. First, the number of carrier and acute isolates was
relatively small and thus might not truly reflect the actual genetic structure in the
bacterial population. This sample size may also affect the interpretation of the phylogenetic distances between acute and carrier isolates and the functional analyses of the specific NS mutations in these isolates. Second, the duration of carriage is impossible to assess as most typhoid carriers do not recall a history of typhoid. My data suggested that the duration of carriage was likely to be highly variable, which consequently led to variable terminal branch lengths. It was not possible to measure the evolutionary rate of *S. Typhi* during the carriage state without knowing the duration of carriage in the gallbladder. Therefore, it is necessary for all future epidemiological investigations of typhoid fever to include a follow-up period of at least one year. This is essential not only to provide better estimate of the evolutionary rate of *S. Typhi* in different ecological niches but also to understand bacterial adaptation during carriage. Despite these obvious limitations, this was the first ever genomic investigation of the genetic characteristics of *S. Typhi* carriage isolates and the phylogenetic relatedness between carrier and acute isolates circulating over the same time period in a typhoid-endemic area. This study also provided valuable evidence for assessing the role of typhoid carriers in disease transmission in Kathmandu, Nepal.

In conclusion, typhoid carriage is not associated with any particular genotype nor driven by fluoroquinolone resistance. Additionally, I found strong evidence that typhoid carriers are likely not an important source of new infections in endemic settings such as Kathmandu. As a result, public health control measures should focus on providing people with safe water, food safety and vaccination for disease prevention. *S. Typhi* is exposed
to a variety of stressful conditions within the gallbladder and undergoes distinctive evolutionary processes for better adaptation.
Chapter 7

General discussion

The emergence and spread of a novel subclade of ciprofloxacin-resistant H58 S. Typhi in Nepal and a neighboring country demonstrates that fluoroquinolone resistance in this pathogen has become a serious problem across South Asia. More importantly, such organisms were found to be associated with fluoroquinolone treatment failure. My data strongly advocate that fluoroquinolones should no longer be used as empirical therapy for typhoid in this region; therefore, more clinical studies should be conducted to evaluate the effectiveness of alternative treatment options. Additionally, with the first typhoid conjugate vaccine Typbar-TCV that can be used for children aged less than two years and has been prequalified by WHO, it is necessary to conduct vaccine trials to measure the efficacy and immunogenicity of this new vaccine in Nepal and to develop a sustainable vaccine procurement and financing mechanisms. Programmatic use of the new typhoid vaccine targeting high-risk populations may not only reduce the disease burden but also minimize the antimicrobial use and selective pressure in the bacterial population.

Meanwhile, molecular typing and phenotypic characterization of S. Typhi needs to be performed routinely in South Asian countries and other reference laboratories outside this region to keep track of the spread of this novel H58 subclade to provide early warning for public health authorities and healthcare providers. These potentially dangerous organisms are likely to be widespread across the Indian subcontinent and may become a global health threat if they reach other vulnerable populations in Asia and Africa, as we have seen previously with the other H58 S. Typhi.
Despite the progress in *S. Typhi* vaccine development, there is still no vaccine against *S. Paratyphi* A. Further, the epidemiological characteristics and burden of disease caused by *S. Paratyphi* A are not well described, and the dynamics of bacterial populations in endemic areas are often overlooked. Using a genomic approach combined with conventional epidemiological tools, my study provided a detailed characterization of the phylogenetic structure and spatiotemporal distribution of *S. Paratyphi* A isolates from Kathmandu. This Nepalese *S. Paratyphi* A population was highly dynamic with evidence of regular bacterial transmission between Nepal and neighboring countries, resulting in clonal expansions of distinct genotypes at different time periods. Lineage A was the most common genotype found in this collection of *S. Paratyphi* A, which consisted of two dominant sub-lineages (A1 and A2), whereby a single sub-lineage (A2) has rapidly replaced all other preexisting genotypes from 2011. The emergence and rapid lineage replacement of sublineage A2 in Nepal is an intriguing phenomenon and requires further research into its virulence and antigenicity. For example, in my laboratory we are working on the characterization of LPS structure, epithelial cellular invasion and macrophage killing of *S. Paratyphi* A isolates belonging to different genotypes. Further, human protective immunity against differing *S. Paratyphi* A genotypes as well as the clinical and epidemiological features associated with infections caused by sub-lineage A2 need to be further investigated. The introduction of the new typhoid conjugate vaccine against *S. Typhi* in areas where *S. Typhi* and *S. Paratyphi* A are co-circulating may additionally lead to an increase in *S. Paratyphi* A infections. While *S. Paratyphi* A infections have become a growing problem in many parts of Asia, vaccines against this
pathogen are lagging far behind. Therefore, routine surveillance and improved public health practices in endemic areas can be an immediate solution for disease control and management.

Most previous typhoid studies have primarily focused on urban slums in low and middle-income countries where high rates of typhoid fever have often been reported. There is a general lack of epidemiological data about the disease burden and antimicrobial resistance of S. Typhi in rural settings with limited access to healthcare. My study provided unprecedented insights into the S. Typhi population structure and epidemiological features of typhoid fever in rural areas in Siem Reap province, Cambodia. Typhoid fever is widespread in rural areas of Cambodia, causing a significant disease burden in children aged less than 15 years. Several spatiotemporal outbreaks were identified in communes located near Tonle Sap Lake and proximity to the lake was associated with increased risk of infection. There was also a wide geographic distribution and high prevalence of MDR H58 S. Typhi with reduced susceptibility to fluoroquinolone in this setting. Despite these novel findings, the study was based on hospital surveillance data, which was limited to a single healthcare setting and lacked complete geospatial data. Therefore, community surveillance at individual or household-level in communes with high typhoid fever incidence (such as Kaoh Chiveang and Kampong Kleang) is necessary to identify specific risk behaviors and spatiotemporal case clusters. Furthermore, the level of fecal contamination and the presence of S. Typhi in water resources in these areas should also be examined. Such valuable information would be essential for guiding public health interventions as well as identifying high-risk
populations for typhoid vaccine introduction, given the fact that the new typhoid conjugate vaccine has been prequalified and registered in Cambodia.

The scope of my study also aimed to address one of the most important typhoid research questions regarding to role of typhoid carriage in disease transmission in endemic settings. My genomic and phylogenetic analyses of S. Typhi and S. Paratyphi A isolates recovered from asymptomatic carriers and acute patients suggested that typhoid carriers are not likely to contribute significantly to new infections in endemic area or the maintenance of bacterial genetic diversity. S. Typhi is likely to be under stress within the gallbladder and forge a distinct evolutionary pathway for better adaptation within the gallbladder. One of the important limitations of my study was a lack of information on the duration of carriage, which is crucial to estimate the mutation rate of S. Typhi during the carrier state and to better understand the phylogenetic relationships between carriage and acute isolates. However, it is very difficult to identify chronic typhoid carriers prospectively considering the very low rate of chronic carriage, intermittent fecal shedding, and a lack of a robust detection method. To follow up on the findings described in my thesis, studies are currently being performed in my laboratory to further characterize and compare the capacity to invade epithelial cells, kill macrophages and form biofilms between the acute and carrier isolates belonging to the same lineages. Such experiments are necessary to better understand how S. Typhi behave within the gallbladder environment during the carrier state. Additional research is also being conducted to identify specific biological markers associated with chronic carriage that can be used to develop a diagnostic test.
In conclusion, the advance of next-generation sequencing has revolutionized epidemiological research and bacterial genomics has become an essential tool for understanding the circulation of pathogens and the impact of antimicrobial resistant organisms on disease outcomes. Using genomic approaches combined with other clinical and epidemiological data, my research contributed significantly to a better understanding of the emergence of antimicrobial resistance and epidemiological features of typhoid fever in both rural and urban settings. For the first time, I measured an effect of AMR and bacterial genotype on the treatment outcome of typhoid fever, which is likely to have a substantial impact on clinical and public health practices.

In conclusion, my thesis strongly advocates the discontinued use of fluoroquinolones as empirical treatment for typhoid fever in South Asia and highlights the necessity for alternative antimicrobial therapies, as well as the introduction of new conjugate typhoid vaccines in this region. Further molecular investigations on disease transmission and AMR surveillance are required to routinely monitor the emergence and spread of these clinically important pathogens. The research methods described in my thesis have proven to be powerful, adaptable, robust and should be applied for similar studies on alternative pathogens in the future. Only by understanding the interaction of antimicrobial treatment and the infecting organism can we begin to curtail the current antimicrobial resistance crisis.
Chapter 8

References


30 Stuart BM, Pullen RL. Typhoid: Clinical Analysis of Three Hundred and Sixty Cases. *Arch Intern Med* 1946; **78**: 629–61.


36 STUART BM, PULLEN RL. Typhoid; clinical analysis of 360 cases. *Arch Intern Med (Chic)* 1946; **78**: 629–61.


49 Levine MM, Tacket CO, Sztein MB. Host-*Salmonella* interaction: Human trials. *Microbes Infect*


Khocharo HK. A comparative study of the typhidot (Dot-EIA) and Widal tests in blood culture positive cases of typhoid fever. *Trop Doct* 2011; 41: 136–8.


DOI:10.7554/eLife.03100.


Wait DA, Sobsey MD. Comparative survival of enteric viruses and bacteria in Atlantic Ocean


118 Crump J a, Crump J a, Luby SP, Luby SP, Mintz ED, Mintz ED. The global burden of typhoid fever. *Bull World Health Organ* 2004; **82**: 346–53.


120 Levine MM, Black RE, Lanata C. Precise estimation of the numbers of chronic carriers of
Salmonella typhi in Santiago, Chile, an endemic area. *J Infect Dis* 1982; **146**: 724–6.


132 Holt KE, Dolecek C, Chau TT, et al. Temporal Fluctuation of Multidrug Resistant *Salmonella*


154 WOODWARD TE, SMADEL JE, LEY HL, GREEN R, MANKIKAR DS. PRELIMINARY


166 Kariuki S, Gilks C, Revathi G, Hart CA. Genotypic analysis of multidrug-resistant Salmonella


177 Pokharel BM, Koirala J, Dahal RK, Mishra SK, Khadga PK, Tuladhar NR. Multidrug-resistant and extended-spectrum beta-lactamase (ESBL)-producing Salmonella enterica (serotypes Typhi and


Munir T, Lodhi M, Ansari JK, Andleeb S, Ahmed M. Extended Spectrum Beta Lactamase producing Cephalosporin resistant Salmonella Typhi, reported from Rawalpindi, Pakistan. J Pak


McClelland M, Sanderson KE, Clifton SW, et al. Comparison of genome degradation in Paratyphi


243 Kidgell C, Reichard U, Wain J, et al. Salmonella typhi, the causative agent of typhoid fever, is


Lu T, Zhao X, Drlica K. Gatifloxacin activity against quinolone-resistant gyrase: allele-specific

293 CLSI. Clinical and Laboratory Standards Institute: Performance standards for antimicrobial susceptibility testing; twentieth informational supplement. 2012.


304 Zhao W, Moest T, Zhao Y, et al. The Salmonella effector protein SifA plays a dual role in

305 Liu Y, Ho KK, Su J, Gong H, Chang AC, Lu S. Potassium transport of Salmonella is important for type III secretion and pathogenesis. Microbiol (United Kingdom) 2013; 159: 1705–19.


318 Crump JA. Building the case for wider use of typhoid vaccines. *Vaccine* 2015; **33**: C1–2.


320 Carrier of typhoid fever-1912. ; i.

321 Gonzalez-Escobedo G, Marshall JM, Gunn JS. Chronic and acute infection of the gall bladder by Salmonella Typhi: understanding the carrier state. DOI:10.1038/nrmicro2490.


324 Prouty AM, Schwesinger WH, Gunn JS. Biofilm formation and interaction with the surfaces of gallstones by Salmonella spp. *Infect Immun* 2002; **70**: 2640–9.


1703–13.


DOI:10.1038/srep07457.


341 Steinsiek S, Bettenbrock K. Glucose transport in Escherichia coli mutant strains with defects in


### Appendix A  *Salmonella* Typhi isolates and their corresponding sequencing metadata in chapter 3

<table>
<thead>
<tr>
<th>Name</th>
<th>Tree ID</th>
<th>Accession number</th>
<th>Trial</th>
<th>Year</th>
<th>Country</th>
<th>gyrA</th>
<th>gyrA</th>
<th>gyrA</th>
<th>gyrB</th>
<th>gyrB</th>
<th>parC</th>
<th>parE</th>
<th>parE</th>
<th>parE</th>
<th>MIC</th>
<th>MIC group*</th>
</tr>
</thead>
<tbody>
<tr>
<td>02TY_001</td>
<td>STY1_AL513382</td>
<td>ERR1079230</td>
<td>1</td>
<td>2011</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>1</td>
</tr>
<tr>
<td>02TY_006</td>
<td>STY6_AL513382</td>
<td>ERR1079231</td>
<td>1</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.38</td>
<td>2</td>
</tr>
<tr>
<td>02TY_007</td>
<td>STY7_AL513382</td>
<td>ERR1079232</td>
<td>1</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>02TY_009</td>
<td>STY9_AL513382</td>
<td>ERR1079233</td>
<td>1</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.38</td>
<td>2</td>
</tr>
<tr>
<td>02TY_015</td>
<td>STY15_MERGE_AL513382</td>
<td>ERR1079234</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.38</td>
<td>2</td>
</tr>
<tr>
<td>02TY_016</td>
<td>STY16_AL513382</td>
<td>ERR1079235</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_018</td>
<td>STY18_AL513382</td>
<td>ERR1079236</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_027</td>
<td>STY27_MERGE_AL513382</td>
<td>ERR1079237</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_028</td>
<td>STY28_MERGE_AL513382</td>
<td>ERR1079238</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_030</td>
<td>STY30_AL513382</td>
<td>ERR1079239</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_031</td>
<td>STY31_AL513382</td>
<td>ERR1079240</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_032</td>
<td>STY32_AL513382</td>
<td>ERR1079241</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>02TY_034</td>
<td>STY34_AL513382</td>
<td>ERR1079242</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>02TY_037</td>
<td>STY37_AL513382</td>
<td>ERR1079243</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>02TY_038</td>
<td>STY38_AL513382</td>
<td>ERR1079244</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>02TY_045</td>
<td>STY45_MERGE_AL513382</td>
<td>ERR1079245</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>02TY_047</td>
<td>STY47_AL513382</td>
<td>ERR1079246</td>
<td>1</td>
<td>2012</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>----------</td>
<td>-------------</td>
<td>------</td>
<td>---------</td>
<td>--------</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>-----</td>
</tr>
<tr>
<td>02TY_048</td>
<td>STY48 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>02TY_051</td>
<td>STY51 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_055</td>
<td>STY55 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_057</td>
<td>STY57 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_058</td>
<td>STY58D1 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_059</td>
<td>STY59 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_060</td>
<td>STY60 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>02TY_061</td>
<td>STY61 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.012</td>
<td></td>
</tr>
<tr>
<td>02TY_078</td>
<td>14STY78 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td></td>
</tr>
<tr>
<td>02TY_080</td>
<td>STY80 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>02TY_083</td>
<td>STY83 AL513382</td>
<td>2012</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
<td></td>
</tr>
<tr>
<td>02TY_119</td>
<td>15STY119D1 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_125</td>
<td>STY125 MERGE AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>02TY_126</td>
<td>6STY126 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_127</td>
<td>7STY127 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td>02TY_132</td>
<td>15STY132 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>02TY_135</td>
<td>16STY135 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_136</td>
<td>STY136 MERGE AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>02TY_140</td>
<td>16STY140 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.023</td>
<td></td>
</tr>
<tr>
<td>02TY_143</td>
<td>9STY143D26 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_144</td>
<td>10STY144 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>02TY_146</td>
<td>11STY146 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_147</td>
<td>17STY147 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_148</td>
<td>17STY148 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>02TY_149</td>
<td>12STY149 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>02TY_150</td>
<td>13STY150 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td>02TY_156</td>
<td>14STY156 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>02TY_157</td>
<td>18STY157 AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.008</td>
<td></td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
<td>Year</td>
<td>Country</td>
<td>N</td>
<td>L</td>
<td>C</td>
<td>Y1</td>
<td>Y2</td>
<td>Y3</td>
<td>Y4</td>
<td>Y5</td>
<td>Y12</td>
<td>Y15</td>
<td>Y20</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------</td>
<td>------</td>
<td>---------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02TY_158</td>
<td>15STY158_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td></td>
<td></td>
</tr>
<tr>
<td>02TY_159</td>
<td>16STY159_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_160</td>
<td>17STY160_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_163</td>
<td>18STY163_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_164</td>
<td>19STY164_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>02TY_169</td>
<td>STY169_MERGE_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>24</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_173</td>
<td>STY173_MERGE_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>02TY_174</td>
<td>19STY174_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_176</td>
<td>22STY176D1_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
</tr>
<tr>
<td>02TY_178</td>
<td>24STY178_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_183</td>
<td>13254_5637</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
</tr>
<tr>
<td>02TY_185</td>
<td>25STY185_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.15</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>02TY_186</td>
<td>26STY186_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>02TY_188</td>
<td>STY188_MERGE_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_190</td>
<td>13254_5639</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>24</td>
<td>3</td>
</tr>
<tr>
<td>02TY_195</td>
<td>STY195_MERGE_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>02TY_196</td>
<td>29STY196_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>02TY_200</td>
<td>30STY200_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>02TY_202</td>
<td>20STY202_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>02TY_210</td>
<td>18STY210_AL513382</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_213</td>
<td>24STY213_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.016</td>
<td>1</td>
</tr>
<tr>
<td>02TY_216</td>
<td>31STY216_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>02TY_219</td>
<td>21STY219D1_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_222</td>
<td>23STY222_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0.38</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>02TY_226</td>
<td>32STY226_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>02TY_229</td>
<td>20STY229_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_232</td>
<td>STY232_MERGE_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>02TY_235</td>
<td>24STY235D1_AL513382</td>
<td>2014</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Description</td>
<td>Code</td>
<td>Country</td>
<td>Year</td>
<td>Result 1</td>
<td>Result 2</td>
<td>Result 3</td>
<td>Result 4</td>
<td>Result 5</td>
<td>Result 6</td>
<td>Result 7</td>
<td>Result 8</td>
<td>Result 9</td>
<td>Result 10</td>
<td>Result 11</td>
<td>Result 12</td>
</tr>
<tr>
<td>-------</td>
<td>-------------</td>
<td>------</td>
<td>---------</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
<td>-----------</td>
</tr>
<tr>
<td>02TY_236</td>
<td>STY236_MERGE_AL513382</td>
<td>ERR1079305</td>
<td>Nepal</td>
<td>2014</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>02TY_239</td>
<td>28STY239_AL513382</td>
<td>ERR1079306</td>
<td>Nepal</td>
<td>2014</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>02TY_240</td>
<td>29STY240_AL513382</td>
<td>ERR1079307</td>
<td>Nepal</td>
<td>2014</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>02TY_241</td>
<td>30STY241_AL513382</td>
<td>ERR1079308</td>
<td>Nepal</td>
<td>2014</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.5</td>
</tr>
<tr>
<td>02TY_244</td>
<td>12STY244_AL513382</td>
<td>ERR1079309</td>
<td>Nepal</td>
<td>2014</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>01TY072</td>
<td>5886_3/1</td>
<td>ERR119817</td>
<td>Nepal</td>
<td>2008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY075</td>
<td>5886_3/2</td>
<td>ERR119818</td>
<td>Nepal</td>
<td>2008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY090</td>
<td>5886_3/4</td>
<td>ERR119820</td>
<td>Nepal</td>
<td>2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY098</td>
<td>5886_3/5</td>
<td>ERR119821</td>
<td>Nepal</td>
<td>2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY101</td>
<td>5886_3/6</td>
<td>ERR119822</td>
<td>Nepal</td>
<td>2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY103</td>
<td>5886_3/8</td>
<td>ERR119824</td>
<td>Nepal</td>
<td>2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>01TY112</td>
<td>5886_3/9</td>
<td>ERR119825</td>
<td>Nepal</td>
<td>2009</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CC_04</td>
<td>10561_2/1</td>
<td>ERR357576</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CC_05</td>
<td>10561_2/8</td>
<td>ERR357583</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CC_08</td>
<td>10561_2/12</td>
<td>ERR357587</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>CC_11</td>
<td>10561_2/19</td>
<td>ERR357594</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_12</td>
<td>10561_2/25</td>
<td>ERR357600</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CC_13</td>
<td>10561_2/31</td>
<td>ERR357606</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_16</td>
<td>10561_2/37</td>
<td>ERR357612</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CC_22</td>
<td>10561_2/44</td>
<td>ERR357619</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>CC_27</td>
<td>10425_1/9</td>
<td>ERR349339</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.38</td>
</tr>
<tr>
<td>CC_31</td>
<td>10561_2/53</td>
<td>ERR357628</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_39</td>
<td>10561_2/57</td>
<td>ERR357632</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_40</td>
<td>10561_2/62</td>
<td>ERR357637</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.38</td>
</tr>
<tr>
<td>CC_41</td>
<td>10561_2/82</td>
<td>ERR357577</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.125</td>
</tr>
<tr>
<td>CC_43</td>
<td>10561_2/99</td>
<td>ERR357584</td>
<td>Nepal</td>
<td>2011</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0.38</td>
</tr>
<tr>
<td>CC_44</td>
<td>10561_2/13</td>
<td>ERR357588</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_45</td>
<td>10425_1/5</td>
<td>ERR349335</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0.008</td>
</tr>
<tr>
<td>CC_49</td>
<td>10561_2/26</td>
<td>ERR357601</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_50</td>
<td>10561_2/32</td>
<td>ERR357607</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.125</td>
</tr>
<tr>
<td>CC_51</td>
<td>10561_2/38</td>
<td>ERR357613</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>CC_57</td>
<td>10561_2/45</td>
<td>ERR357620</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.38</td>
</tr>
<tr>
<td>CC_60</td>
<td>10561_2/50</td>
<td>ERR357625</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_61</td>
<td>10426_1/6</td>
<td>ERR349528</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_65</td>
<td>10561_2/58</td>
<td>ERR357633</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.125</td>
</tr>
<tr>
<td>CC_67</td>
<td>10561_2/63</td>
<td>ERR357638</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>CC_69</td>
<td>10425_1/2</td>
<td>ERR349332</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.125</td>
</tr>
<tr>
<td>CC_74</td>
<td>10425_1/3</td>
<td>ERR349333</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>CC_75</td>
<td>10561_2/14</td>
<td>ERR357589</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
</tr>
<tr>
<td>CC_80</td>
<td>10425_1/6</td>
<td>ERR349336</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_81</td>
<td>10561_2/27</td>
<td>ERR357602</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_83</td>
<td>10561_2/33</td>
<td>ERR357608</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>CC_87</td>
<td>10561_2/39</td>
<td>ERR357614</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.25</td>
<td>2</td>
</tr>
<tr>
<td>CC_88</td>
<td>10561_2/46</td>
<td>ERR357621</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_89</td>
<td>10561_2/51</td>
<td>ERR357626</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_90</td>
<td>10561_2/54</td>
<td>ERR357629</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_92</td>
<td>10561_2/59</td>
<td>ERR357634</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_93</td>
<td>10561_2/64</td>
<td>ERR357639</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_94</td>
<td>10561_2/3</td>
<td>ERR357578</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_95</td>
<td>10561_2/10</td>
<td>ERR357585</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
</tr>
<tr>
<td>CC_97</td>
<td>10561_2/15</td>
<td>ERR357590</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.19</td>
<td>2</td>
</tr>
<tr>
<td>CC_99</td>
<td>10561_2/20</td>
<td>ERR357595</td>
<td>0</td>
<td>2011</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0.094</td>
<td>2</td>
</tr>
<tr>
<td>Ind082356_2008</td>
<td>10349_1/34</td>
<td>ERR343282</td>
<td>0</td>
<td>2008</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Ind101104_2010</td>
<td>10349_1/52</td>
<td>ERR343300</td>
<td>0</td>
<td>2010</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>32</td>
</tr>
<tr>
<td>Code</td>
<td>Accession</td>
<td>ERR</td>
<td>Year</td>
<td>Country</td>
<td>Species</td>
<td>Genotype</td>
<td>Susceptibility Group</td>
<td>MIC</td>
<td>MDR</td>
<td>Other Key</td>
<td>Comments</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>-----------</td>
<td>----------------------</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
<td>----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndBCR175_2011</td>
<td>8616_4/48</td>
<td>ERR420424</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 1 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndBCR191_2011</td>
<td>8616_4/50</td>
<td>ERR420426</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndBCR211_2011</td>
<td>8616_4/51</td>
<td>ERR420427</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndBCR62_2009</td>
<td>8616_4/41</td>
<td>ERR420417</td>
<td>2009</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST127_2011</td>
<td>10492_166</td>
<td>ERR352259</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST128_2011</td>
<td>10493_169</td>
<td>ERR352434</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST147_2012</td>
<td>10492_16/13</td>
<td>ERR352266</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST167_2011</td>
<td>10562_2/13</td>
<td>ERR357768</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>24 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST197_2010</td>
<td>10492_1/25</td>
<td>ERR352278</td>
<td>2010</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST216_2011</td>
<td>10492_1/31</td>
<td>ERR352284</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST247_2011</td>
<td>10562_2/32</td>
<td>ERR357787</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST248_2011</td>
<td>10492_1/43</td>
<td>ERR352296</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>24 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST254_2010</td>
<td>10562_2/34</td>
<td>ERR357789</td>
<td>2010</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST400_2012</td>
<td>10562_2/79</td>
<td>ERR357834</td>
<td>2012</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>24 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndMDUST408_2012</td>
<td>10492_1/82</td>
<td>ERR352335</td>
<td>2012</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IndSP80_2011</td>
<td>8616_4/58</td>
<td>ERR420434</td>
<td>2011</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEN_2512</td>
<td>13254_5/40</td>
<td>ERR586915</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEN_2533</td>
<td>2533_AL513382</td>
<td>ERR1079310</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>0.38 2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NEN_2546</td>
<td>2546_AL513382</td>
<td>ERR1079311</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1 0 0 0 1 0 0 1 0 0</td>
<td>6 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_510</td>
<td>13254_5/42</td>
<td>ERR586917</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_514</td>
<td>13254_5/43</td>
<td>ERR586918</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1 0 0 0 1 0 1 0 0 0</td>
<td>8 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_515</td>
<td>13254_5/44</td>
<td>ERR586919</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_516</td>
<td>13254_5/45</td>
<td>ERR586920</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1 0 0 0 1 0 1 0 0 0</td>
<td>8 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_521</td>
<td>13254_5/46</td>
<td>ERR586921</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>8 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTHERS_535</td>
<td>535_AL513382</td>
<td>ERR1079312</td>
<td>2013</td>
<td>Nepal</td>
<td>1</td>
<td>0</td>
<td>1 0 0 0 1 0 1 0 0 0</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>UnkMDUST146_2012</td>
<td>10562_2/10</td>
<td>ERR357765</td>
<td>2012</td>
<td>Unknown</td>
<td>1</td>
<td>1</td>
<td>0 0 0 0 1 0 0 0 0 1</td>
<td>32 3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ciprofloxacin MIC group: 1; susceptible, 2; intermediate, 3; resistant
## Appendix B  *Salmonella* Paratyphi A isolates and their corresponding metadata in chapter 4

<table>
<thead>
<tr>
<th>Strain</th>
<th>Acute/Carrier</th>
<th>Lineage</th>
<th>Genotype</th>
<th>Lat</th>
<th>Long</th>
<th>GyrA</th>
<th>Date</th>
<th>Age</th>
<th>Sex</th>
<th>Main_water</th>
<th>Treat_water</th>
<th>Anorexia</th>
<th>Nausea</th>
<th>Abdominal pain</th>
<th>Diarrhoea</th>
<th>Constipation</th>
<th>Headache</th>
</tr>
</thead>
<tbody>
<tr>
<td>DM035</td>
<td>Acute</td>
<td>A</td>
<td>A1.6</td>
<td>27.69116</td>
<td>85.34435</td>
<td>S83F</td>
<td>17-Jun-05</td>
<td>23</td>
<td>F</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM038</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.65838</td>
<td>85.32773</td>
<td>S83F</td>
<td>19-Jun-05</td>
<td>25</td>
<td>M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM042</td>
<td>Acute</td>
<td>A</td>
<td>A1.6</td>
<td>27.6898</td>
<td>85.34475</td>
<td>S83F</td>
<td>20-Jun-05</td>
<td>30</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM049</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.68463</td>
<td>85.32151</td>
<td>WT</td>
<td>21-Jun-05</td>
<td>22</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM056</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.69708</td>
<td>85.33345</td>
<td>S83F</td>
<td>25-Jun-05</td>
<td>11</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM067</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.6692</td>
<td>85.32035</td>
<td>S83F</td>
<td>27-Jun-05</td>
<td>16</td>
<td>M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM072</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.67321</td>
<td>85.32898</td>
<td>S83F</td>
<td>28-Jun-05</td>
<td>31</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM080</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.67098</td>
<td>85.3219</td>
<td>S83F</td>
<td>29-Jun-05</td>
<td>6</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM082</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.67166</td>
<td>85.32711</td>
<td>S83F</td>
<td>29-Jun-05</td>
<td>50</td>
<td>F</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM089</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.67448</td>
<td>85.31856</td>
<td>S83F</td>
<td>30-Jun-05</td>
<td>13</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM093</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.676</td>
<td>85.33893</td>
<td>WT</td>
<td>30-Jun-05</td>
<td>21</td>
<td>F</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM098</td>
<td>Acute</td>
<td>A</td>
<td>A1.8</td>
<td>27.67525</td>
<td>85.33198</td>
<td>S83F</td>
<td>3-Jul-05</td>
<td>18</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM111</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.65938</td>
<td>85.32823</td>
<td>S83F</td>
<td>6-Jul-05</td>
<td>21</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM116</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.6421</td>
<td>85.32248</td>
<td>S83F</td>
<td>8-Jul-05</td>
<td>32</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM118</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.69901</td>
<td>85.35241</td>
<td>S83F</td>
<td>8-Jul-05</td>
<td>30</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM158</td>
<td>Acute</td>
<td>A</td>
<td>A1.6</td>
<td>27.68921</td>
<td>85.34438</td>
<td>S83F</td>
<td>16-Jul-05</td>
<td>29</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM175</td>
<td>Acute</td>
<td>A</td>
<td>A1.10</td>
<td>27.67495</td>
<td>85.30568</td>
<td>S83F</td>
<td>17-Jul-05</td>
<td>13</td>
<td>F</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM178</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66806</td>
<td>85.3289</td>
<td>S83F</td>
<td>19-Jul-05</td>
<td>8</td>
<td>F</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM183</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.67451</td>
<td>85.31931</td>
<td>S83F</td>
<td>19-Jul-05</td>
<td>22</td>
<td>F</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM186</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.67746</td>
<td>85.33521</td>
<td>S83F</td>
<td>20-Jul-05</td>
<td>15</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM188</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.66326</td>
<td>85.32745</td>
<td>D87Y</td>
<td>20-Jul-05</td>
<td>28</td>
<td>M</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM192</td>
<td>Acute</td>
<td>A</td>
<td>A1.10</td>
<td>27.65668</td>
<td>85.3298</td>
<td>S83F</td>
<td>21-Jul-05</td>
<td>14</td>
<td>F</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM210</td>
<td>Acute</td>
<td>A</td>
<td>A1.10</td>
<td>27.6751</td>
<td>85.30683</td>
<td>S83F</td>
<td>26-Jul-05</td>
<td>20</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM211</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.67746</td>
<td>85.33521</td>
<td>S83F</td>
<td>26-Jul-05</td>
<td>16</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>DM219</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66865</td>
<td>85.32715</td>
<td>S83F</td>
<td>28-Jul-05</td>
<td>10</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM227</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66311</td>
<td>85.31691</td>
<td>SK3F</td>
<td>31-Jul-05</td>
<td>14</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM233</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.66948</td>
<td>85.32518</td>
<td>SK3F</td>
<td>31-Jul-05</td>
<td>4</td>
<td>F</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM247</td>
<td>Acute</td>
<td>A</td>
<td>A1.8</td>
<td>27.67205</td>
<td>85.32166</td>
<td>SK3F</td>
<td>2-Aug-05</td>
<td>20</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM253</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.66886</td>
<td>85.32515</td>
<td>SK3F</td>
<td>4-Aug-05</td>
<td>21</td>
<td>M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM254</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66913</td>
<td>85.32666</td>
<td>SK3F</td>
<td>4-Aug-05</td>
<td>12</td>
<td>F</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM256</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.67683</td>
<td>85.3219</td>
<td>SK3F</td>
<td>4-Aug-05</td>
<td>17</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM265</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66198</td>
<td>85.32601</td>
<td>SK3F</td>
<td>5-Aug-05</td>
<td>14</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM274</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.68061</td>
<td>85.32708</td>
<td>SK3F</td>
<td>8-Aug-05</td>
<td>23</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM295</td>
<td>Acute</td>
<td>A</td>
<td>A1.8</td>
<td>27.68211</td>
<td>85.3234</td>
<td>SK3F</td>
<td>12-Aug-05</td>
<td>23</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM305</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67885</td>
<td>85.32426</td>
<td>SK3F</td>
<td>17-Aug-05</td>
<td>20</td>
<td>M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM313</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.67413</td>
<td>85.30495</td>
<td>WT</td>
<td>19-Aug-05</td>
<td>17</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM323</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.67336</td>
<td>85.32105</td>
<td>SK3F</td>
<td>22-Aug-05</td>
<td>30</td>
<td>M</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM329</td>
<td>Acute</td>
<td>A</td>
<td>A1.10</td>
<td>27.67471</td>
<td>85.30631</td>
<td>SK3F</td>
<td>25-Aug-05</td>
<td>30</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM356</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.67951</td>
<td>85.32698</td>
<td>SK3F</td>
<td>30-Aug-05</td>
<td>24</td>
<td>M</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM358</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.64466</td>
<td>85.3157</td>
<td>SK3F</td>
<td>31-Aug-05</td>
<td>15</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM363</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.67413</td>
<td>85.30495</td>
<td>WT</td>
<td>1-Sep-05</td>
<td>16</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM387</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.68</td>
<td>85.3272</td>
<td>SK3F</td>
<td>7-Sep-05</td>
<td>19</td>
<td>M</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DM390</td>
<td>Acute</td>
<td>A</td>
<td>A1.8</td>
<td>27.67256</td>
<td>85.3215</td>
<td>SK3F</td>
<td>8-Sep-05</td>
<td>25</td>
<td>M</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED027</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.66598</td>
<td>85.31043</td>
<td>SK3F</td>
<td>20-May-06</td>
<td>22</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED034</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.6673</td>
<td>85.33213</td>
<td>SK3F</td>
<td>22-May-06</td>
<td>21</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED035</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67546</td>
<td>85.3212</td>
<td>SK3F</td>
<td>22-May-06</td>
<td>10</td>
<td>F</td>
<td>Municipal</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED037</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.6748</td>
<td>85.32711</td>
<td>SK3F</td>
<td>23-May-06</td>
<td>12</td>
<td>M</td>
<td>Well</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>ED039</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67965</td>
<td>85.30196</td>
<td>SK3F</td>
<td>23-May-06</td>
<td>26</td>
<td>M</td>
<td>Municipal</td>
<td>Boil</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED043</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67363</td>
<td>85.32805</td>
<td>SK3F</td>
<td>24-May-06</td>
<td>8</td>
<td>M</td>
<td>Well</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED044</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.67761</td>
<td>85.34775</td>
<td>SK3F</td>
<td>24-May-06</td>
<td>21</td>
<td>M</td>
<td>Mix</td>
<td>Boil</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED051</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.67986</td>
<td>85.32321</td>
<td>SK3F</td>
<td>28-May-06</td>
<td>39</td>
<td>M</td>
<td>Municipal</td>
<td>Other</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED053</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.67731</td>
<td>85.33545</td>
<td>SK3F</td>
<td>28-May-06</td>
<td>19</td>
<td>M</td>
<td>Well</td>
<td>Chlorine</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Type</td>
<td>Acute</td>
<td>A</td>
<td>A1.4</td>
<td>27.67181</td>
<td>85.30208</td>
<td>SK3F</td>
<td>1-Jun-06</td>
<td>17</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>---</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>------</td>
<td>----------</td>
<td>----</td>
<td>---</td>
<td>-----------</td>
<td>--------</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>ED065</td>
<td>Acute</td>
<td>A</td>
<td>A1.4</td>
<td>27.66946</td>
<td>85.31273</td>
<td>SK3F</td>
<td>4-Jun-06</td>
<td>35</td>
<td>M</td>
<td>Municipal</td>
<td>Boil</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED070</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.6352</td>
<td>85.31878</td>
<td>SK3F</td>
<td>5-Jun-06</td>
<td>7</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED074</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.68463</td>
<td>85.32151</td>
<td>WT</td>
<td>6-Jun-06</td>
<td>4</td>
<td>F</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ED079</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67256</td>
<td>85.3323</td>
<td>SK3F</td>
<td>6-Jun-06</td>
<td>21</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED099</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.66757</td>
<td>85.33275</td>
<td>WT</td>
<td>10-Jun-06</td>
<td>28</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED116</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.65751</td>
<td>85.31991</td>
<td>SK3F</td>
<td>25-Jun-06</td>
<td>23</td>
<td>F</td>
<td>Well</td>
<td>Mix</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ED164</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67776</td>
<td>85.32176</td>
<td>SK3F</td>
<td>26-Jun-06</td>
<td>34</td>
<td>M</td>
<td>Municipal</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED168</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67761</td>
<td>85.32058</td>
<td>SK3F</td>
<td>27-Jun-06</td>
<td>13</td>
<td>F</td>
<td>Mix</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED179</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67891</td>
<td>85.32911</td>
<td>SK3F</td>
<td>2-Jul-06</td>
<td>21</td>
<td>F</td>
<td>Stone spout</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED182</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.67718</td>
<td>85.32768</td>
<td>SK3F</td>
<td>3-Jul-06</td>
<td>24</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED191</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.67318</td>
<td>85.32173</td>
<td>SK3F</td>
<td>5-Jul-06</td>
<td>13</td>
<td>M</td>
<td>Well</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED192</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67595</td>
<td>85.32538</td>
<td>SK3F</td>
<td>5-Jul-06</td>
<td>8</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED193</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.68211</td>
<td>85.3234</td>
<td>SK3F</td>
<td>5-Jul-06</td>
<td>20</td>
<td>M</td>
<td>Mix</td>
<td>Other</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED194</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67546</td>
<td>85.3212</td>
<td>SK3F</td>
<td>5-Jul-06</td>
<td>19</td>
<td>F</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED199</td>
<td>Acute</td>
<td>A</td>
<td>A1.7</td>
<td>27.66366</td>
<td>85.32826</td>
<td>WT</td>
<td>6-Jul-06</td>
<td>35</td>
<td>F</td>
<td>Stone spout</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED200</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.41299</td>
<td>85.1876</td>
<td>SK3F</td>
<td>23-Jul-06</td>
<td>7</td>
<td>F</td>
<td>Municipal</td>
<td>Chlorine</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED221</td>
<td>Acute</td>
<td>A</td>
<td>A1.4</td>
<td>27.66781</td>
<td>85.32068</td>
<td>SK3F</td>
<td>23-Jul-06</td>
<td>26</td>
<td>M</td>
<td>Tanker</td>
<td>Untreated</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ED226</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66475</td>
<td>85.32506</td>
<td>SK3F</td>
<td>24-Jul-06</td>
<td>9</td>
<td>M</td>
<td>Well</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED227</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.68061</td>
<td>85.32708</td>
<td>SK3F</td>
<td>25-Jul-06</td>
<td>18</td>
<td>M</td>
<td>Stone spout</td>
<td>Chlorine</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED228</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.67761</td>
<td>85.32176</td>
<td>SK3F</td>
<td>26-Jul-06</td>
<td>27</td>
<td>F</td>
<td>Well</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED229</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.68531</td>
<td>85.31443</td>
<td>SK3F</td>
<td>26-Jul-06</td>
<td>17</td>
<td>F</td>
<td>Municipal</td>
<td>Boil</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED230</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.66928</td>
<td>85.32481</td>
<td>SK3F</td>
<td>28-Jul-06</td>
<td>26</td>
<td>M</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED239</td>
<td>Acute</td>
<td>A</td>
<td>A1.3</td>
<td>27.67653</td>
<td>85.32718</td>
<td>SK3F</td>
<td>30-Jul-06</td>
<td>24</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ED317</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.86355</td>
<td>85.31371</td>
<td>SK3F</td>
<td>6-Aug-06</td>
<td>33</td>
<td>F</td>
<td>Mix</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.67208</td>
<td>85.34118</td>
<td>SSF</td>
<td>6-Aug-06</td>
<td>24</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>---</td>
<td>------</td>
<td>---------</td>
<td>----------</td>
<td>-----</td>
<td>----------</td>
<td>----</td>
<td>---</td>
<td>--------------</td>
<td>------------</td>
<td>-----</td>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>ED320</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.65438</td>
<td>85.32065</td>
<td>SSF</td>
<td>6-Aug-06</td>
<td>20</td>
<td>M</td>
<td>Mix</td>
<td>Chlorine</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED321</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67246</td>
<td>85.32088</td>
<td>SSF</td>
<td>7-Aug-06</td>
<td>26</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED327</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.67821</td>
<td>85.31538</td>
<td>SSF</td>
<td>9-Aug-06</td>
<td>12</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED330</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66311</td>
<td>85.31691</td>
<td>SSF</td>
<td>10-Aug-06</td>
<td>17</td>
<td>F</td>
<td>Bottled</td>
<td>Other</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED333</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.67105</td>
<td>85.32501</td>
<td>SSF</td>
<td>13-Aug-06</td>
<td>24</td>
<td>F</td>
<td>Well</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED338</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.67106</td>
<td>85.3047</td>
<td>SSF</td>
<td>15-Aug-06</td>
<td>33</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED340</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.67303</td>
<td>85.32156</td>
<td>SSF</td>
<td>16-Aug-06</td>
<td>11</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED345</td>
<td>Acute</td>
<td>A</td>
<td>A1.4</td>
<td>27.68505</td>
<td>85.30958</td>
<td>SSF</td>
<td>17-Aug-06</td>
<td>25</td>
<td>M</td>
<td>Municipal</td>
<td>Mix</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED347</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.6657</td>
<td>85.32391</td>
<td>SSF</td>
<td>17-Aug-06</td>
<td>16</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED348</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66415</td>
<td>85.31873</td>
<td>SSF</td>
<td>18-Aug-06</td>
<td>23</td>
<td>F</td>
<td>Well</td>
<td>Other</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED356</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66821</td>
<td>85.3221</td>
<td>SSF</td>
<td>21-Aug-06</td>
<td>25</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED361</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67396</td>
<td>85.31955</td>
<td>SSF</td>
<td>22-Aug-06</td>
<td>9</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED362</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66913</td>
<td>85.32336</td>
<td>SSF</td>
<td>23-Aug-06</td>
<td>15</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED363</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.67048</td>
<td>85.31215</td>
<td>SSF</td>
<td>23-Aug-06</td>
<td>25</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED366</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67031</td>
<td>85.3251</td>
<td>SSF</td>
<td>23-Aug-06</td>
<td>24</td>
<td>F</td>
<td>Well</td>
<td>Boil</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>ED375</td>
<td>Acute</td>
<td>A</td>
<td>A1.5</td>
<td>27.6799</td>
<td>85.31151</td>
<td>SSF</td>
<td>28-Aug-06</td>
<td>22</td>
<td>M</td>
<td>Municipal</td>
<td>Chlorine</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED383</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66998</td>
<td>85.32285</td>
<td>SSF</td>
<td>31-Aug-06</td>
<td>3</td>
<td>M</td>
<td>Well</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED408</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.65583</td>
<td>85.33271</td>
<td>SSF</td>
<td>17-Dec-06</td>
<td>26</td>
<td>M</td>
<td>Stone spout</td>
<td>Boil</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED416</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.67343</td>
<td>85.3222</td>
<td>WT</td>
<td>4-Jan-07</td>
<td>19</td>
<td>F</td>
<td>Well</td>
<td>Boil</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED417</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.68121</td>
<td>85.32736</td>
<td>SSF</td>
<td>5-Jan-07</td>
<td>19</td>
<td>F</td>
<td>Stone spout</td>
<td>Mix</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>ED432</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.67593</td>
<td>85.31933</td>
<td>SSF</td>
<td>2-Mar-07</td>
<td>25</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED443</td>
<td>Acute</td>
<td>A</td>
<td>A1</td>
<td>27.68886</td>
<td>85.31315</td>
<td>SSF</td>
<td>16-Mar-07</td>
<td>24</td>
<td>M</td>
<td>Bottled</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED446</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68018</td>
<td>85.32763</td>
<td>D87Y</td>
<td>19-Mar-07</td>
<td>25</td>
<td>M</td>
<td>Stone spout</td>
<td>Boil</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED459</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.67735</td>
<td>85.32795</td>
<td>SSF</td>
<td>4-Apr-07</td>
<td>23</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED494</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66525</td>
<td>85.33801</td>
<td>SSF</td>
<td>6-May-07</td>
<td>22</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>ED503</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.68463</td>
<td>85.32151</td>
<td>SSF</td>
<td>16-May-07</td>
<td>23</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Acute</td>
<td>A1.11</td>
<td>27.669</td>
<td>85.32908</td>
<td>S83F</td>
<td>1-Jan-07</td>
<td>17</td>
<td>F</td>
<td>Mix</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>-----------</td>
<td>------</td>
<td>----------</td>
<td>----</td>
<td>---</td>
<td>-----</td>
<td>--------</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ED579</td>
<td>Acute</td>
<td>A1.11</td>
<td>27.6729</td>
<td>85.3215</td>
<td>S83F</td>
<td>11-Jul-07</td>
<td>16</td>
<td>M</td>
<td>Well</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED624</td>
<td>Acute</td>
<td>A1.11</td>
<td>27.68061</td>
<td>85.32708</td>
<td>S83F</td>
<td>10-Aug-07</td>
<td>20</td>
<td>M</td>
<td>Municipal</td>
<td>Chlorine</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED632</td>
<td>Acute</td>
<td>A11.1</td>
<td>27.670069</td>
<td>85.32481</td>
<td>S83F</td>
<td>20-Aug-07</td>
<td>33</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED646</td>
<td>Acute</td>
<td>A1.1</td>
<td>27.6724</td>
<td>85.32131</td>
<td>S83F</td>
<td>5-Sep-07</td>
<td>11</td>
<td>M</td>
<td>Well</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED680</td>
<td>Acute</td>
<td>A1.11</td>
<td>27.66795</td>
<td>85.29868</td>
<td>S83F</td>
<td>2-Jun-08</td>
<td>20</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>ED849</td>
<td>Acute</td>
<td>A1.4</td>
<td>27.67208</td>
<td>85.30555</td>
<td>S83F</td>
<td>22-Jul-08</td>
<td>15</td>
<td>F</td>
<td>Stone spout</td>
<td>Mix</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>GB113</td>
<td>Carrier</td>
<td>C</td>
<td>C</td>
<td>27.13393</td>
<td>85.15166</td>
<td>WT</td>
<td>17-Jan-08</td>
<td>45</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB193</td>
<td>Carrier</td>
<td>A</td>
<td>A2</td>
<td>27.699308</td>
<td>85.290304</td>
<td>WT</td>
<td>10-Mar-08</td>
<td>43</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB214</td>
<td>Carrier</td>
<td>F</td>
<td>F</td>
<td>27.67391</td>
<td>85.3218</td>
<td>WT</td>
<td>24-Mar-08</td>
<td>38</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB224</td>
<td>Carrier</td>
<td>C</td>
<td>C</td>
<td>27.67391</td>
<td>85.3218</td>
<td>WT</td>
<td>27-Mar-08</td>
<td>42</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB245</td>
<td>Carrier</td>
<td>A1.11</td>
<td>27.13733</td>
<td>85.00836</td>
<td>S83F</td>
<td>11-Apr-08</td>
<td>33</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB280</td>
<td>Carrier</td>
<td>A</td>
<td>A1</td>
<td>27.15806</td>
<td>85.15575</td>
<td>S83F</td>
<td>4-May-08</td>
<td>34</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB493</td>
<td>Carrier</td>
<td>A</td>
<td>A1</td>
<td>27.670389</td>
<td>85.437631</td>
<td>S83F</td>
<td>3-Oct-08</td>
<td>28</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB58</td>
<td>Carrier</td>
<td>A</td>
<td>A1.10</td>
<td>27.672926</td>
<td>85.428279</td>
<td>S83F</td>
<td>20-Dec-07</td>
<td>28</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB625</td>
<td>Carrier</td>
<td>A</td>
<td>A1.3</td>
<td>27.678325</td>
<td>85.316358</td>
<td>S83F</td>
<td>16-Jan-09</td>
<td>57</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB63</td>
<td>Carrier</td>
<td>A</td>
<td>A1.7</td>
<td>27.693811</td>
<td>85.218566</td>
<td>S83F</td>
<td>21-Dec-07</td>
<td>35</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB637</td>
<td>Carrier</td>
<td>C</td>
<td>C</td>
<td>27.693335</td>
<td>85.281691</td>
<td>S83F</td>
<td>23-Jan-09</td>
<td>24</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB640</td>
<td>Carrier</td>
<td>A</td>
<td>A2</td>
<td>27.720193</td>
<td>85.329767</td>
<td>S83F</td>
<td>30-Jan-09</td>
<td>18</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB672</td>
<td>Carrier</td>
<td>A</td>
<td>A2</td>
<td>27.621528</td>
<td>85.302213</td>
<td>N/A</td>
<td>25-Mar-09</td>
<td>31</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB726</td>
<td>Carrier</td>
<td>A</td>
<td>A2</td>
<td>27.67076</td>
<td>85.32031</td>
<td>S83Y</td>
<td>9-Apr-09</td>
<td>44</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB20</td>
<td>Carrier</td>
<td>A</td>
<td>A1</td>
<td>27.665874</td>
<td>85.328512</td>
<td>S83F</td>
<td>12-Jun-09</td>
<td>40</td>
<td>F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GB75</td>
<td>Carrier</td>
<td>A</td>
<td>A1.7</td>
<td>27.600485</td>
<td>85.13981</td>
<td>S83F</td>
<td>19-Jun-09</td>
<td>49</td>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPA010</td>
<td>Acute</td>
<td>A1.11</td>
<td>27.6779</td>
<td>85.31366</td>
<td>S83F</td>
<td>4-Oct-11</td>
<td>22</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>SPA014</td>
<td>Acute</td>
<td>A1.1</td>
<td>27.67681</td>
<td>85.31805</td>
<td>S83F</td>
<td>4-Nov-11</td>
<td>28</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>SPA015</td>
<td>Acute</td>
<td>A1.11</td>
<td>27.66745</td>
<td>85.30191</td>
<td>S83F</td>
<td>13-Apr-11</td>
<td>15</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>SPA017</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.67895</td>
<td>85.32656</td>
<td>S83F</td>
<td>15-Apr-11</td>
<td>19</td>
<td>M</td>
<td>Stone spout</td>
<td>Mix</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA018</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.6671</td>
<td>85.36555</td>
<td>S83F</td>
<td>20-Apr-11</td>
<td>9</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA019</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66545</td>
<td>85.32481</td>
<td>S83F</td>
<td>28-Apr-11</td>
<td>25</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA02</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66033</td>
<td>85.31735</td>
<td>S83F</td>
<td>21-Mar-11</td>
<td>17</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA020</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67618</td>
<td>85.34698</td>
<td>S83F</td>
<td>28-Apr-11</td>
<td>18</td>
<td>M</td>
<td>Bottled</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA021</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.65503</td>
<td>85.31999</td>
<td>S83F</td>
<td>5-May-11</td>
<td>24</td>
<td>F</td>
<td>Bottled</td>
<td>Mix</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA022</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.69058</td>
<td>85.31155</td>
<td>S83F</td>
<td>5-Mar-11</td>
<td>20</td>
<td>M</td>
<td>Bottled</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA023</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67995</td>
<td>85.3481</td>
<td>S83F</td>
<td>5-May-11</td>
<td>15</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA024</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.66573</td>
<td>85.3115</td>
<td>S83F</td>
<td>5-Aug-11</td>
<td>9</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA025</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.66845</td>
<td>85.34206</td>
<td>S83F</td>
<td>15-May-11</td>
<td>22</td>
<td>M</td>
<td>Bottled</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA026</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66611</td>
<td>85.3171</td>
<td>S83F</td>
<td>22-May-11</td>
<td>10</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA027</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67223</td>
<td>85.03216</td>
<td>S83F</td>
<td>24-May-11</td>
<td>20</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA028</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67246</td>
<td>85.322</td>
<td>S83F</td>
<td>26-May-11</td>
<td>23</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA029</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67298</td>
<td>85.31961</td>
<td>S83F</td>
<td>20-May-11</td>
<td>18</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA030</td>
<td>Acute</td>
<td>A</td>
<td>A1.2</td>
<td>27.66778</td>
<td>85.3242</td>
<td>S83F</td>
<td>22-May-11</td>
<td>20</td>
<td>M</td>
<td>Municipal</td>
<td>Boil</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA031</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.69586</td>
<td>85.31235</td>
<td>S83F</td>
<td>26-May-11</td>
<td>23</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA032</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67233</td>
<td>85.3216</td>
<td>S83F</td>
<td>24-May-11</td>
<td>20</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA033</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67246</td>
<td>85.322</td>
<td>S83F</td>
<td>6-Jun-11</td>
<td>16</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA034</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67758</td>
<td>85.32498</td>
<td>S83F</td>
<td>19-Jun-11</td>
<td>21</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA035</td>
<td>Acute</td>
<td>A</td>
<td>A1.11</td>
<td>27.67373</td>
<td>85.31835</td>
<td>S83F</td>
<td>20-Jun-11</td>
<td>25</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA036</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67381</td>
<td>85.32286</td>
<td>S83F</td>
<td>7-Oct-11</td>
<td>7</td>
<td>M</td>
<td>Municipal</td>
<td>Boil</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA037</td>
<td>Acute</td>
<td>A</td>
<td>A1.1</td>
<td>27.68108</td>
<td>85.32763</td>
<td>S83F</td>
<td>14-Jul-11</td>
<td>25</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA038</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67088</td>
<td>85.33695</td>
<td>S83F</td>
<td>18-Jul-11</td>
<td>8</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA039</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66313</td>
<td>85.30933</td>
<td>S83F</td>
<td>17-Jul-11</td>
<td>8</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA040</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67536</td>
<td>85.34578</td>
<td>S83F</td>
<td>22-Jul-11</td>
<td>55</td>
<td>M</td>
<td>Mix</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Code</td>
<td>Type</td>
<td>Case</td>
<td>StartMonth</td>
<td>StartDay</td>
<td>StartYear</td>
<td>EndMonth</td>
<td>EndDay</td>
<td>EndYear</td>
<td>Gender</td>
<td>Age</td>
<td>Source</td>
<td>Treatment</td>
<td>Filter</td>
<td>Boil</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>------</td>
<td>------------</td>
<td>----------</td>
<td>------------</td>
<td>----------</td>
<td>--------</td>
<td>---------</td>
<td>--------</td>
<td>-----</td>
<td>--------</td>
<td>-----------</td>
<td>--------</td>
<td>------</td>
<td>-----------</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>SPA058</td>
<td>Acute</td>
<td>A A1.1</td>
<td>27.6792</td>
<td>85.34925</td>
<td>2011</td>
<td>22-Jul</td>
<td>16</td>
<td>2011</td>
<td>F</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA06</td>
<td>Acute</td>
<td>A A1.2</td>
<td>27.68125</td>
<td>85.30543</td>
<td>2011</td>
<td>3-Mar</td>
<td>11</td>
<td>2011</td>
<td>F</td>
<td>Well</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA062</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65415</td>
<td>85.32045</td>
<td>2011</td>
<td>24-Jul</td>
<td>10</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Boil</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA063</td>
<td>Acute</td>
<td>A A2</td>
<td>27.66641</td>
<td>85.34001</td>
<td>2011</td>
<td>24-Jul</td>
<td>18</td>
<td>2011</td>
<td>M</td>
<td>Bottled</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA064</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65408</td>
<td>85.32073</td>
<td>2011</td>
<td>25-Jul</td>
<td>25</td>
<td>2011</td>
<td>M</td>
<td>Mix</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA07</td>
<td>Acute</td>
<td>A A2</td>
<td>27.68031</td>
<td>85.3365</td>
<td>2011</td>
<td>3-Mar</td>
<td>18</td>
<td>2011</td>
<td>M</td>
<td>Bottled</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA072</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65323</td>
<td>85.3214</td>
<td>2011</td>
<td>8-Feb</td>
<td>19</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA076</td>
<td>Acute</td>
<td>A A2</td>
<td>27.65968</td>
<td>85.31146</td>
<td>2011</td>
<td>8-Jul</td>
<td>6</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA077</td>
<td>Acute</td>
<td>A A1.5</td>
<td>27.66098</td>
<td>85.30935</td>
<td>2011</td>
<td>14-Aug</td>
<td>7</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA078</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65855</td>
<td>85.32275</td>
<td>2011</td>
<td>15-Aug</td>
<td>10</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA079</td>
<td>Acute</td>
<td>C C</td>
<td>27.64206</td>
<td>85.33405</td>
<td>2011</td>
<td>15-Aug</td>
<td>33</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>SPA082</td>
<td>Acute</td>
<td>A A2</td>
<td>27.66486</td>
<td>85.33193</td>
<td>2011</td>
<td>17-Aug</td>
<td>21</td>
<td>2011</td>
<td>M</td>
<td>Bottled</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA085</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65855</td>
<td>85.32275</td>
<td>2011</td>
<td>15-Aug</td>
<td>21</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA086</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65051</td>
<td>85.32133</td>
<td>2011</td>
<td>25-Aug</td>
<td>10</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA09</td>
<td>Acute</td>
<td>A A1.2</td>
<td>27.6666</td>
<td>85.3295</td>
<td>2011</td>
<td>4-Nov</td>
<td>9</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA096</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65415</td>
<td>85.31986</td>
<td>2011</td>
<td>9-Sep</td>
<td>12</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA098</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.65886</td>
<td>85.3134</td>
<td>2011</td>
<td>13-Sep</td>
<td>25</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA100</td>
<td>Acute</td>
<td>A A2</td>
<td>27.67816</td>
<td>85.32581</td>
<td>2011</td>
<td>15-Sep</td>
<td>25</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Untreated</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>SPA102</td>
<td>Acute</td>
<td>A A2</td>
<td>27.68808</td>
<td>85.31613</td>
<td>2011</td>
<td>23-Sep</td>
<td>11</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Mix</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>TY018</td>
<td>Acute</td>
<td>A A2</td>
<td>27.6762</td>
<td>85.32461</td>
<td>2011</td>
<td>6-Aug</td>
<td>18</td>
<td>2011</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>TY035</td>
<td>Acute</td>
<td>A A1.3</td>
<td>27.66776</td>
<td>85.3051</td>
<td>2011</td>
<td>4-Sep</td>
<td>17</td>
<td>2011</td>
<td>M</td>
<td>Stone spout</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>TY037</td>
<td>Acute</td>
<td>A A1.11</td>
<td>27.66083</td>
<td>85.32318</td>
<td>2011</td>
<td>7-Sep</td>
<td>5</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>TY051</td>
<td>Acute</td>
<td>A A1.11</td>
<td>27.67546</td>
<td>85.29706</td>
<td>2011</td>
<td>31-Oct</td>
<td>14</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY004</td>
<td>Acute</td>
<td>A A2</td>
<td>27.68808</td>
<td>85.31613</td>
<td>2011</td>
<td>23/9/2011</td>
<td>11</td>
<td>2011</td>
<td>F</td>
<td>Municipal</td>
<td>Boil</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>02TY013</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.66751</td>
<td>85.32836</td>
<td>2011</td>
<td>25/12/2011</td>
<td>22</td>
<td>2011</td>
<td>M</td>
<td>Stone spout</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>02TY036</td>
<td>Acute</td>
<td>A A2.1</td>
<td>27.66438</td>
<td>85.30686</td>
<td>2011</td>
<td>22/4/2012</td>
<td>30</td>
<td>2011</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY041</td>
<td>Acute</td>
<td>A A2</td>
<td>27.88046</td>
<td>85.33966</td>
<td>2012</td>
<td>29/4/2012</td>
<td>23</td>
<td>2012</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Code</td>
<td>Type</td>
<td>A</td>
<td>Z</td>
<td>Lattitude</td>
<td>Longitude</td>
<td>Provider</td>
<td>Filter</td>
<td>Boil</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>---</td>
<td>---</td>
<td>-----------</td>
<td>-----------</td>
<td>----------</td>
<td>--------</td>
<td>------</td>
<td>-----------</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>02TY044</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68915</td>
<td>85.32486</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>02TY046</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68053</td>
<td>85.33798</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY049</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.67411</td>
<td>85.32641</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY052</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.66096</td>
<td>85.31985</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY053</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66735</td>
<td>85.31911</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY062</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68858</td>
<td>85.32641</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY063</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68858</td>
<td>85.32641</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY067</td>
<td>Acute</td>
<td>A</td>
<td>A1.9</td>
<td>27.68263</td>
<td>85.3381</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY068</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.693</td>
<td>85.34703</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY072</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.67253</td>
<td>85.3352</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY074</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68193</td>
<td>85.32373</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY075</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67121</td>
<td>85.3166</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY076</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.66033</td>
<td>85.33945</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY077</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.6673</td>
<td>85.34081</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY081</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67988</td>
<td>85.33608</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY084</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68456</td>
<td>85.34125</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY085</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67305</td>
<td>85.33271</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY086</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.6698</td>
<td>85.3161</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>02TY091</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67656</td>
<td>85.32735</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY094</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68203</td>
<td>85.34038</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02TY100</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68236</td>
<td>85.31561</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02TY101</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.67696</td>
<td>85.32135</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY109</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66198</td>
<td>85.30661</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY129</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.65771</td>
<td>85.31621</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY152</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.69551</td>
<td>85.33811</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY153</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.69343</td>
<td>85.34256</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>02TY155</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.69343</td>
<td>85.34256</td>
<td>Sk3F</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Acute</td>
<td>C</td>
<td>C</td>
<td>27.66913</td>
<td>85.32658</td>
<td>S83F</td>
<td>14/08/2013</td>
<td>10</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>-----</td>
<td>-------</td>
<td>---</td>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td>------</td>
<td>-------------</td>
<td>----</td>
<td>---</td>
<td>-----------</td>
<td>--------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>-----</td>
<td>----</td>
<td>-----</td>
</tr>
<tr>
<td>02TY180</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.68955</td>
<td>85.33551</td>
<td>S83F</td>
<td>28/08/2013</td>
<td>25</td>
<td>M</td>
<td>Bottled</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY212</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.68063</td>
<td>85.31225</td>
<td>S83F</td>
<td>12/1/14</td>
<td>34</td>
<td>M</td>
<td>Municipal</td>
<td>Untreated</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY215</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.66568</td>
<td>85.32441</td>
<td>S83F</td>
<td>29/01/2014</td>
<td>32</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY211</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.67211</td>
<td>85.31981</td>
<td>S83F</td>
<td>31/03/2014</td>
<td>24</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY223</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.69096</td>
<td>85.33345</td>
<td>S83F</td>
<td>4/3/14</td>
<td>17</td>
<td>M</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY224</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.69075</td>
<td>85.3488</td>
<td>S83F</td>
<td>17/03/2014</td>
<td>27</td>
<td>M</td>
<td>Stone sprout</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY231</td>
<td>Acute</td>
<td>A</td>
<td>A2</td>
<td>27.66741</td>
<td>85.35096</td>
<td>S83F</td>
<td>1/5/14</td>
<td>18</td>
<td>M</td>
<td>Bottled</td>
<td>Untreated</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>02TY242</td>
<td>Acute</td>
<td>A</td>
<td>A2.1</td>
<td>27.68543</td>
<td>85.3415</td>
<td>S83F</td>
<td>16/06/2014</td>
<td>15</td>
<td>F</td>
<td>Municipal</td>
<td>Filter</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
</tbody>
</table>
# Appendix C

**Salmonella Typhi isolates and their corresponding metadata in chapter 5**

<table>
<thead>
<tr>
<th>ID</th>
<th>Age (years)</th>
<th>Sex</th>
<th>Year</th>
<th>Province</th>
<th>District</th>
<th>Commune</th>
<th>Longitude</th>
<th>Latitude</th>
<th>Lineages</th>
<th>Sublineages</th>
<th>IncH1 plasmid</th>
<th>AMP</th>
<th>CHL</th>
<th>SXT</th>
<th>Na</th>
<th>AZI</th>
<th>M1C</th>
<th>CIP</th>
<th>MIC</th>
<th>Acc No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.3</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Sala Kamraeuk</td>
<td>103.867133</td>
<td>13.345557</td>
<td>I</td>
<td>I</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>I</td>
<td>8</td>
<td>0.094</td>
<td>ERR340759</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>13.6</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Nokor Pheas</td>
<td>103.719479</td>
<td>13.624206</td>
<td>II</td>
<td>II</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>S</td>
<td>12</td>
<td>0.016</td>
<td>ERR360845</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14.8</td>
<td>M</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Ta Saom</td>
<td>103.677526</td>
<td>13.671199</td>
<td>II</td>
<td>II</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>S</td>
<td>8</td>
<td>0.012</td>
<td>ERR340758</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>12.0</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Sala Kamraeuk</td>
<td>103.867133</td>
<td>13.345557</td>
<td>II</td>
<td>II</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>16</td>
<td>0.19</td>
<td>ERR340774</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7.1</td>
<td>F</td>
<td>2007</td>
<td>Banteay Meanchey</td>
<td>O Chrauv</td>
<td>Poipet</td>
<td>102.659645</td>
<td>13.642996</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.19</td>
<td>ERR340760</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1.5</td>
<td>F</td>
<td>2011</td>
<td>Battambang</td>
<td>Samlaut</td>
<td>Chamlong</td>
<td>102.859385</td>
<td>12.583148</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360770</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>14.7</td>
<td>F</td>
<td>2012</td>
<td>Kampong Chhnang</td>
<td>Cheung Prey</td>
<td>Khnaior Dambang</td>
<td>105.014179</td>
<td>12.018816</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319473</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>3.8</td>
<td>F</td>
<td>2011</td>
<td>Kampong Chhnang</td>
<td>Kampong Chhnang</td>
<td>Khsam</td>
<td>104.660713</td>
<td>12.279773</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR319445</td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>12.7</td>
<td>M</td>
<td>2011</td>
<td>Kampong Thom</td>
<td>Stoeng Sen</td>
<td>Ou Kanthor</td>
<td>104.816132</td>
<td>12.667063</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319445</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>13.7</td>
<td>M</td>
<td>2010</td>
<td>Kampong Thom</td>
<td>Stoeng Sen</td>
<td>Srayov</td>
<td>104.858972</td>
<td>12.59656</td>
<td>III</td>
<td>IIIa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR360952</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>4.2</td>
<td>F</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>103.843962</td>
<td>13.273706</td>
<td>III</td>
<td>IIIa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR319416</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>13.2</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>III</td>
<td>IIIa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>8</td>
<td>0.008</td>
<td>ERR319459</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2010</td>
<td>Banteay Meanchey</td>
<td>Preah Neale</td>
<td>Prasat</td>
<td>103.365501</td>
<td>13.457949</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360754</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>5.9</td>
<td>F</td>
<td>2010</td>
<td>Banteay Meanchey</td>
<td>Preah Neale</td>
<td>Preah Neale</td>
<td>103.237876</td>
<td>13.593116</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319425</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>9.8</td>
<td>F</td>
<td>2010</td>
<td>Banteay Meanchey</td>
<td>Preah Neale</td>
<td>Preah Neale</td>
<td>103.369654</td>
<td>13.620483</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR360956</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>11.7</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Kouk Doung</td>
<td>103.792493</td>
<td>13.701174</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.25</td>
<td>ERR319439</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>8.1</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Tbaeng</td>
<td>104.038638</td>
<td>13.552495</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR360846</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>14.1</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Tbaeng</td>
<td>104.038638</td>
<td>13.552495</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR319434</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>4.6</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Krouch Kor</td>
<td>103.587557</td>
<td>13.703525</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.25</td>
<td>ERR319437</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>13.3</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Krouch Kor</td>
<td>103.587557</td>
<td>13.703525</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR319441</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>12.2</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Bakong</td>
<td>103.991452</td>
<td>13.355908</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR319429</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Sex</td>
<td>Year</td>
<td>Province</td>
<td>District</td>
<td>Village</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Province</td>
<td>District</td>
<td>Village</td>
<td>Village Number</td>
<td>Classification</td>
<td>District Code</td>
<td>Type</td>
<td>Area (ha)</td>
<td>Remarks</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>-----------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>----------------</td>
<td>---------------</td>
<td>--------------</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>III</td>
<td>IIIb</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>III</td>
<td>IIIb</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>0</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Nokor</td>
<td>103.889397</td>
<td>13.423837</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Siem Reab</td>
<td>103.843962</td>
<td>13.277306</td>
<td>III</td>
<td>IIIb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Siem Reab</td>
<td>103.843962</td>
<td>13.277306</td>
<td>III</td>
<td>IIIb</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Svay Leu</td>
<td>104.280929</td>
<td>13.45211</td>
<td>III</td>
<td>IIIb</td>
<td>0</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>M</td>
<td>2011</td>
<td>Banteay Meanchey</td>
<td>Mongkol Borei</td>
<td>Koy Maeng</td>
<td>103.10134</td>
<td>13.550598</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>2011</td>
<td>Banteay Meanchey</td>
<td>Preah Netro Preah</td>
<td>Chob Veari</td>
<td>103.167669</td>
<td>13.66311</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>2011</td>
<td>Banteay Meanchey</td>
<td>Preah Netro Preah</td>
<td>Phnom Lieb</td>
<td>103.297435</td>
<td>13.520053</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>2011</td>
<td>Banteay Meanchey</td>
<td>Preah Netro Preah</td>
<td>Phnom Lieb</td>
<td>103.297435</td>
<td>13.520053</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>F</td>
<td>2011</td>
<td>Battambang</td>
<td>Phnom Preak</td>
<td>Pech Chenda</td>
<td>102.538466</td>
<td>13.268541</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Siem Reab</td>
<td>103.623954</td>
<td>13.663276</td>
<td>III</td>
<td>IIIc</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>6</td>
<td>0.006</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Chi Krench</td>
<td>Anlong Samnar</td>
<td>104.229051</td>
<td>12.94545</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Chomloas Dai</td>
<td>103.46967</td>
<td>13.641287</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Kralanh</td>
<td>103.432035</td>
<td>13.55668</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Doun Kaev</td>
<td>103.774015</td>
<td>13.478241</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Kaev Pooar</td>
<td>103.716096</td>
<td>13.321008</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Prey Chruk</td>
<td>103.542856</td>
<td>13.503103</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Trei Nhoar</td>
<td>103.706054</td>
<td>13.574489</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Tuet Vil</td>
<td>103.807153</td>
<td>13.413196</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Yeang</td>
<td>103.624304</td>
<td>13.563253</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Chom Klnies</td>
<td>103.829966</td>
<td>13.175937</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>6</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45</td>
<td>F</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>III</td>
<td>IIIc</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>12</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Sotr Nikorn</td>
<td>Popel</td>
<td>104.163519</td>
<td>13.39285</td>
<td>III</td>
<td>IIIc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>F</td>
<td>2008</td>
<td>Banteay Meanchey</td>
<td>O Chrauv</td>
<td>Nimitt</td>
<td>102.659645</td>
<td>13.642996</td>
<td>IV</td>
<td>IVa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>16</td>
<td>0.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Gender</td>
<td>Year</td>
<td>Province</td>
<td>District</td>
<td>Sub-District</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Risk Level</td>
<td>Risk Type</td>
<td>Exposure</td>
<td>Value</td>
<td>Error Code</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>--------</td>
<td>------</td>
<td>-----------</td>
<td>----------</td>
<td>--------------</td>
<td>----------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------</td>
<td>-----------</td>
<td>-------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>F</td>
<td>2008</td>
<td>Oddar Meanchey</td>
<td>Anlong Veng</td>
<td>Anlong Veng</td>
<td>104.153127</td>
<td>14.240485</td>
<td>IV</td>
<td>IVa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR319410</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>F</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>IV</td>
<td>IVa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>12</td>
<td>0.25</td>
<td>ERR340767</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>F</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>IV</td>
<td>IVa</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>8</td>
<td>0.008</td>
<td>ERR319411</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV</td>
<td>IVa</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.19</td>
<td>ERR360890</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>52</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV</td>
<td>IVa</td>
<td>0</td>
<td>S</td>
<td>R</td>
<td>S</td>
<td>R</td>
<td>16</td>
<td>0.5</td>
<td>ERR319406</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53</td>
<td>M</td>
<td>2010</td>
<td>Banteay Meanchey</td>
<td>O Chrouv</td>
<td>Changha</td>
<td>102.892727</td>
<td>13.647435</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360955</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>M</td>
<td>2008</td>
<td>Banteay Meanchey</td>
<td>Preah Netr Preah</td>
<td>Prasat</td>
<td>103.365501</td>
<td>13.457949</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360742</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55</td>
<td>M</td>
<td>2008</td>
<td>Banteay Meanchey</td>
<td>Preah Netr Preah</td>
<td>Prasat</td>
<td>103.365501</td>
<td>13.457949</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360889</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56</td>
<td>F</td>
<td>2008</td>
<td>Battambang</td>
<td>ek Phnum</td>
<td>Kaob Chivaeng</td>
<td>103.610762</td>
<td>13.21135</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.25</td>
<td>ERR360743</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57</td>
<td>F</td>
<td>2009</td>
<td>Kampong Cham</td>
<td>Chamkar Leu</td>
<td>Bos Khmaor</td>
<td>105.335766</td>
<td>12.170203</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.5</td>
<td>ERR319415</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58</td>
<td>M</td>
<td>2007</td>
<td>Kampong Thom</td>
<td>Stoung</td>
<td>Banteay Stoung</td>
<td>104.6337</td>
<td>12.943935</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR340757</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59</td>
<td>M</td>
<td>2008</td>
<td>Kampong Thom</td>
<td>Stoeng Sen</td>
<td>Achar Leak</td>
<td>104.903907</td>
<td>12.727871</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR360950</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>F</td>
<td>2007</td>
<td>Pursat</td>
<td>Bakan</td>
<td>Me Tuck</td>
<td>103.782131</td>
<td>12.803292</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR340770</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>M</td>
<td>2012</td>
<td>Pursat</td>
<td>Bakan</td>
<td>Me Tuck</td>
<td>103.782131</td>
<td>12.803292</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360999</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Srae Khvav</td>
<td>103.672335</td>
<td>13.793712</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR360739</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>63</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Chi Krieng</td>
<td>Anlong Sannar</td>
<td>104.229051</td>
<td>12.94545</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340766</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Chi Krieng</td>
<td>Anlong Sannar</td>
<td>104.229051</td>
<td>12.94545</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340761</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65</td>
<td>M</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Chi Krieng</td>
<td>Anlong Sannar</td>
<td>104.229051</td>
<td>12.94545</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR319419</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Chi Krieng</td>
<td>Chi Kraeng</td>
<td>104.254404</td>
<td>13.074933</td>
<td>IV</td>
<td>IV</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>0</td>
<td>12</td>
<td>0.5</td>
<td>ERR319431</td>
<td></td>
<td></td>
</tr>
<tr>
<td>67</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.445813</td>
<td>13.433224</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340772</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.445813</td>
<td>13.433224</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360738</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.445813</td>
<td>13.433224</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360740</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.445813</td>
<td>13.433224</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319405</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71</td>
<td>F</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.445813</td>
<td>13.433224</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.5</td>
<td>ERR319409</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>M</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Chong Khnies</td>
<td>103.829966</td>
<td>13.175937</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR360744</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73</td>
<td>M</td>
<td>2008</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>Sambuor</td>
<td>103.7935</td>
<td>13.30069</td>
<td>IV</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR319403</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>#</td>
<td>Name</td>
<td>Age</td>
<td>Year</td>
<td>Location</td>
<td>Sex</td>
<td>District</td>
<td>Province</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Status</td>
<td>Marital Status</td>
<td>Gender</td>
<td>Members</td>
<td>Children</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td>--------------</td>
<td>-----</td>
<td>----------</td>
<td>----------</td>
<td>-----------</td>
<td>-----------</td>
<td>--------</td>
<td>---------------</td>
<td>--------</td>
<td>---------</td>
<td>----------</td>
<td>-------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74</td>
<td>F</td>
<td>15.2</td>
<td>2010</td>
<td>Siem Reap</td>
<td>F</td>
<td>Siem Reap</td>
<td>Siem Reap</td>
<td>103.77659</td>
<td>13.303244</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.38</td>
<td>ERR319428</td>
<td></td>
<td></td>
</tr>
<tr>
<td>75</td>
<td>M</td>
<td>12.0</td>
<td>2007</td>
<td>Siem Reap</td>
<td>M</td>
<td>Siem Reap</td>
<td>Sotr Nikom</td>
<td>104.08779</td>
<td>13.182907</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340768</td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>F</td>
<td>15.1</td>
<td>2008</td>
<td>Siem Reap</td>
<td>F</td>
<td>Sotr Nikom</td>
<td>Dan Run</td>
<td>104.08779</td>
<td>13.182907</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR360948</td>
<td></td>
<td></td>
</tr>
<tr>
<td>77</td>
<td>F</td>
<td>2.4</td>
<td>2007</td>
<td>Siem Reap</td>
<td>F</td>
<td>Sotr Nikom</td>
<td>Kampong Kheang</td>
<td>104.14206</td>
<td>13.046753</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR340771</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>F</td>
<td>4.5</td>
<td>2007</td>
<td>Siem Reap</td>
<td>F</td>
<td>Sotr Nikom</td>
<td>Kampong Kheang</td>
<td>104.14206</td>
<td>13.046753</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340773</td>
<td></td>
<td></td>
</tr>
<tr>
<td>79</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2007</td>
<td>Siem Reap</td>
<td>#N/A</td>
<td>Sotr Nikom</td>
<td>Kampong Kheang</td>
<td>104.14206</td>
<td>13.046753</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR340775</td>
<td></td>
<td></td>
</tr>
<tr>
<td>80</td>
<td>M</td>
<td>0.6</td>
<td>2007</td>
<td>Siem Reap</td>
<td>M</td>
<td>Sotr Nikom</td>
<td>Kampong Kheang</td>
<td>104.14206</td>
<td>13.046753</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR340763</td>
<td></td>
<td></td>
</tr>
<tr>
<td>81</td>
<td>M</td>
<td>3.2</td>
<td>2008</td>
<td>Siem Reap</td>
<td>M</td>
<td>Sotr Nikom</td>
<td>Kampong Kheang</td>
<td>104.14206</td>
<td>13.046753</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR360891</td>
<td></td>
<td></td>
</tr>
<tr>
<td>82</td>
<td>F</td>
<td>6.1</td>
<td>2008</td>
<td>Banteay Meas</td>
<td>F</td>
<td>Preah Netr</td>
<td>Preah Prasat</td>
<td>103.36501</td>
<td>13.457949</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319404</td>
<td></td>
<td></td>
</tr>
<tr>
<td>83</td>
<td>F</td>
<td>5.2</td>
<td>2010</td>
<td>Battambang</td>
<td>F</td>
<td>Preah Netr</td>
<td>Preah Prasat</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ERR319433</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>M</td>
<td>4.1</td>
<td>2009</td>
<td>Siem Reap</td>
<td>M</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.44581</td>
<td>13.433224</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319412</td>
<td></td>
<td></td>
</tr>
<tr>
<td>85</td>
<td>F</td>
<td>6.6</td>
<td>2009</td>
<td>Siem Reap</td>
<td>F</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.44581</td>
<td>13.433224</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319413</td>
<td></td>
<td></td>
</tr>
<tr>
<td>86</td>
<td>M</td>
<td>14.7</td>
<td>2010</td>
<td>Battambang</td>
<td>M</td>
<td>Preah Netr</td>
<td>Tuek Chour</td>
<td>103.36965</td>
<td>13.620483</td>
<td>IVb</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.5</td>
<td>ERR319426</td>
<td></td>
<td></td>
</tr>
<tr>
<td>87</td>
<td>F</td>
<td>3.3</td>
<td>2010</td>
<td>Battambang</td>
<td>F</td>
<td>Preah Netr</td>
<td>Phnum Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.38</td>
<td>ERR360748</td>
<td></td>
<td></td>
</tr>
<tr>
<td>88</td>
<td>F</td>
<td>2.6</td>
<td>2011</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360761</td>
<td></td>
<td></td>
</tr>
<tr>
<td>89</td>
<td>F</td>
<td>5.6</td>
<td>2011</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360772</td>
<td></td>
<td></td>
</tr>
<tr>
<td>90</td>
<td>F</td>
<td>6.8</td>
<td>2008</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR360949</td>
<td></td>
<td></td>
</tr>
<tr>
<td>91</td>
<td>F</td>
<td>1.9</td>
<td>2009</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>16</td>
<td>0.38</td>
<td>ERR319420</td>
<td></td>
<td></td>
</tr>
<tr>
<td>92</td>
<td>M</td>
<td>1.2</td>
<td>2010</td>
<td>Battambang</td>
<td>M</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.38</td>
<td>ERR319436</td>
<td></td>
<td></td>
</tr>
<tr>
<td>93</td>
<td>M</td>
<td>4.1</td>
<td>2011</td>
<td>Battambang</td>
<td>M</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319469</td>
<td></td>
<td></td>
</tr>
<tr>
<td>94</td>
<td>M</td>
<td>5.7</td>
<td>2008</td>
<td>Battambang</td>
<td>M</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.5</td>
<td>ERR319408</td>
<td></td>
<td></td>
</tr>
<tr>
<td>95</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2012</td>
<td>Battambang</td>
<td>#N/A</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ERR319475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>96</td>
<td>F</td>
<td>11.6</td>
<td>2012</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Kaoh Chiveang</td>
<td>103.61076</td>
<td>13.21135</td>
<td>IV</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ERR319476</td>
<td></td>
<td></td>
</tr>
<tr>
<td>97</td>
<td>F</td>
<td>5.9</td>
<td>2011</td>
<td>Battambang</td>
<td>F</td>
<td>Phnum</td>
<td>Prey Chas</td>
<td>103.39292</td>
<td>13.304805</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360771</td>
<td></td>
<td></td>
</tr>
<tr>
<td>98</td>
<td>M</td>
<td>12.1</td>
<td>2012</td>
<td>Preah Vihear</td>
<td>M</td>
<td>Kulen</td>
<td>Srayang</td>
<td>104.51133</td>
<td>13.76033</td>
<td>IV</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319475</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>Type</td>
<td>Year</td>
<td>Province</td>
<td>Location</td>
<td>Latitude</td>
<td>Longitude</td>
<td>Stability Level</td>
<td>Construction Level</td>
<td>Water Level</td>
<td>Threat Level</td>
<td>Value</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>------</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>----------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Nokor Pheas</td>
<td>103.719479</td>
<td>13.624206</td>
<td>IV</td>
<td>IVb</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360964</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>101</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Chi Keng</td>
<td>Chi Kraeng</td>
<td>104.244404</td>
<td>13.074933</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR369093</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>102</td>
<td>M</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Sambuor</td>
<td>103.44813</td>
<td>13.433224</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR319417</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>103</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kampong Phluk</td>
<td>103.986709</td>
<td>13.235758</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360769</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kampong Phluk</td>
<td>103.986709</td>
<td>13.235758</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319467</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>105</td>
<td>M</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Kaeo Poar</td>
<td>103.716096</td>
<td>13.321008</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.38</td>
<td>ERR319421</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>106</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Kaeo Poar</td>
<td>103.716096</td>
<td>13.321008</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319440</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>107</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Kaeo Poar</td>
<td>103.659697</td>
<td>13.415687</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR319424</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>108</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Pou Treay</td>
<td>103.53096</td>
<td>13.374667</td>
<td>IV</td>
<td>IVb</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>M</td>
<td>2011</td>
<td>Banteay Meanchey</td>
<td>Serei Saophonj</td>
<td>Kamlak</td>
<td>102.965143</td>
<td>13.674269</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360760</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>110</td>
<td>M</td>
<td>2012</td>
<td>Banteay Meanchey</td>
<td>Serei Saophonj</td>
<td>Preah Ponlea</td>
<td>102.900584</td>
<td>13.579188</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360787</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>111</td>
<td>M</td>
<td>2012</td>
<td>Kampong Thom</td>
<td>Kampong Svay</td>
<td>Kampong Svay</td>
<td>104.980724</td>
<td>12.765143</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319472</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>112</td>
<td>M</td>
<td>2002</td>
<td>Oddar Meanchey</td>
<td>Chong Kal</td>
<td>Pongro</td>
<td>103.611356</td>
<td>14.057113</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360972</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>113</td>
<td>M</td>
<td>2012</td>
<td>Oddar Meanchey</td>
<td>Samraong</td>
<td>Osmach</td>
<td>103.662372</td>
<td>14.270904</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360972</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>114</td>
<td>F</td>
<td>2012</td>
<td>Oddar Meanchey</td>
<td>Samraong</td>
<td>Osmach</td>
<td>103.662372</td>
<td>14.270904</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360972</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>115</td>
<td>M</td>
<td>2012</td>
<td>Oddar Meanchey</td>
<td>Samraong</td>
<td>Osmach</td>
<td>103.662372</td>
<td>14.270904</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319495</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>116</td>
<td>F</td>
<td>2011</td>
<td>Oddar Meanchey</td>
<td>Samraong</td>
<td>Samraong</td>
<td>103.543846</td>
<td>14.154499</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360765</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>117</td>
<td>M</td>
<td>2011</td>
<td>Preah Vihear</td>
<td>Choam Ksan</td>
<td>Kuantou</td>
<td>104.6417</td>
<td>14.280355</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360756</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>118</td>
<td>F</td>
<td>2012</td>
<td>Preah Vihear</td>
<td>Kampong</td>
<td>Snyay</td>
<td>104.511339</td>
<td>13.76033</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360986</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>119</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Char Chhuk</td>
<td>103.623954</td>
<td>13.663726</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360779</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>120</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Char Chhuk</td>
<td>103.623954</td>
<td>13.663726</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360981</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>121</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Serei Saophonj</td>
<td>103.623954</td>
<td>13.663726</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319465</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>122</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Kambon</td>
<td>103.792493</td>
<td>13.701174</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319464</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>123</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Nokor Pheas</td>
<td>103.719479</td>
<td>13.624206</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360762</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>124</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Ta Saom</td>
<td>103.677526</td>
<td>13.671199</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360973</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>125</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Ta Saom</td>
<td>103.677526</td>
<td>13.671199</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360977</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>M/F</td>
<td>Year</td>
<td>Province</td>
<td>District</td>
<td>Site Name</td>
<td>Latitude, Longitude</td>
<td>Elev.</td>
<td>Elevation</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Error</td>
<td>Location</td>
<td>Latitude, Longitude</td>
<td>Elev.</td>
<td>Elevation</td>
<td>Std. Dev.</td>
<td>Mean</td>
<td>Error</td>
<td>Location</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>------------</td>
<td>---------------------</td>
<td>-------</td>
<td>------------</td>
<td>-----------</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
<td>---------------------</td>
<td>-------</td>
<td>------------</td>
<td>-----------</td>
<td>------</td>
<td>-------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>126</td>
<td>6.7</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Ta Saom</td>
<td>103.677526</td>
<td>13.671199</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR319486</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>127</td>
<td>12.3</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Ta Saom</td>
<td>103.677526</td>
<td>13.671199</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR319487</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>128</td>
<td>12.3</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor</td>
<td>Leang Dai</td>
<td>103.876266</td>
<td>13.502826</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR360775</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>129</td>
<td>12.9</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Leang Dai</td>
<td>103.876266</td>
<td>13.502826</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360979</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>1.8</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Angkor</td>
<td>Peak Snaeng</td>
<td>103.885243</td>
<td>13.610438</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>8</td>
<td>0.25</td>
<td>ERR319435</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>131</td>
<td>4.7</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Khun Sanday</td>
<td>103.952295</td>
<td>13.569677</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360966</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>132</td>
<td>3.9</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Khun Ream</td>
<td>104.032137</td>
<td>13.668959</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR360778</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>133</td>
<td>9.3</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Khun Ream</td>
<td>104.032137</td>
<td>13.668959</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319494</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>134</td>
<td>11.3</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Runcrok</td>
<td>103.981952</td>
<td>13.508645</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360783</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>135</td>
<td>2.0</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Run Ta Aek</td>
<td>104.039073</td>
<td>13.465833</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td>ERR319444</td>
<td></td>
<td></td>
</tr>
<tr>
<td>136</td>
<td>13.1</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Run Ta Aek</td>
<td>104.039073</td>
<td>13.465833</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319479</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>137</td>
<td>13.0</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Banteay Srei</td>
<td>Tbaeng</td>
<td>104.038638</td>
<td>13.552405</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360992</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>138</td>
<td>3.8</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Chi Krong</td>
<td>Anlong Sannar</td>
<td>104.229051</td>
<td>12.94554</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR360959</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>139</td>
<td>15.4</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Chi Krong</td>
<td>Ruessi Lok</td>
<td>104.227712</td>
<td>13.251327</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319482</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>140</td>
<td>14.2</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Kralanh</td>
<td>Chomloas Dai</td>
<td>103.46967</td>
<td>13.641287</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360965</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>141</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kandaek</td>
<td>103.931669</td>
<td>13.299934</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360745</td>
<td></td>
<td></td>
</tr>
<tr>
<td>142</td>
<td>7.8</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kandaek</td>
<td>103.931669</td>
<td>13.299934</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360781</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>143</td>
<td>6.6</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kandaek</td>
<td>103.931669</td>
<td>13.299934</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ERR360968</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>144</td>
<td>7.0</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kandaek</td>
<td>103.931669</td>
<td>13.299934</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.19</td>
<td>ERR319485</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>145</td>
<td>13.4</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Kantreang</td>
<td>104.03968</td>
<td>13.347416</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR360990</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>146</td>
<td>9.0</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Prasat Bakong</td>
<td>Rohuos</td>
<td>103.98127</td>
<td>13.310897</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR360759</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>147</td>
<td>5.8</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Reul</td>
<td>103.73586</td>
<td>13.536268</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>ERR319453</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>148</td>
<td>3.0</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Samraong Yea</td>
<td>103.69054</td>
<td>13.361494</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR360984</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>149</td>
<td>11.8</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Sasar Sdam</td>
<td>103.630882</td>
<td>13.509808</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>12</td>
<td>0.38</td>
<td>ERR319414</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>10.9</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Trei Nhaor</td>
<td>103.706054</td>
<td>13.574489</td>
<td>IV</td>
<td>IVc</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR331208</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

214
<table>
<thead>
<tr>
<th>No.</th>
<th>Age</th>
<th>Gender</th>
<th>Year</th>
<th>Province</th>
<th>District</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Region</th>
<th>Village</th>
<th>Gender/Name</th>
<th>Age</th>
<th>Sex</th>
<th>Hair</th>
<th>Height</th>
<th>Weight</th>
<th>BMI</th>
<th>Status</th>
<th>Code</th>
</tr>
</thead>
<tbody>
<tr>
<td>152</td>
<td>10.9</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok Nheak</td>
<td>103.706054</td>
<td>13.574489</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360788</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>153</td>
<td>5.9</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok Nheak</td>
<td>103.706054</td>
<td>13.574489</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360776</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>154</td>
<td>11.5</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Puok Nheak</td>
<td>103.706054</td>
<td>13.574489</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>155</td>
<td>9.3</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Chrey</td>
<td>103.888751</td>
<td>13.283577</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360777</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>156</td>
<td>8.2</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360763</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>157</td>
<td>1.6</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360971</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>158</td>
<td>6.7</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR319483</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>159</td>
<td>9.0</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Nokor Thum</td>
<td>103.889937</td>
<td>13.423837</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360976</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>7.2</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Sala Kamrauk</td>
<td>103.867133</td>
<td>13.345557</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360963</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>161</td>
<td>9.8</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Sambuor</td>
<td>103.7935</td>
<td>13.30069</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360961</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>162</td>
<td>10.4</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Sambuor</td>
<td>103.7935</td>
<td>13.30069</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR319461</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>163</td>
<td>11.0</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>103.843962</td>
<td>13.277306</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360991</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>164</td>
<td>15.7</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360789</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>165</td>
<td>3.7</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>12 0.25</td>
<td>ERR360847</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>166</td>
<td>1.7</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>12 0.38</td>
<td>ERR360848</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>167</td>
<td>6.3</td>
<td>M</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>12 0.25</td>
<td>ERR360960</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>168</td>
<td>11.4</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360970</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>169</td>
<td>6.4</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360785</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>170</td>
<td>2.9</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360780</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>171</td>
<td>14.1</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR331205</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>172</td>
<td>6.2</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360983</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>173</td>
<td>14.1</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Svay Dangkum</td>
<td>103.822374</td>
<td>13.311014</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR319489</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>174</td>
<td>13.2</td>
<td>F</td>
<td>2012</td>
<td>Soth Nikom</td>
<td>Dan Daek</td>
<td>104.120345</td>
<td>13.255388</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360967</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>175</td>
<td>11.6</td>
<td>F</td>
<td>2012</td>
<td>Soth Nikom</td>
<td>Dan Daek</td>
<td>104.120345</td>
<td>13.255388</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360975</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>176</td>
<td>15.1</td>
<td>F</td>
<td>2012</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.087793</td>
<td>13.182907</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR331206</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>177</td>
<td>3.8</td>
<td>F</td>
<td>2012</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.087793</td>
<td>13.182907</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>0 0 0 0</td>
<td>ERR360786</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>178</td>
<td>6.4</td>
<td>M</td>
<td>2010</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.087793</td>
<td>13.182907</td>
<td>IV IVc</td>
<td>R R R R</td>
<td>12 0.25</td>
<td>ERR360954</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Age</td>
<td>Sex</td>
<td>Year</td>
<td>District</td>
<td>Province</td>
<td>Village</td>
<td>Longitude</td>
<td>Latitude</td>
<td>Date</td>
<td>Longitude_1</td>
<td>Latitude_1</td>
<td>Tag</td>
<td>ID</td>
<td>Reference</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>----------</td>
<td>----------</td>
<td>---------</td>
<td>-----------</td>
<td>----------</td>
<td>------</td>
<td>-------------</td>
<td>-------------</td>
<td>-----</td>
<td>-----</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>179</td>
<td>10.0</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.087793</td>
<td>13.182907</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319456</td>
<td></td>
</tr>
<tr>
<td>180</td>
<td>4.0</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Sotr Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360755</td>
<td></td>
</tr>
<tr>
<td>181</td>
<td>7.6</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360774</td>
<td></td>
</tr>
<tr>
<td>182</td>
<td>8.6</td>
<td>F</td>
<td>2010</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360958</td>
<td></td>
</tr>
<tr>
<td>183</td>
<td>2.2</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360961</td>
<td></td>
</tr>
<tr>
<td>184</td>
<td>3.0</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360980</td>
<td></td>
</tr>
<tr>
<td>185</td>
<td>#N/A</td>
<td>#N/A</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>ERR360985</td>
<td></td>
</tr>
<tr>
<td>186</td>
<td>4.8</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360994</td>
<td></td>
</tr>
<tr>
<td>187</td>
<td>9.3</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319442</td>
<td></td>
</tr>
<tr>
<td>188</td>
<td>5.1</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td>ERR319443</td>
<td></td>
</tr>
<tr>
<td>189</td>
<td>2.0</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319449</td>
<td></td>
</tr>
<tr>
<td>190</td>
<td>5.7</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td>ERR319450</td>
<td></td>
</tr>
<tr>
<td>191</td>
<td>7.0</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>12</td>
<td>0.38</td>
<td>ERR319451</td>
<td></td>
</tr>
<tr>
<td>192</td>
<td>7.4</td>
<td>M</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319468</td>
<td></td>
</tr>
<tr>
<td>193</td>
<td>11.8</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0.38</td>
<td>ERR319474</td>
<td></td>
</tr>
<tr>
<td>194</td>
<td>7.7</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319477</td>
<td></td>
</tr>
<tr>
<td>195</td>
<td>3.8</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319480</td>
<td></td>
</tr>
<tr>
<td>196</td>
<td>10.4</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319481</td>
<td></td>
</tr>
<tr>
<td>197</td>
<td>12.9</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319490</td>
<td></td>
</tr>
<tr>
<td>198</td>
<td>6.5</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Kcleang</td>
<td>104.142061</td>
<td>13.046753</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319492</td>
<td></td>
</tr>
<tr>
<td>199</td>
<td>9.5</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Khchas</td>
<td>104.06283</td>
<td>13.254239</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR360996</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>4.8</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Sotr Nikom</td>
<td>Khnar Pou</td>
<td>104.089546</td>
<td>13.423928</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0.38</td>
<td>ERR331207</td>
<td></td>
</tr>
<tr>
<td>201</td>
<td>7.0</td>
<td>M</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Khnar Pou</td>
<td>104.089546</td>
<td>13.423928</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319488</td>
<td></td>
</tr>
<tr>
<td>202</td>
<td>3.1</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kien Sangkae</td>
<td>104.160906</td>
<td>13.208966</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0.25</td>
<td>ERR360978</td>
<td></td>
</tr>
<tr>
<td>203</td>
<td>10.8</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Varin</td>
<td>Lvea Krang</td>
<td>103.912669</td>
<td>13.900364</td>
<td>IV</td>
<td>IVc</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>ERR319478</td>
<td></td>
</tr>
<tr>
<td>204</td>
<td>8.7</td>
<td>F</td>
<td>2010</td>
<td>Banteaynea</td>
<td>nchey</td>
<td>O Chrauv</td>
<td>102.659645</td>
<td>13.642906</td>
<td>non_HS8</td>
<td>non_HS8</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>6</td>
<td>0.006</td>
<td>ERR360747</td>
</tr>
<tr>
<td>No.</td>
<td>11.1</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Char Chhuk</td>
<td>103.623954</td>
<td>13.663726</td>
<td>non_H5</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR340765</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>-----</td>
<td>------</td>
<td>-----------</td>
<td>-------------</td>
<td>-----------</td>
<td>------------</td>
<td>-----------</td>
<td>--------</td>
<td>----</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>206</td>
<td>7.7</td>
<td>F</td>
<td>2012</td>
<td>Siem Reap</td>
<td>Angkor Chum</td>
<td>Doun Peang</td>
<td>103.705808</td>
<td>13.728858</td>
<td>non_H5</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>0</td>
<td>0.006</td>
<td>ERR360095</td>
<td></td>
</tr>
<tr>
<td>207</td>
<td>7.6</td>
<td>F</td>
<td>2011</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Sasar Sdam</td>
<td>103.630882</td>
<td>13.509808</td>
<td>non_H5</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>0</td>
<td>0</td>
<td>ERR319462</td>
<td></td>
</tr>
<tr>
<td>208</td>
<td>3.8</td>
<td>M</td>
<td>2009</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Kouk Chak</td>
<td>103.841705</td>
<td>13.429301</td>
<td>non_H5</td>
<td>0</td>
<td>S</td>
<td>S</td>
<td>S</td>
<td>R</td>
<td>0</td>
<td>0.25</td>
<td>ERR319418</td>
<td></td>
</tr>
<tr>
<td>209</td>
<td>8.4</td>
<td>F</td>
<td>2007</td>
<td>Siem Reap</td>
<td>Siem Reab</td>
<td>Sla Kram</td>
<td>103.885703</td>
<td>13.367605</td>
<td>non_H5</td>
<td>1</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>R</td>
<td>0</td>
<td>0.38</td>
<td>ERR340769</td>
<td></td>
</tr>
<tr>
<td>210</td>
<td>12.3</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Ampil</td>
<td>103.946822</td>
<td>13.39642561</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>211</td>
<td>9.4</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi KrÖng</td>
<td>Anlong Samnar</td>
<td>104.2274314</td>
<td>12.94743614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>212</td>
<td>5.2</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi KrÖng</td>
<td>Anlong Samnar</td>
<td>104.2274314</td>
<td>12.94743614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>213</td>
<td>6.2</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi KrÖng</td>
<td>Anlong Samnar</td>
<td>104.2274314</td>
<td>12.94743614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>214</td>
<td>2.2</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi KrÖng</td>
<td>Anlong Samnar</td>
<td>104.2274314</td>
<td>12.94743614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>215</td>
<td>4.8</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi KrÖng</td>
<td>Anlong Samnar</td>
<td>104.2274314</td>
<td>12.94743614</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>216</td>
<td>9.6</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Bakong</td>
<td>103.9894232</td>
<td>13.35739817</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>217</td>
<td>11.0</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nik™m</td>
<td>Chan Sar</td>
<td>104.1103659</td>
<td>13.37891805</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>218</td>
<td>0.3</td>
<td>F</td>
<td>2014</td>
<td>Banteay Meanchey</td>
<td>Preah Nžr Preah</td>
<td>Chob Yeari</td>
<td>103.1649315</td>
<td>13.66601257</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>219</td>
<td>6.6</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Chong Knies</td>
<td>103.8401794</td>
<td>13.25630554</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>220</td>
<td>1.3</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Chong Knies</td>
<td>103.8401794</td>
<td>13.25630554</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>221</td>
<td>12.3</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soe SiPom</td>
<td>Chrouy Neang Nuon</td>
<td>103.5517513</td>
<td>13.8336277</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>222</td>
<td>8.3</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Darn Daek</td>
<td>104.1184591</td>
<td>13.25764664</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>223</td>
<td>10.3</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.0859347</td>
<td>13.18512897</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>224</td>
<td>10.7</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.0859347</td>
<td>13.18512897</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>225</td>
<td>10.0</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Dan Run</td>
<td>104.0859347</td>
<td>13.18512897</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>226</td>
<td>10.7</td>
<td>M</td>
<td>2014</td>
<td>Kampong Thom</td>
<td>Stoung</td>
<td>Kampong Chen Tboung</td>
<td>104.5560628</td>
<td>12.93187723</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>227</td>
<td>1.2</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.1403332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>228</td>
<td>7.2</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.1403332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>229</td>
<td>1.4</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.1403332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No.</td>
<td>Age</td>
<td>Gender</td>
<td>Year</td>
<td>District</td>
<td>Village</td>
<td>Latitude</td>
<td>Longitude</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>-----</td>
<td>--------</td>
<td>------</td>
<td>----------</td>
<td>---------</td>
<td>----------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>230</td>
<td>2.7</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>231</td>
<td>4.7</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>232</td>
<td>6.9</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>233</td>
<td>1.9</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>234</td>
<td>6.5</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>235</td>
<td>7.9</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>236</td>
<td>2.6</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>237</td>
<td>8.4</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>238</td>
<td>11.9</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nikom</td>
<td>Kampong Khleang</td>
<td>104.140332</td>
<td>13.04884688</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>239</td>
<td>14.1</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Kandaek</td>
<td>103.9296375</td>
<td>13.3029341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>240</td>
<td>10.4</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Kandaek</td>
<td>103.9296375</td>
<td>13.3029341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>241</td>
<td>7.7</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Kandaek</td>
<td>103.9296375</td>
<td>13.3029341</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>242</td>
<td>9.1</td>
<td>F</td>
<td>2013</td>
<td>Battambang</td>
<td>Prasat Bak™ng</td>
<td>Kaoh Chiveang</td>
<td>103.5919063</td>
<td>13.2188273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>243</td>
<td>1.7</td>
<td>F</td>
<td>2013</td>
<td>Battambang</td>
<td>Prasat Bak™ng</td>
<td>Kaoh Chiveang</td>
<td>103.5919063</td>
<td>13.2188273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>244</td>
<td>12.3</td>
<td>F</td>
<td>2013</td>
<td>Battambang</td>
<td>Prasat Bak™ng</td>
<td>Kaoh Chiveang</td>
<td>103.5919063</td>
<td>13.2188273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td>1.7</td>
<td>F</td>
<td>2013</td>
<td>Battambang</td>
<td>Prasat Bak™ng</td>
<td>Kaoh Chiveang</td>
<td>103.5919063</td>
<td>13.2188273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>246</td>
<td>3.0</td>
<td>F</td>
<td>2014</td>
<td>Battambang</td>
<td>Prasat Bak™ng</td>
<td>Kaoh Chiveang</td>
<td>103.5919063</td>
<td>13.2188273</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>247</td>
<td>7.7</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Sotr Nôt™m</td>
<td>Khchas</td>
<td>104.0609096</td>
<td>13.25652359</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>248</td>
<td>12.0</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Sotr Nôt™m</td>
<td>Khchas</td>
<td>104.0609096</td>
<td>13.25652359</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>249</td>
<td>11.4</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Svay Leu</td>
<td>Phnum</td>
<td>104.1521159</td>
<td>13.55969797</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>6.2</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Sotr Nôt™m</td>
<td>Kien Sangkao</td>
<td>104.1590757</td>
<td>13.2118358</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>251</td>
<td>8.6</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Kouk Chak</td>
<td>103.8395359</td>
<td>13.4319055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>252</td>
<td>4.2</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Kouk Chak</td>
<td>103.8395359</td>
<td>13.4319055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>253</td>
<td>12.4</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Kouk Chak</td>
<td>103.8395359</td>
<td>13.4319055</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>254</td>
<td>13.0</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>ìngk™ng</td>
<td>Kouk Doung</td>
<td>103.790108</td>
<td>13.7038741</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>255</td>
<td>12.0</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Lvea</td>
<td>103.6574108</td>
<td>13.41912907</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>256</td>
<td>9.0</td>
<td>M</td>
<td>2014</td>
<td>Banteay Meanchey</td>
<td>O Chrauv</td>
<td>Poipet</td>
<td>102.659645</td>
<td>13.642996</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>257</td>
<td>13.0</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Soth Nik™m</td>
<td>Poipet</td>
<td>104.1615677</td>
<td>13.3951848</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>258</td>
<td>14.3</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soth Nik™m</td>
<td>Poipet</td>
<td>104.1615677</td>
<td>13.3951848</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>259</td>
<td>5.5</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Poak</td>
<td>Poak</td>
<td>103.5286413</td>
<td>13.37723486</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>260</td>
<td>7.8</td>
<td>F</td>
<td>2014</td>
<td>Battambang</td>
<td>Battambang</td>
<td>Preaek</td>
<td>103.2117684</td>
<td>13.08410463</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>261</td>
<td>0.0</td>
<td>F</td>
<td>2014</td>
<td>Preah Vihear</td>
<td>Rotanak</td>
<td>Robieb</td>
<td>105.0841422</td>
<td>13.38869225</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>262</td>
<td>11.2</td>
<td>M</td>
<td>2013</td>
<td>Battambang</td>
<td>Battambang</td>
<td>Preaek</td>
<td>103.2150136</td>
<td>13.09952554</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>263</td>
<td>6.6</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Chi Krong</td>
<td>Ruessei Lok</td>
<td>104.2258928</td>
<td>13.25354713</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>264</td>
<td>3.1</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Chi Krong</td>
<td>Ruessei Lok</td>
<td>104.2258928</td>
<td>13.25354713</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>265</td>
<td>6.7</td>
<td>F</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sala Kamraeuk</td>
<td>103.8650309</td>
<td>13.347982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>266</td>
<td>4.4</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sala Kamraeuk</td>
<td>103.8650309</td>
<td>13.347982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>267</td>
<td>6.8</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sala Kamraeuk</td>
<td>103.8650309</td>
<td>13.347982</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>268</td>
<td>11.3</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Siem Reab</td>
<td>103.8437332</td>
<td>13.30476315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>269</td>
<td>11.5</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Siem Reab</td>
<td>103.8437332</td>
<td>13.30476315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>270</td>
<td>10.5</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Siem Reab</td>
<td>103.8437332</td>
<td>13.30476315</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>271</td>
<td>11.1</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sla Kram</td>
<td>103.8836083</td>
<td>13.37003502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>272</td>
<td>11.7</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sla Kram</td>
<td>103.8836083</td>
<td>13.37003502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>273</td>
<td>8.1</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sla Kram</td>
<td>103.8836083</td>
<td>13.37003502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>274</td>
<td>13.3</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sla Kram</td>
<td>103.8836083</td>
<td>13.37003502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>275</td>
<td>5.5</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Sla Kram</td>
<td>103.8836083</td>
<td>13.37003502</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>276</td>
<td>10.0</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Varin</td>
<td>Srae Nisy</td>
<td>104.0130835</td>
<td>13.83855219</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>277</td>
<td>11.0</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Svay Dangkum</td>
<td>103.8202681</td>
<td>13.313428</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>278</td>
<td>10.0</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Svay Dangkum</td>
<td>103.8202681</td>
<td>13.313428</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>279</td>
<td>9.6</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Svay Dangkum</td>
<td>103.8202681</td>
<td>13.313428</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>280</td>
<td>9.5</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Svay Dangkum</td>
<td>103.8202681</td>
<td>13.313428</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>281</td>
<td>6.1</td>
<td>M</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Siem RŽab</td>
<td>Svay Dangkum</td>
<td>103.8202681</td>
<td>13.313428</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>282</td>
<td>9.4</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Soir Nâ™m</td>
<td>Ta Yaek</td>
<td>104.1847881</td>
<td>13.30881913</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>283</td>
<td>15.3</td>
<td>M</td>
<td>2013</td>
<td>Siem Reap</td>
<td>Prasat Bak™ng</td>
<td>Trapeang Thum</td>
<td>104.044328</td>
<td>13.30570753</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>284</td>
<td>7.5</td>
<td>F</td>
<td>2014</td>
<td>Siem Reap</td>
<td>Puok</td>
<td>Tuck Vil</td>
<td>103.8049637</td>
<td>13.41569467</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Appendix D

*Salmonella* Typhi isolates and their corresponding metadata in chapter 6

<table>
<thead>
<tr>
<th>ID</th>
<th>Study_code</th>
<th>Date</th>
<th>Sex</th>
<th>Age_year</th>
<th>Year</th>
<th>Acute/Carrier</th>
<th>Genotype</th>
<th>H58 status</th>
<th>gyrA_mutation</th>
</tr>
</thead>
<tbody>
<tr>
<td>01TY052</td>
<td>01TY</td>
<td>4-Nov-08</td>
<td>Male</td>
<td>4</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY061</td>
<td>01TY</td>
<td>30-Nov-08</td>
<td>Male</td>
<td>16</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY065</td>
<td>01TY</td>
<td>2-Dec-08</td>
<td>Female</td>
<td>23</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY066</td>
<td>01TY</td>
<td>9-Dec-08</td>
<td>Female</td>
<td>11</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY069</td>
<td>01TY</td>
<td>11-Dec-08</td>
<td>Male</td>
<td>24</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY072</td>
<td>01TY</td>
<td>8-Dec</td>
<td>Male</td>
<td>18</td>
<td>2008</td>
<td>Acute</td>
<td>4.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY073</td>
<td>01TY</td>
<td>16-Dec-08</td>
<td>Female</td>
<td>24</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY075</td>
<td>01TY</td>
<td>8-Dec</td>
<td>Male</td>
<td>14</td>
<td>2008</td>
<td>Acute</td>
<td>3.2.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY077</td>
<td>01TY</td>
<td>26-Dec-08</td>
<td>Female</td>
<td>6</td>
<td>2008</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY080</td>
<td>01TY</td>
<td>9-Jan-09</td>
<td>Male</td>
<td>0</td>
<td>2009</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY083</td>
<td>01TY</td>
<td>14-Jan-09</td>
<td>Male</td>
<td>6</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY086</td>
<td>01TY</td>
<td>30-Jan-09</td>
<td>Male</td>
<td>13</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY088</td>
<td>01TY</td>
<td>9-Feb-09</td>
<td>Male</td>
<td>19</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY090</td>
<td>01TY</td>
<td>9-Feb</td>
<td>Female</td>
<td>29</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY095</td>
<td>01TY</td>
<td>19-Feb-09</td>
<td>Male</td>
<td>16</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY096</td>
<td>01TY</td>
<td>20-Feb-09</td>
<td>Male</td>
<td>15</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY098</td>
<td>01TY</td>
<td>9-Feb</td>
<td>Female</td>
<td>21</td>
<td>2009</td>
<td>Acute</td>
<td>3.2.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY099</td>
<td>01TY</td>
<td>24-Feb-09</td>
<td>Female</td>
<td>19</td>
<td>2009</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY101</td>
<td>01TY</td>
<td>9-Feb</td>
<td>Male</td>
<td>24</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY102</td>
<td>01TY</td>
<td>2-Mar-09</td>
<td>Male</td>
<td>18</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY103</td>
<td>01TY</td>
<td>9-Mar</td>
<td>Female</td>
<td>14</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY104</td>
<td>01TY</td>
<td>9-Mar</td>
<td>Male</td>
<td>15</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY105</td>
<td>01TY</td>
<td>9-Mar-09</td>
<td>Male</td>
<td>22</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY107</td>
<td>01TY</td>
<td>11-Mar-09</td>
<td>Male</td>
<td>21</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>------------</td>
<td>------</td>
<td>----</td>
<td>------</td>
<td>-------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>01TY112</td>
<td>01TY</td>
<td>9-Mar</td>
<td>Female</td>
<td>15</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY122</td>
<td>01TY</td>
<td>9-Apr</td>
<td>Female</td>
<td>11</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY128</td>
<td>01TY</td>
<td>3-May-09</td>
<td>Male</td>
<td>21</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY132</td>
<td>01TY</td>
<td>17-May-09</td>
<td>Female</td>
<td>28</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY134</td>
<td>01TY</td>
<td>18-May-09</td>
<td>Female</td>
<td>14</td>
<td>2009</td>
<td>Acute</td>
<td>3.0.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY139</td>
<td>01TY</td>
<td>14-Jun-09</td>
<td>Male</td>
<td>19</td>
<td>2009</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY142</td>
<td>01TY</td>
<td>19-Jun-09</td>
<td>Male</td>
<td>15</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY150</td>
<td>01TY</td>
<td>25-Jun-09</td>
<td>Female</td>
<td>29</td>
<td>2009</td>
<td>Acute</td>
<td>2.2.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY151</td>
<td>01TY</td>
<td>25-Jun-09</td>
<td>Female</td>
<td>6</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY152</td>
<td>01TY</td>
<td>25-Jun-09</td>
<td>Male</td>
<td>28</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY154</td>
<td>01TY</td>
<td>8-Jul-09</td>
<td>Female</td>
<td>17</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY157</td>
<td>01TY</td>
<td>12-Jul-09</td>
<td>Female</td>
<td>24</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY163</td>
<td>01TY</td>
<td>17-Jul-09</td>
<td>Female</td>
<td>27</td>
<td>2009</td>
<td>Acute</td>
<td>3.0.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY170</td>
<td>01TY</td>
<td>24-Jul-09</td>
<td>Female</td>
<td>17</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY172</td>
<td>01TY</td>
<td>26-Jul-09</td>
<td>Male</td>
<td>10</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY191</td>
<td>01TY</td>
<td>11-Aug-09</td>
<td>Male</td>
<td>18</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY205</td>
<td>01TY</td>
<td>28-Aug-09</td>
<td>Male</td>
<td>6</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY208</td>
<td>01TY</td>
<td>30-Aug-09</td>
<td>Female</td>
<td>28</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY211</td>
<td>01TY</td>
<td>30-Aug-09</td>
<td>Female</td>
<td>4</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY216</td>
<td>01TY</td>
<td>1-Sep-09</td>
<td>Male</td>
<td>25</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY218</td>
<td>01TY</td>
<td>3-Sep-09</td>
<td>Male</td>
<td>11</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY226</td>
<td>01TY</td>
<td>10-Sep-09</td>
<td>Female</td>
<td>7</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY230</td>
<td>01TY</td>
<td>1-Oct-09</td>
<td>Female</td>
<td>11</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY233</td>
<td>01TY</td>
<td>4-Oct-09</td>
<td>Male</td>
<td>7</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY242</td>
<td>01TY</td>
<td>22-Oct-09</td>
<td>Male</td>
<td>3</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY256</td>
<td>01TY</td>
<td>26-Nov-09</td>
<td>Male</td>
<td>8</td>
<td>2009</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>Patient ID</td>
<td>Type</td>
<td>Date</td>
<td>Gender</td>
<td>Age</td>
<td>Year</td>
<td>Status</td>
<td>Age Group</td>
<td>Genotype</td>
<td>Other Genotypes</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>------------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
<td>-----------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>01TY257</td>
<td>01TY</td>
<td>30-Nov-09</td>
<td>Male</td>
<td>19</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY258</td>
<td>01TY</td>
<td>2-Dec-09</td>
<td>Male</td>
<td>19</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY261</td>
<td>01TY</td>
<td>6-Dec-09</td>
<td>Male</td>
<td>7</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY264</td>
<td>01TY</td>
<td>10-Dec-09</td>
<td>Male</td>
<td>21</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY267</td>
<td>01TY</td>
<td>14-Dec-09</td>
<td>Female</td>
<td>4</td>
<td>2009</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY274</td>
<td>01TY</td>
<td>18-Jan-10</td>
<td>Male</td>
<td>9</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY283</td>
<td>01TY</td>
<td>5-Mar-10</td>
<td>Male</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY286</td>
<td>01TY</td>
<td>14-Mar-10</td>
<td>Male</td>
<td>16</td>
<td>2010</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY301</td>
<td>01TY</td>
<td>5-Apr-10</td>
<td>Male</td>
<td>16</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY305</td>
<td>01TY</td>
<td>11-Apr-10</td>
<td>Female</td>
<td>22</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY311</td>
<td>01TY</td>
<td>18-Apr-10</td>
<td>Female</td>
<td>9</td>
<td>2010</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY312</td>
<td>01TY</td>
<td>18-Apr-10</td>
<td>Male</td>
<td>18</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY313</td>
<td>01TY</td>
<td>18-Apr-10</td>
<td>Male</td>
<td>13</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY315</td>
<td>01TY</td>
<td>20-Apr-10</td>
<td>Female</td>
<td>11</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY327</td>
<td>01TY</td>
<td>4-May-10</td>
<td>Male</td>
<td>0</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY332</td>
<td>01TY</td>
<td>16-May-10</td>
<td>Female</td>
<td>14</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY337</td>
<td>01TY</td>
<td>21-May-10</td>
<td>Male</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY342</td>
<td>01TY</td>
<td>2-Jun-10</td>
<td>Male</td>
<td>0</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY347</td>
<td>01TY</td>
<td>10-Jun-10</td>
<td>Female</td>
<td>20</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY349</td>
<td>01TY</td>
<td>14-Jun-10</td>
<td>Male</td>
<td>14</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY366</td>
<td>01TY</td>
<td>6-Jul-10</td>
<td>Female</td>
<td>9</td>
<td>2010</td>
<td>Acute</td>
<td>3.2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY370</td>
<td>01TY</td>
<td>8-Jul-10</td>
<td>Male</td>
<td>18</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY371</td>
<td>01TY</td>
<td>8-Jul-10</td>
<td>Male</td>
<td>12</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY374</td>
<td>01TY</td>
<td>11-Jul-10</td>
<td>Female</td>
<td>8</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>01TY375</td>
<td>01TY</td>
<td>12-Jul-10</td>
<td>Male</td>
<td>8</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>01TY391</td>
<td>01TY</td>
<td>21-Jul-10</td>
<td>Male</td>
<td>13</td>
<td>2010</td>
<td>Acute</td>
<td>3.1</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY394</td>
<td>01TY</td>
<td>23-Jul-10</td>
<td>Male</td>
<td>8</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>Ref</td>
<td>GB</td>
<td>Country</td>
<td>Date</td>
<td>Age</td>
<td>Year</td>
<td>Acute</td>
<td>Subtype</td>
<td>Site</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>01TY397</td>
<td>01TY</td>
<td>23-Jul-10</td>
<td>Male</td>
<td>20</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>01TY405</td>
<td>01TY</td>
<td>1-Aug-10</td>
<td>Female</td>
<td>4</td>
<td>2010</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY408</td>
<td>01TY</td>
<td>2-Aug-10</td>
<td>Male</td>
<td>9</td>
<td>2010</td>
<td>Acute</td>
<td>2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY410</td>
<td>01TY</td>
<td>3-Aug-10</td>
<td>Female</td>
<td>31</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY421</td>
<td>01TY</td>
<td>9-Aug-10</td>
<td>Male</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY426</td>
<td>01TY</td>
<td>10-Aug-10</td>
<td>Male</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>01TY432</td>
<td>01TY</td>
<td>13-Aug-10</td>
<td>Male</td>
<td>19</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY433</td>
<td>01TY</td>
<td>15-Aug-10</td>
<td>Male</td>
<td>19</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY441</td>
<td>01TY</td>
<td>18-Aug-10</td>
<td>Male</td>
<td>28</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY443</td>
<td>01TY</td>
<td>19-Aug-10</td>
<td>Male</td>
<td>18</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY444</td>
<td>01TY</td>
<td>20-Aug-10</td>
<td>Female</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY445</td>
<td>01TY</td>
<td>20-Aug-10</td>
<td>Male</td>
<td>18</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY447</td>
<td>01TY</td>
<td>22-Aug-10</td>
<td>Female</td>
<td>8</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY448</td>
<td>01TY</td>
<td>22-Aug-10</td>
<td>Female</td>
<td>13</td>
<td>2010</td>
<td>Acute</td>
<td>3.2.2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY472</td>
<td>01TY</td>
<td>9-Sep-10</td>
<td>Male</td>
<td>13</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>01TY476</td>
<td>01TY</td>
<td>13-Sep-10</td>
<td>Male</td>
<td>22</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY481</td>
<td>01TY</td>
<td>19-Sep-10</td>
<td>Male</td>
<td>17</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>01TY484</td>
<td>01TY</td>
<td>22-Sep-10</td>
<td>Male</td>
<td>16</td>
<td>2010</td>
<td>Acute</td>
<td>3.3</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>01TY492</td>
<td>01TY</td>
<td>3-Oct-10</td>
<td>Male</td>
<td>13</td>
<td>2010</td>
<td>Acute</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>GB003</td>
<td>GB</td>
<td>12-Jun-09</td>
<td>Female</td>
<td>40</td>
<td>2009</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>GB005</td>
<td>GB</td>
<td>5-Nov-07</td>
<td>Female</td>
<td>34</td>
<td>2007</td>
<td>Carrier</td>
<td>3.2.2</td>
<td>non-H58</td>
<td>D87N</td>
</tr>
<tr>
<td>GB026</td>
<td>GB</td>
<td>23-Dec-07</td>
<td>Female</td>
<td>27</td>
<td>2007</td>
<td>Carrier</td>
<td>2</td>
<td>non-H58</td>
<td>WT</td>
</tr>
<tr>
<td>GB031</td>
<td>GB</td>
<td>26-Jun-09</td>
<td>Female</td>
<td>53</td>
<td>2009</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>GB044</td>
<td>GB</td>
<td>6-Dec-07</td>
<td>Female</td>
<td>30</td>
<td>2007</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83Y</td>
</tr>
<tr>
<td>GB071</td>
<td>GB</td>
<td>22-Mar-08</td>
<td>Female</td>
<td>42</td>
<td>2008</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
<tr>
<td>GB076</td>
<td>GB</td>
<td>17-Jul-09</td>
<td>Female</td>
<td>35</td>
<td>2009</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>WT</td>
</tr>
<tr>
<td>GB125</td>
<td>GB</td>
<td>21-Aug-09</td>
<td>Female</td>
<td>25</td>
<td>2009</td>
<td>Carrier</td>
<td>4.3.1</td>
<td>H58</td>
<td>S83F</td>
</tr>
</tbody>
</table>

224
<table>
<thead>
<tr>
<th>Code</th>
<th>Country</th>
<th>Date</th>
<th>Gender</th>
<th>Age</th>
<th>Year</th>
<th>Carrier</th>
<th>Strain</th>
</tr>
</thead>
<tbody>
<tr>
<td>GB169</td>
<td>GB</td>
<td>14-Sep-09</td>
<td>Female</td>
<td>27</td>
<td>2009</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB199</td>
<td>GB</td>
<td>29-Sep-07</td>
<td>Female</td>
<td>15</td>
<td>2007</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB266</td>
<td>GB</td>
<td>14-Mar-08</td>
<td>Male</td>
<td>50</td>
<td>2008</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB276</td>
<td>GB</td>
<td>1-May-08</td>
<td>Female</td>
<td>27</td>
<td>2008</td>
<td>Carrier</td>
<td>3.3</td>
</tr>
<tr>
<td>GB281</td>
<td>GB</td>
<td>4-May-08</td>
<td>Female</td>
<td>27</td>
<td>2008</td>
<td>Carrier</td>
<td>2.2.2</td>
</tr>
<tr>
<td>GB335</td>
<td>GB</td>
<td>6-Jun-08</td>
<td>Female</td>
<td>33</td>
<td>2008</td>
<td>Carrier</td>
<td>3.3</td>
</tr>
<tr>
<td>GB336</td>
<td>GB</td>
<td>10-Jun-08</td>
<td>Female</td>
<td>30</td>
<td>2008</td>
<td>Carrier</td>
<td>3.3</td>
</tr>
<tr>
<td>GB368</td>
<td>GB</td>
<td>30-Jun-08</td>
<td>Female</td>
<td>30</td>
<td>2008</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB387</td>
<td>GB</td>
<td>8-Feb-10</td>
<td>Female</td>
<td>60</td>
<td>2010</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB428</td>
<td>GB</td>
<td>11-Mar-10</td>
<td>Female</td>
<td>34</td>
<td>2010</td>
<td>Carrier</td>
<td>2.2</td>
</tr>
<tr>
<td>GB441</td>
<td>GB</td>
<td>19-Mar-10</td>
<td>Male</td>
<td>38</td>
<td>2010</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB496</td>
<td>GB</td>
<td>22-Apr-10</td>
<td>Female</td>
<td>20</td>
<td>2010</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB580</td>
<td>GB</td>
<td>12-Dec-08</td>
<td>Female</td>
<td>54</td>
<td>2008</td>
<td>Carrier</td>
<td>2.2</td>
</tr>
<tr>
<td>GB650</td>
<td>GB</td>
<td>6-Aug-10</td>
<td>Female</td>
<td>34</td>
<td>2010</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>GB705</td>
<td>GB</td>
<td>16-Sep-10</td>
<td>Male</td>
<td>23</td>
<td>2010</td>
<td>Carrier</td>
<td>3.2.2</td>
</tr>
<tr>
<td>GB710</td>
<td>GB</td>
<td>20-Sep-10</td>
<td>Female</td>
<td>67</td>
<td>2010</td>
<td>Carrier</td>
<td>4.3.1</td>
</tr>
<tr>
<td>SNP Site</td>
<td>Ref base</td>
<td>SNP base</td>
<td>ancestral Codon</td>
<td>derived Codon</td>
<td>Ancestral AA</td>
<td>Derived AA</td>
<td>CDS name</td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>----------</td>
<td>----------------</td>
<td>--------------</td>
<td>--------------</td>
<td>------------</td>
<td>----------</td>
</tr>
<tr>
<td>2952</td>
<td>N</td>
<td>G</td>
<td>T</td>
<td>CGA</td>
<td>CTA</td>
<td>R</td>
<td>L</td>
</tr>
<tr>
<td>25425</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GAT</td>
<td>AAT</td>
<td>D</td>
<td>N</td>
</tr>
<tr>
<td>47769</td>
<td>N</td>
<td>T</td>
<td>C</td>
<td>GTG</td>
<td>GCG</td>
<td>V</td>
<td>A</td>
</tr>
<tr>
<td>50775</td>
<td>N</td>
<td>A</td>
<td>G</td>
<td>ATG</td>
<td>ACG</td>
<td>M</td>
<td>T</td>
</tr>
<tr>
<td>54513</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCG</td>
<td>GTG</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>58654</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GGC</td>
<td>AGC</td>
<td>G</td>
<td>S</td>
</tr>
<tr>
<td>62410</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>CGG</td>
<td>CAG</td>
<td>R</td>
<td>Q</td>
</tr>
<tr>
<td>63682</td>
<td>N</td>
<td>G</td>
<td>T</td>
<td>CGC</td>
<td>AGC</td>
<td>R</td>
<td>S</td>
</tr>
<tr>
<td>71918</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCG</td>
<td>GTG</td>
<td>A</td>
<td>V</td>
</tr>
<tr>
<td>83223</td>
<td>N</td>
<td>A</td>
<td>G</td>
<td>TCC</td>
<td>CCC</td>
<td>S</td>
<td>P</td>
</tr>
<tr>
<td>92753</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>AGC</td>
<td>AAC</td>
<td>S</td>
<td>N</td>
</tr>
<tr>
<td>101269</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>CGC</td>
<td>TGC</td>
<td>R</td>
<td>C</td>
</tr>
<tr>
<td>103365</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCG</td>
<td>ACG</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>103932</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCG</td>
<td>ACG</td>
<td>A</td>
<td>T</td>
</tr>
<tr>
<td>113304</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>CCG</td>
<td>CTG</td>
<td>P</td>
<td>L</td>
</tr>
<tr>
<td>118304</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>CCT</td>
<td>CTT</td>
<td>P</td>
<td>L</td>
</tr>
<tr>
<td>144484</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GTT</td>
<td>ATT</td>
<td>V</td>
<td>I</td>
</tr>
<tr>
<td>151468</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GGC</td>
<td>GAC</td>
<td>G</td>
<td>D</td>
</tr>
<tr>
<td>Accession</td>
<td>Source</td>
<td>Description</td>
<td>Protein Function</td>
<td>Mod Factor</td>
<td>Log2 Ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>--------</td>
<td>-------------</td>
<td>------------------</td>
<td>------------</td>
<td>------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0156</td>
<td></td>
<td>preprotein translocase SecA subunit</td>
<td>Transport Other</td>
<td>4.A.6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0173</td>
<td></td>
<td>aromatic amino acid transport protein AroP</td>
<td>Transport amino acid and amines</td>
<td>4.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0194</td>
<td></td>
<td>hypothetical ABC transporter ATP-binding protein</td>
<td>Transport/binding proteins</td>
<td>4.A</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0225</td>
<td></td>
<td>putative membrane protein</td>
<td>Membranes</td>
<td>3.C.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0238</td>
<td></td>
<td>methionine aminopeptidase</td>
<td>Protein translation and modification</td>
<td>3.A.8</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0244</td>
<td></td>
<td>undecaprenyl pyrophosphate synthetase</td>
<td>Sugar-nucleotide biosynthesis, conversions</td>
<td>1.C.3</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0254</td>
<td></td>
<td>DNA polymerase III, alpha chain</td>
<td>DNA-replication, repair, restriction/repair</td>
<td>3.B.4</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0259</td>
<td></td>
<td>lysine decarboxylase</td>
<td>Pyruvate dehydrogenase</td>
<td>1.B.2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0288</td>
<td></td>
<td>hypothetical protein SPI-6</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0290</td>
<td></td>
<td>conserved hypothetical protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0300</td>
<td></td>
<td>hypothetical protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0307</td>
<td></td>
<td>hypothetical protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0319</td>
<td></td>
<td>Rhs-family protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0324</td>
<td></td>
<td>Rhs-family protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0328</td>
<td></td>
<td>hypothetical protein</td>
<td>Hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0355</td>
<td></td>
<td>phosphoheptose isomerase</td>
<td>Surface polysaccharides &amp; antigens</td>
<td>3.C.2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0396</td>
<td></td>
<td>type III restriction-modification system StyLTI enzyme mod</td>
<td>DNA-replication, repair, restriction/repair</td>
<td>3.A.7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0397</td>
<td></td>
<td>ferrioxamine B receptor precursor</td>
<td>Transport Other</td>
<td>4.A.6</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0429</td>
<td></td>
<td>exonuclease SbcC</td>
<td>Degradation of DNA</td>
<td>3.B.2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0481</td>
<td></td>
<td>probable membrane component of 2-aminoethylphosphonate transporter</td>
<td>Transport amino acid and amines</td>
<td>4.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY0481</td>
<td></td>
<td>cytochrome o ubiquinol oxidase C subunit</td>
<td>Aerobic Respiration</td>
<td>1.B.7</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene</td>
<td>Accession</td>
<td>Protein Name</td>
<td>Function</td>
<td>Score</td>
<td>Fold Change</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>-----------</td>
<td>--------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>-------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>490825</td>
<td>N C T GTA</td>
<td>AmpG protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>490875</td>
<td>N G A GCG</td>
<td>AmpG protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>502112</td>
<td>N G A GTA</td>
<td>peptidyl-prolyl cis-trans isomerase D</td>
<td>3.A.8 Protein translation and modification</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>518336</td>
<td>N G A CCC</td>
<td>ylaB</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>520037</td>
<td>N C A GGG</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>540170</td>
<td>N G A ATG</td>
<td>ferrochelatase</td>
<td>1.G.12 Heme and porphyrin</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>554496</td>
<td>N C T GGC</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>7</td>
<td>7.291666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>554650</td>
<td>N G A GCA</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>556290</td>
<td>N T G GTA</td>
<td>putative ABC transporter ATP-binding protein</td>
<td>3.A.8 Protein translation and modification</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>557966</td>
<td>N G A CCA</td>
<td>ybbN</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>589822</td>
<td>N C A GCT</td>
<td>phosphoribosylaminomimidazole carboxylase ATPase</td>
<td>1.F.1 Purine ribonucleotide biosynthesis</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>596512</td>
<td>N A C AAG</td>
<td>fimA type-1 filamential protein, a chain precursor</td>
<td>3.C.3 Surface structures</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>608472</td>
<td>N C T GGG</td>
<td>grtB</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>617891</td>
<td>2 C T TGG</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>618849</td>
<td>N A G TAC</td>
<td>outer membrane esterase</td>
<td>4.1 Pathogenicity</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>621564</td>
<td>N G A TCC</td>
<td>oxygen-insensitive NAD(P)H nitroreductase</td>
<td>5.D Drug/Analogue sensitivity</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>624836</td>
<td>N C T GGC</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>625772</td>
<td>N C T GAT</td>
<td>ferrienterobactin receptor precursor</td>
<td>4.A.6 Transport Other</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>639491</td>
<td>N T C CAG</td>
<td>ferrienterobactin-binding periplasmic protein</td>
<td>4.A.6 Transport Other</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>641681</td>
<td>N G A GTC</td>
<td>2,3-dihydroxybenzoate-AMP ligase</td>
<td>1.G.14 Iron uptake and storage</td>
<td>5</td>
<td>5.208333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>643446</td>
<td>N G T GCG</td>
<td>isochorismutase</td>
<td>1.G.14 Iron uptake and storage</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>648137</td>
<td>N C T CGC</td>
<td>putative oxidoreductase</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>664501</td>
<td>N A G CTG</td>
<td>regulator of nucleoside diphosphate kinase</td>
<td>1.F.5 Miscellaneous</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Type</td>
<td>Description</td>
<td>Location</td>
<td>Log2 Fold Change</td>
<td>p-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
<td>--------------------------------------------------</td>
<td>----------</td>
<td>-----------------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>670557</td>
<td></td>
<td>4.3-carboxylate anerobic carrier</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>676249</td>
<td></td>
<td>Central Intermediary metabolism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>678229</td>
<td></td>
<td>putative membrane protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>691764</td>
<td></td>
<td>conserved hypothetical protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>692954</td>
<td></td>
<td>DNA polymerase III, delta subunit</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>698331</td>
<td></td>
<td>probable permease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>715716</td>
<td></td>
<td>asparagine synthetase B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>722451</td>
<td></td>
<td>pta system, N-acetylglucosamine-specific IIABC component</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>724837</td>
<td></td>
<td>glutaminyl-tRNA synthetase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>725967</td>
<td></td>
<td>putative outer membrane protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>737776</td>
<td></td>
<td>possible 5-nitroimidazole antibiotic resistance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>750451</td>
<td></td>
<td>deoxyribonucleotidylase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>752359</td>
<td></td>
<td>PTR2-family transport protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>752996</td>
<td></td>
<td>PTR2-family transport protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>761027</td>
<td></td>
<td>putative glycosyl transferase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>761098</td>
<td></td>
<td>putative glycosyl transferase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>770444</td>
<td></td>
<td>putative membrane protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>771204</td>
<td></td>
<td>citrate synthase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>848432</td>
<td></td>
<td>HlyD-family secretion protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>862862</td>
<td></td>
<td>putative membrane protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>883410</td>
<td></td>
<td>putative ABC transporter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>884918</td>
<td></td>
<td>hypothetical ABC transporter permease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>888699</td>
<td></td>
<td>putative oxidoreductase</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>901044</td>
<td></td>
<td>ribosomal protein S6 modification protein</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accession</td>
<td>Type</td>
<td>Description</td>
<td>Function</td>
<td>E-value</td>
<td>Score</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>---------</td>
<td>--------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>T</td>
<td>C GETG GCG V A STY0910</td>
<td>putrescine-binding periplasmic protein precursor</td>
<td>4.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>ACG ATG T M STY0911</td>
<td>putrescine transport ATP-binding protein PotG</td>
<td>4.A.1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>CCG CTG P L STY0921</td>
<td>arginine transport system permease protein ArtM</td>
<td>4.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>GCC ACC A T STY0931</td>
<td>pyruvate dehydrogenase</td>
<td>1.A.1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>GAC AAC D N STY0954</td>
<td>transport ATP-binding protein CydC</td>
<td>4.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GCC ACC A T STY0958</td>
<td>cell division protein FtsK</td>
<td>4.C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GGT AGT G S ycaD</td>
<td>probable transport protein</td>
<td>4.A</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>A G</td>
<td>AAT GAT N D STY0979</td>
<td>putative lipoprotein</td>
<td>3.C.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>CCG TCG P S lpxK</td>
<td>tetraacyl-diaccharide 4'-kinase</td>
<td>3.C.2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>GAC AAC D N STY1002</td>
<td>outer membrane protein F precursor</td>
<td>3.C</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GCC ACC A T STY1078</td>
<td>aminopeptidase N</td>
<td>3.B.3</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>T G</td>
<td>AAC ACC N T pipA</td>
<td>conserved hypothetical protein</td>
<td>SPI-5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>TGG TAG W * pipD</td>
<td>putative secreted peptidase</td>
<td>SPI-5</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>CAT TAT H Y lpaA</td>
<td>4-hydroxyphenylacetate 3-monoxygenase operon regulatory protein</td>
<td>1.A.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GGT AGT G S STY1167</td>
<td>conserved hypothetical protein</td>
<td>5.H.b</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>T C</td>
<td>GTT GCT V A STY1170</td>
<td>putative oxidoreductase</td>
<td>5.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GTC ATC V I STY1172</td>
<td>putative 2-hydroxyacid dehydrogenase in pho-csgg intergenic region</td>
<td>5.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>T C</td>
<td>TGG CGG W R STY1175</td>
<td>putative membrane protein</td>
<td>3.C.1</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>CCG CTG P L STY1176</td>
<td>assembly/transport component in curli production</td>
<td>3.C.3</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>C T</td>
<td>ACG AAC S N STY1186</td>
<td>glucans biosynthesis protein</td>
<td>3.F</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>CAT TAT H Y STY1199</td>
<td>conserved hypothetical protein</td>
<td>5.H.a</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GGC AGC G S STY1236</td>
<td>3-oxoacyl-[acyl-carrier-protein] synthase II</td>
<td>1.H</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>G A</td>
<td>GGC AGC G S STY1240</td>
<td>DNA polymerase III, delta' subunit</td>
<td>3.A.7</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Type</td>
<td>Description</td>
<td>Function</td>
<td>Score</td>
<td>Log Odds Ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>-------------</td>
<td>----------</td>
<td>-------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1194365</td>
<td>N A G GAG GGG E G</td>
<td>DNA polymerase III, delta' subunit</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1195346</td>
<td>N G A GCC ACC A T</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>4</td>
<td>4.166666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1208580</td>
<td>N C T GGT AGT G S</td>
<td>transcription-repair coupling factor (TrcF)</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1210415</td>
<td>N C T CGT CAT R H</td>
<td>transcription-repair coupling factor (TrcF)</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1216818</td>
<td>N G A GTG ATG V M</td>
<td>cobB putative regulatory protein</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1221269</td>
<td>2 C T TGG TAG W *</td>
<td>spermidine/putrescine transport system permease protein PotB</td>
<td>4.A.1 Transport amino acid and amines</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1230675</td>
<td>N C T GCG ACG A T</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>3</td>
<td>3.125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1233131</td>
<td>N G A CCC TCC P S</td>
<td>putative ribosomal large subunit pseudouridine synthase B</td>
<td>3.A.3 Ribosomes - maturation and modification</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1239073</td>
<td>N A G AGT GGT S G</td>
<td>putative invasin</td>
<td>4.I Pathogenicity</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1239209</td>
<td>N C T CCT CTM P</td>
<td>putative invasin</td>
<td>4.I Pathogenicity</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1283533</td>
<td>N C A TGA TTA L</td>
<td>indole-3-glycerol phosphate synthase</td>
<td>1.D.4 Aromatic Amino Acid</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1288407</td>
<td>N A G ACC GCC T A</td>
<td>transcriptional regulatory protein TyrR</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1306213</td>
<td>N C A TGG TAG W *</td>
<td>putative mandelate racemase/muconate lactonizing enzyme family protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1342083</td>
<td>N C T GGA GAA G E</td>
<td>putative transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Start</td>
<td>End</td>
<td>Blat Score</td>
<td>Legacy</td>
<td>Description</td>
<td>E-Value</td>
<td>Log2FoldChange</td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>-------</td>
<td>------------</td>
<td>--------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td>STY1399</td>
<td>1351519</td>
<td>1351519</td>
<td>60</td>
<td>conserved hypothetical protein</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1408</td>
<td>1358778</td>
<td>1358778</td>
<td>60</td>
<td>putative chemo-receptor protein</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1427</td>
<td>1378120</td>
<td>1378120</td>
<td>60</td>
<td>acyl carrier protein phosphodiesterase</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1444</td>
<td>1397389</td>
<td>1397389</td>
<td>60</td>
<td>ATP-dependent helicase HrpA</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1451</td>
<td>1399474</td>
<td>1399474</td>
<td>60</td>
<td>putative glycolate oxidase</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1452</td>
<td>1402707</td>
<td>1402707</td>
<td>60</td>
<td>putative ribulose-5-phosphate 3-epimerase</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1460</td>
<td>1410199</td>
<td>1410199</td>
<td>60</td>
<td>putative phosphotransferase enzyme</td>
<td>Hypothetical protein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1490</td>
<td>1424538</td>
<td>1424538</td>
<td>60</td>
<td>putative peptide</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
<td>2</td>
<td>2.0833333333</td>
</tr>
<tr>
<td>STY1492</td>
<td>1446984</td>
<td>1446984</td>
<td>60</td>
<td>putative regulatory protein</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1494</td>
<td>1453048</td>
<td>1453048</td>
<td>60</td>
<td>putative protein</td>
<td>1.B.7.a Aerobic Respiration</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1502</td>
<td>1458457</td>
<td>1458457</td>
<td>60</td>
<td>putative secreted protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1504</td>
<td>1462308</td>
<td>1462308</td>
<td>60</td>
<td>putative hydrolase</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1505</td>
<td>1463445</td>
<td>1463445</td>
<td>60</td>
<td>putative glycogen debranching protein</td>
<td>3.B.4 Degradations of polysaccharides</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1509</td>
<td>1467534</td>
<td>1467534</td>
<td>60</td>
<td>putative protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
</tr>
<tr>
<td>STY1520</td>
<td>1485774</td>
<td>1485774</td>
<td>60</td>
<td>putative ABC transporter periplasmic binding protein</td>
<td>4.A Transport/binding proteins</td>
<td>4</td>
<td>6.25</td>
</tr>
</tbody>
</table>

232
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Codon</th>
<th>Amino Acid</th>
<th>Protein Name</th>
<th>Functional Category</th>
<th>Description</th>
<th>Identifier</th>
<th>Score</th>
<th>E-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>NGAAGT</td>
<td>GGTAGT</td>
<td>GS</td>
<td>putative secreted stress response protein</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGACCG</td>
<td>CTGCTG</td>
<td>PL</td>
<td>fumarate hydratase class II</td>
<td>1.3 B Tricarboxylic acid cycle</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGGC</td>
<td>AGCAGG</td>
<td>GS</td>
<td>putative ferredoxin-like protein, cytoplasmic membrane</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NGAGCG</td>
<td>GTAGTA</td>
<td>V</td>
<td>putative perredoxin-like protein</td>
<td>1.3 Central Intermediary metabolism</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

233
<table>
<thead>
<tr>
<th>Accession</th>
<th>Symbol</th>
<th>Description</th>
<th>Functions</th>
<th>E-value</th>
<th>Fold Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1836701</td>
<td>N T C</td>
<td>GAG GGG</td>
<td>STY1947</td>
<td>ribonuclease D</td>
<td>3.B.1 Degradation of RNA</td>
</tr>
<tr>
<td>1843733</td>
<td>N G A</td>
<td>GGT AGT</td>
<td>STY1954</td>
<td>para-aminobenzoate synthase component I</td>
<td>1.G.2 Folic Acid</td>
</tr>
<tr>
<td>1856466</td>
<td>N G A</td>
<td>CCT TCT</td>
<td>STY1976</td>
<td>tail-specific protease precursor</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
</tr>
<tr>
<td>1863892</td>
<td>N C T</td>
<td>GGT AGT</td>
<td>STY2004</td>
<td>putative hydrolase</td>
<td>5.1 Unknown</td>
</tr>
<tr>
<td>1884909</td>
<td>N G A</td>
<td>CGT TGT</td>
<td>STY2085</td>
<td>oligopeptidase</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
</tr>
<tr>
<td>1905241</td>
<td>N T A</td>
<td>ATC TTC</td>
<td>STY2087</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
</tr>
<tr>
<td>1951699</td>
<td>N C T</td>
<td>GGC GAC</td>
<td>STY2098</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
</tr>
<tr>
<td>1965470</td>
<td>N G A</td>
<td>GAC AAC</td>
<td>STY2114</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
</tr>
<tr>
<td>1996037</td>
<td>2 C T</td>
<td>TGG TAG</td>
<td>STY2146</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
</tr>
<tr>
<td>1996313</td>
<td>N C T</td>
<td>CGG CAG</td>
<td>STY2146</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
</tr>
<tr>
<td>1997612</td>
<td>N T C</td>
<td>ACC GCC</td>
<td>STY2149</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
</tr>
<tr>
<td>2001948</td>
<td>N G T</td>
<td>CCC ACC</td>
<td>STY2154</td>
<td>excinuclease ABC subunit C</td>
<td>3.B.2 Degradation of DNA</td>
</tr>
<tr>
<td>2013806</td>
<td>N T A</td>
<td>CTG CAG</td>
<td>STY2174</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
</tr>
<tr>
<td>2018745</td>
<td>N T C</td>
<td>TCC S S</td>
<td>STY2218</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
</tr>
<tr>
<td>2053440</td>
<td>N G A</td>
<td>CCG TCG</td>
<td>STY2219</td>
<td>nicotinate-nucleotide–dimethylbenzimidazole phosphoribosyltransferase</td>
<td>1.G.13 Cobalamin</td>
</tr>
<tr>
<td>2077040</td>
<td>N A G</td>
<td>AAA GAA</td>
<td>STY2276</td>
<td>diol dehydratase small subunit</td>
<td>1.A.1 Degradation of carbohydrates</td>
</tr>
<tr>
<td>2102317</td>
<td>N G A</td>
<td>GCA ACA</td>
<td>STY2310</td>
<td>putative G-antigen transporter</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
</tr>
<tr>
<td>2121684</td>
<td>N G A</td>
<td>TCT TTT</td>
<td>STY2320</td>
<td>putative glycosyltransferase</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
</tr>
<tr>
<td>2128153</td>
<td>N T C</td>
<td>AAT GAT</td>
<td>STY2330</td>
<td>putative glycosyltransferase</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
</tr>
<tr>
<td>2139553</td>
<td>N A G</td>
<td>TTC CTC</td>
<td>STY2340</td>
<td>GDP-fucose synthetase</td>
<td>1.C.3 Sugar-nucleotide biosynthesis, conversions</td>
</tr>
<tr>
<td>2149716</td>
<td>N G A</td>
<td>CCG CTG</td>
<td>STY2350</td>
<td>putative tyrosine-protein kinase</td>
<td>3.A.8 Protein translation and modification</td>
</tr>
<tr>
<td>Gene</td>
<td>Start</td>
<td>End</td>
<td>Description</td>
<td>Function</td>
<td>Identity</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
<td>-------</td>
<td>-----------------------------------</td>
<td>-------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>putative tyrosine-protein kinase</td>
<td>3.A.8 Protein translation and modification</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative tyrosine-protein kinase</td>
<td>3.A.8 Protein translation and modification</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative RND-family transporter protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative RND-family transporter protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>7</td>
<td>7.291666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative sugar kinase</td>
<td>1.C Central Intermediary metabolism</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative gntR-family transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative fimbrial chaperone protein</td>
<td>3.C.3 Surface structures</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical lipoprotein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical lipoprotein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical lipoprotein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative two-component system response regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>hypothetical lipoprotein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>0.041666667</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

235
<table>
<thead>
<tr>
<th>Accession</th>
<th>Gene Name</th>
<th>Function</th>
<th>Location</th>
<th>Description</th>
<th>Class</th>
<th>Function</th>
<th>E-value</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>STY2436</td>
<td>putative transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>119</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2437</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>148</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2441</td>
<td>1-phosphofructokinase</td>
<td>1.B.1 Glycolysis</td>
<td>173</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2443</td>
<td>sugar efflux transporter</td>
<td>4.A Transport/binding proteins</td>
<td>208</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2444</td>
<td>anaerobic glycerol-3-phosphate dehydrogenase subunit A</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>235</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2450</td>
<td>NADH dehydrogenase I chain J</td>
<td>1.B.7.a Aerobic Respiration</td>
<td>263</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2455</td>
<td>putative amino acid decarboxylase</td>
<td>1.D Amino Acid Biosynthesis</td>
<td>291</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2477</td>
<td>conserved hypothetical protein</td>
<td>5.I Unknown</td>
<td>319</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2528</td>
<td>putative lipopolysaccharide</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>346</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2530</td>
<td>NADH dehydrogenase I chain G</td>
<td>1.B.7.a Aerobic Respiration</td>
<td>374</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2550</td>
<td>putative amino acid decarboxylase</td>
<td>1.D Amino Acid Biosynthesis</td>
<td>402</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2564</td>
<td>glucokinase</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>429</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2568</td>
<td>nucleoside permease NupC</td>
<td>4.A.6 Transport Other</td>
<td>457</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2567</td>
<td>cell division protein</td>
<td>4.C Cell division</td>
<td>485</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2567</td>
<td>cell division protein</td>
<td>4.C Cell division</td>
<td>513</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2578</td>
<td>sulphate transport ATP-binding protein CysA</td>
<td>4.A.5 Transport Anions</td>
<td>541</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2579</td>
<td>sulphate transport system permease protein CysW</td>
<td>4.A.5 Transport Anions</td>
<td>569</td>
<td>1.041666667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Chromosome Position</td>
<td>Description</td>
<td>Function/Category</td>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------</td>
<td>----------------------------------</td>
<td>--------------------------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2697</td>
<td>2531257</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2711</td>
<td>2545678</td>
<td>transketolase</td>
<td>1.B.5.b Non-oxidative branch</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2713</td>
<td>2560281</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2741</td>
<td>2578401</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2742</td>
<td>258401</td>
<td>putative exported protein</td>
<td>1.C Central Intermediary metabolism</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2760</td>
<td>2606065</td>
<td>putative exported protein</td>
<td>4.I Pathogenicity</td>
<td>7.29166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2764</td>
<td>2616492</td>
<td>putative GTP-binding protein</td>
<td>5.I Unknown</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2769</td>
<td>2621661</td>
<td>putative DNA-binding protein</td>
<td>5.I Unknown</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2770</td>
<td>2623817</td>
<td>putative hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2771</td>
<td>2624579</td>
<td>putative exported protein</td>
<td>1.C Central Intermediary metabolism</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2811</td>
<td>2671118</td>
<td>putative sensor kinase protein</td>
<td>2 Broad regulatory function</td>
<td>5.20833333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2815</td>
<td>2678943</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2820</td>
<td>2683796</td>
<td>putative transmembrane transport protein</td>
<td>4.A Transport/binding proteins</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2820</td>
<td>2684372</td>
<td>putative transmembrane transport protein</td>
<td>4.A Transport/binding proteins</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2832</td>
<td>2694835</td>
<td>sigma-E factor negative regulatory protein</td>
<td>2 Broad regulatory function</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2857</td>
<td>2727753</td>
<td>phospho-2-dehydro-3-deoxyheptonate aldolase, tyr-sensitive</td>
<td>1.D.4 Aromatic Amino Acid</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2875</td>
<td>274407</td>
<td>large repetitive protein</td>
<td>SPI-9</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2875</td>
<td>2746073</td>
<td>large repetitive protein</td>
<td>SPI-9</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY2875</td>
<td>2747171</td>
<td>large repetitive protein</td>
<td>SPI-9</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene Name</td>
<td>Description</td>
<td>Function</td>
<td>Protein ID</td>
<td>Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>----------</td>
<td>------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gab protein homolog</td>
<td>putative GAB DTP gene cluster repressor</td>
<td>5.1 Unknown</td>
<td>STY2910</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>putative GAB DTP gene cluster repressor</td>
<td>GabA permease (4-amino butyrate transport carrier)</td>
<td>4.A.3 Transport carbohydrates, organic acids and alcohols</td>
<td>STY2913</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gamma-glutamylcysteine synthetase</td>
<td>1.G.10 Thioredoxin</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RNA polymerase sigma subunit RpoS (sigma-38)</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP sulfurylase (ATP:sulfate adenyltransferase)</td>
<td>1.C.5 Sulphur Metabolism</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP sulfurylase (ATP:sulfate adenyltransferase)</td>
<td>1.C.5 Sulphur Metabolism</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ATP sulfurylase (ATP:sulfate adenyltransferase)</td>
<td>1.C.5 Sulphur Metabolism</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>enolase</td>
<td>1.B.1 Glycolysis</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outer membrane usher protein</td>
<td>3.C Cell envelope</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fimbrial subunit</td>
<td>3.C.3 Surface structures</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L-fucose permease</td>
<td>4.A.3 Transport carbohydrates, organic acids and alcohols</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fucos operon facC protein</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>5</td>
<td>5.208333333</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>l-fucose operon activator</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>0.04166667</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GenBank Accession</td>
<td>Nucleotide</td>
<td>Protein</td>
<td>Function</td>
<td>GO Term</td>
<td>Count</td>
<td>E-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------</td>
<td>---------</td>
<td>-----------------------------------------------</td>
<td>----------------------------------------------</td>
<td>-------</td>
<td>---------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3002057</td>
<td>N G A CCG CTG</td>
<td>P L recB</td>
<td>exonuclease V subunit</td>
<td>3.B.2 Degradation of DNA</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3006403</td>
<td>N G A CCG TCG</td>
<td>P S STY3133</td>
<td>protease III precursor (pilin)</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3021088</td>
<td>N G A GTT ATT</td>
<td>V I STY3149</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3025188</td>
<td>N G A CCG TCG</td>
<td>P S STY3153</td>
<td>2-acylglycerophosphoethanolamine acyl carrier protein synthetase</td>
<td>1.H Fatty acid biosynthesis</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3035138</td>
<td>N G A CGC TGC R C STY3162</td>
<td>2-keto-3-deoxygluconate oxidoreductase</td>
<td>1.H Fatty acid biosynthesis</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3039476</td>
<td>N G A ATG ATA M I STY3167</td>
<td>probable amino acid transport protein</td>
<td>4.A.1 Transport amino acid and amines</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3046278</td>
<td>2 G A CAG TAG Q * stdB</td>
<td>probable outer membrane fimbrial usher protein</td>
<td>3.C.3 Surface structures</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3068338</td>
<td>N G A AGC AAC S N STY3201</td>
<td>flavodoxin II</td>
<td>1.B.7.a Aerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3072213</td>
<td>N A C ACC CCC T P STY3207</td>
<td>6-phospho-beta-glucosidase</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3088543</td>
<td>N G A GTG ATG V M STY3220</td>
<td>chromosome initiation inhibitor</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3098424</td>
<td>N T A CTC CAC L H STY3232</td>
<td>possible ABC-transport protein, ATP-binding component</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3111928</td>
<td>N T C TCG CCG S P STY3247</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3128126</td>
<td>N G A GGG AGG G R STY3268</td>
<td>nucleoside permease</td>
<td>4.A.6 Transport Other</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3128789</td>
<td>N C T CCG TCG P S STY3268</td>
<td>nucleoside permease</td>
<td>4.A.6 Transport Other</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3136486</td>
<td>N G A GTA ATA V I STY3285</td>
<td>putative exported protein</td>
<td>SPI-8</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3138646</td>
<td>N G A GCA GTA A V STY3290</td>
<td>hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3140066</td>
<td>N G A GCC ACC A T STY3293</td>
<td>LysR-family transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3147262</td>
<td>2 G A CAG TAG Q * STY3300</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3151115</td>
<td>N G A GCG GTG A V STY3305</td>
<td>hexuronate transporter</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3183175</td>
<td>2 G A TGG TAG W * STY3340</td>
<td>possible membrane transport protein</td>
<td>4.A Transport/binding proteins</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3195738</td>
<td>N C T AGC AAC S N STY3351</td>
<td>topoisomerase IV subunit A</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3217295</td>
<td>N C T GCG GTG A V STY3377</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Position</td>
<td>Sequence</td>
<td>Description</td>
<td>Location</td>
<td>Function</td>
<td>Impact Factor</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>----------</td>
<td>-------------</td>
<td>----------</td>
<td>----------</td>
<td>---------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3219040</td>
<td>N G A GCG ACG A T STY3378</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3224368</td>
<td>N G A TCA TTA S L STY3380</td>
<td>adenyl-transferase</td>
<td>3.A.8 Protein translation and modification</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3246839</td>
<td>N T C AAT GAT N D STY3400</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3247366</td>
<td>N G A GCC ACC A T STY3401</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3254191</td>
<td>N G A CGT CAT R H STY3408</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3258933</td>
<td>2 C T CAA TAA Q * STY3416</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3259259</td>
<td>N G A GCG ACG A T STY3417</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3288370</td>
<td>N G A GCC ACC A T STY3444</td>
<td>galactitol-1-phosphate dehydrogenase</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3292861</td>
<td>N T C TTT TCT F S yraN</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3361086</td>
<td>N G T CCG ACG P T STY3519</td>
<td>putative static acid transporter</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3371173</td>
<td>N G A GAA AAA E K degS</td>
<td>serine protease</td>
<td>3.B.3 Degradation of proteins, peptides, glycoproteins</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3382903</td>
<td>2 C T CAG TAG Q * STY3543</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3444980</td>
<td>N T G ATT AGT I S STY3592</td>
<td>putative hydrolase</td>
<td>5.I Unknown</td>
<td>5</td>
<td>5.20833333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3455162</td>
<td>N G A CGC TGC R C recQ</td>
<td>ATP-dependent DNA helicase</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3506652</td>
<td>N C T CAT TAT H Y STY3649</td>
<td>LysR-family regulatory protein for ilvC expression</td>
<td>1.D.6 Pyruvate Family</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3507253</td>
<td>N C T CCG CTG P L STY3649</td>
<td>LysR-family regulatory protein for ilvC expression</td>
<td>1.D.6 Pyruvate Family</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3508762</td>
<td>N A C TCT GTCT S A STY3652</td>
<td>threonine deaminase</td>
<td>1.D.6 Pyruvate Family</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3566895</td>
<td>N G A CCA TCA P S nfi</td>
<td>putative endonuclease</td>
<td>3.B.2 Degradation of DNA</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3570766</td>
<td>N G A GTC ATC V I STY3721</td>
<td>thiamine biosynthesis protein</td>
<td>1.G.8 Thiamine</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3574272</td>
<td>N C T CGT TGT R C STY3725</td>
<td>thiamine biosynthesis protein</td>
<td>1.G.8 Thiamine</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3603732</td>
<td>N G A GCG ACG A T STY3745</td>
<td>tRNA (uracil-5)-methyltransferase</td>
<td>3.A.5 Aminomocyl tRNA synthetase</td>
<td>5</td>
<td>5.20833333</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3614668</td>
<td>N G A TGC TAC C Y STY3754</td>
<td>phosphoenolpyruvate carboxylase</td>
<td>1.B.8 Fermentation</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3616596</td>
<td>N C T CCC TCC P S STY3755</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3634106</td>
<td>N C T GGG GAG A T STY3768</td>
<td>bifunctional aspartokinase II/homoserine dehydrogenase II can I write</td>
<td>1.D.2 Aspartate Family</td>
<td>7</td>
<td>7.29166667</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GenBank ID</td>
<td>Gene</td>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Fold Change</td>
<td>E-value</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>-----------</td>
<td>------------------------------------------------------------------------------</td>
<td>---------------------------------</td>
<td>-------------</td>
<td>-----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3642604</td>
<td>STY3776</td>
<td>NGA CGA</td>
<td>transcriptional repressor</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3653696</td>
<td>STY3788</td>
<td>NGA GGA</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3656589</td>
<td>STY3792</td>
<td>NAT CTTG</td>
<td>putative aldolase</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3669388</td>
<td>STY3806</td>
<td>NGA CGC</td>
<td>putative membrane protein</td>
<td>3.C.i Membranes lipoprotein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3700304</td>
<td>STY3842</td>
<td>NGA CGC</td>
<td>formate dehydrogenase-O, major subunit</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3702097</td>
<td>STY3842</td>
<td>NGA CGC</td>
<td>formate dehydrogenase-O, gamma subunit</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3714939</td>
<td>STY3859</td>
<td>NAT CTTG</td>
<td>putative glycosyl hydrolase</td>
<td>5.1 Unknown</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3719502</td>
<td>STY3861</td>
<td>NGA GCC</td>
<td>putative membrane permease</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3719775</td>
<td>STY3862</td>
<td>NGA GCC</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3729652</td>
<td>STY3874</td>
<td>NAT CGG</td>
<td>glutamine synthetase</td>
<td>1.D.1 Glutamate Family</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3756565</td>
<td>STY3896</td>
<td>NGA CGC</td>
<td>high affinity ribose transport protein</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3757030</td>
<td>STY3896</td>
<td>NAT GGC</td>
<td>high affinity ribose transport protein</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3770081</td>
<td>STY3908</td>
<td>NAT GCC</td>
<td>ATP synthase subunit C</td>
<td>1.A Degradation</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3786029</td>
<td>STY3923</td>
<td>NAT ATG</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3810322</td>
<td>STY3943</td>
<td>NAT TTC</td>
<td>DNA gyrase subunit B</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>3</td>
<td>3.125</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3812138</td>
<td>STY3944</td>
<td>NAT GCA</td>
<td>putative LysR-family transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3828144</td>
<td>STY3957</td>
<td>NAT GAT</td>
<td>TorD protein</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3828146</td>
<td>STY3957</td>
<td>NAT GAT</td>
<td>TorD protein</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3828148</td>
<td>STY3957</td>
<td>NAT TTT</td>
<td>TorD protein</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3828149</td>
<td>STY3957</td>
<td>NAT TTT</td>
<td>TorD protein</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3848849</td>
<td>STY3981</td>
<td>NAT GGC</td>
<td>multidrug resistance protein</td>
<td>5.D Drug/Analogue sensitivity</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3865841</td>
<td>STY4000</td>
<td>NAT CGT</td>
<td>putative PTS system IIC component</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3872688</td>
<td>STY4009</td>
<td>NAT AAG</td>
<td>putative glycosyl hydrolase</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.04166667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3886919</td>
<td>STY4000</td>
<td>NAT TAC</td>
<td>Magnesium transport ATPase, P-type 2</td>
<td>SPI-3</td>
<td>2</td>
<td>2.08333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Strain</td>
<td>First Letter</td>
<td>Last Letter</td>
<td>Function</td>
<td>Description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4057</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>5.1 Unknown</td>
<td>Putative beta-lactamase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4106</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>3.C.1 Membranes</td>
<td>Putative lipoprotein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4115</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>5.1 Unknown</td>
<td>Putative sugar kinase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4118</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>2 Broad regulatory function</td>
<td>Putative transcriptional regulator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4148</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>5.1 Unknown</td>
<td>Putative acetyltransferase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4160</td>
<td>N</td>
<td>A</td>
<td>G</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>Putative amino acid permease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4173</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>4.A.1 Transport amines</td>
<td>Putative amino acid permease</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4193</td>
<td>N</td>
<td>A</td>
<td>A</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4193</td>
<td>N</td>
<td>A</td>
<td>A</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4196</td>
<td>N</td>
<td>A</td>
<td>A</td>
<td>2 Broad regulatory function</td>
<td>Hypothetical hyd-family transcriptional regulator</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4205</td>
<td>N</td>
<td>T</td>
<td>C</td>
<td>1.G.10 Thioredoxin</td>
<td>Glutathione reductase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4222</td>
<td>N</td>
<td>C</td>
<td>G</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4222</td>
<td>N</td>
<td>C</td>
<td>G</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4228</td>
<td>N</td>
<td>C</td>
<td>A</td>
<td>5.H.b Hypothetical protein</td>
<td>Conserved hypothetical protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4222</td>
<td>N</td>
<td>C</td>
<td>A</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4222</td>
<td>N</td>
<td>C</td>
<td>A</td>
<td>3.C.1 Membranes</td>
<td>Putative membrane protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4228</td>
<td>N</td>
<td>C</td>
<td>A</td>
<td>5.H.b Hypothetical protein</td>
<td>Conserved hypothetical protein</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>STY4136</td>
<td>N</td>
<td>G</td>
<td>G</td>
<td>1.C Central Intermediary metabolism</td>
<td>Putative gluconokinase</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

242
<table>
<thead>
<tr>
<th>Gene ID</th>
<th>Classification</th>
<th>Protein ID</th>
<th>Description</th>
<th>Expression Level</th>
<th>Functional Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>4137928</td>
<td></td>
<td>glgB</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>2</td>
<td>3.A.10 Polysaccharides - (cytoplasmic)</td>
</tr>
<tr>
<td>4137929</td>
<td></td>
<td>glgB</td>
<td>1,4-alpha-glucan branching enzyme</td>
<td>2</td>
<td>3.A.10 Polysaccharides - (cytoplasmic)</td>
</tr>
<tr>
<td>4141884</td>
<td></td>
<td>glgC</td>
<td>glucose-1-phosphate adenylyltransferase</td>
<td>1</td>
<td>3.A.10 Polysaccharides - (cytoplasmic)</td>
</tr>
<tr>
<td>4150649</td>
<td></td>
<td>glpR</td>
<td>glycerol-3-phosphate regulon repressor</td>
<td>1</td>
<td>1.B.7.b Anaerobic Respiration</td>
</tr>
<tr>
<td>4155227</td>
<td>maltodextrin phosphorylase</td>
<td>STY4282</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>2</td>
<td>2.083333333</td>
</tr>
<tr>
<td>4178175</td>
<td>putative hydrolase</td>
<td>yrlG</td>
<td>5.1 Unknown</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4190284</td>
<td>DamX protein</td>
<td>damX</td>
<td>4.C Cell division</td>
<td>6</td>
<td>6.25</td>
</tr>
<tr>
<td>4192170</td>
<td>ribulose-phosphate 3-epimerase</td>
<td>ype</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4219044</td>
<td>probable ABC transporter ATP-binding protein</td>
<td>yheS</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4225205</td>
<td>conserved hypothetical protein</td>
<td>yheO</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4249038</td>
<td>potassium transport protein</td>
<td>trkA</td>
<td>4.A.2 Transport Cations</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4255456</td>
<td>conserved hypothetical protein</td>
<td>STY4392</td>
<td>5.H.b Hypothetical protein</td>
<td>2</td>
<td>2.083333333</td>
</tr>
<tr>
<td>4282989</td>
<td>putative membrane protein</td>
<td>STY4415</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4289522</td>
<td>putative lipoprotein</td>
<td>STY4421</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.083333333</td>
</tr>
<tr>
<td>4293291</td>
<td>maltose transport inner membrane protein</td>
<td>STY4424</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>2</td>
<td>2.083333333</td>
</tr>
<tr>
<td>4296632</td>
<td>maltose/maltodextrin transport ATP-binding protein</td>
<td>STY4426</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4296800</td>
<td>maltose/maltodextrin transport ATP-binding protein</td>
<td>STY4426</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4301966</td>
<td>glycerol-3-phosphate acyltransferase</td>
<td>STY4431</td>
<td>3.A.11 Phospholipids</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4308235</td>
<td>putative exported protein</td>
<td>STY4438</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4337777</td>
<td>large repetitive protein</td>
<td>STY4459</td>
<td>SPI-4</td>
<td>2</td>
<td>2.083333333</td>
</tr>
<tr>
<td>4355566</td>
<td>putative membrane protein</td>
<td>STY4469</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4358445</td>
<td>putative membrane protein</td>
<td>STY4472</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4363926</td>
<td>cytochrome c-type biogenesis protein</td>
<td>STY4479</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4385697</td>
<td>arginine decarboxylase</td>
<td>adi</td>
<td>1.A.2 Degradation of amino acids</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>Accession</td>
<td>First Amino Acid</td>
<td>Second Amino Acid</td>
<td>Third Amino Acid</td>
<td>Fourth Amino Acid</td>
<td>Protein Name</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>N C T GGG AGG G R STY4510</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N G A GTA ATA V I STY4521</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N G A CGG CAG R Q STY4529</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N G A GCA ACA A T STY4534</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N C T GCT GTT A V ssB single strand binding protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N T C TCC CCC S P pilO putative pilus assembly protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N G A GCC GTC A V STY4579</td>
<td>putative membrane protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N A G ATC GTC I V STY4583</td>
<td>conserved hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N C T GCA GTA A V STY4584</td>
<td>conserved hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>2 G A TGG TAG W * STY4587</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
</tr>
<tr>
<td>N G A GAG AAG E K STY4594</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
</tr>
<tr>
<td>N A G TTT CTT F L tviE pilus biosynthesis protein TviE, Glycosyl transferses group 1</td>
<td>V2A operon</td>
<td>6</td>
<td>6.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N G A GCC GTC A V tviE pilus biosynthesis protein TviE, Glycosyl transferses group 1</td>
<td>V2A operon</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N A G TGT CGT C R tviE pilus biosynthesis protein TviE, Glycosyl transferses group 1</td>
<td>V2A operon</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N A G GTT GCT V A tviE pilus biosynthesis protein TviE, Glycosyl transferses group 1</td>
<td>V2A operon</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2 C A GGA TGA G * tviD pilus biosynthesis protein</td>
<td>V2A operon</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N T G AAT ACT N T tviD pilus biosynthesis protein</td>
<td>V2A operon</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N G A GCC GTC A V tviD pilus biosynthesis protein</td>
<td>V2A operon</td>
<td>2</td>
<td>2.083333333</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N G A CTG TTT L F yjeJ hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gene ID</td>
<td>Location</td>
<td>Gene</td>
<td>Transcript</td>
<td>Protein</td>
<td>Gene Function</td>
</tr>
<tr>
<td>---------</td>
<td>----------</td>
<td>------</td>
<td>------------</td>
<td>---------</td>
<td>---------------</td>
</tr>
<tr>
<td>4563879</td>
<td>N C T GCG GTG A V yjeM</td>
<td>putative amino acid permease</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4574488</td>
<td>N C T AGC AAC S N yjeS</td>
<td>putative 4Fe-4S binding protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4575322</td>
<td>N G A GGC AGC G S yjeF</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>6</td>
<td>6.25</td>
</tr>
<tr>
<td>4580561</td>
<td>N C T TCA TTA S L miaA</td>
<td>tRNA delta-2-isopentenyliyryophosphate (IPP) transferase</td>
<td>3.A.5 Aminoacyl tRNA synthetase</td>
<td>2</td>
<td>2.08333333</td>
</tr>
<tr>
<td>4580600</td>
<td>N G A GGA GAA G E miaA</td>
<td>tRNA delta-2-isopentenyliyryophosphate (IPP) transferase</td>
<td>3.A.5 Aminoacyl tRNA synthetase</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4595445</td>
<td>N C T GCT GTT A V aidB</td>
<td>probable acyl Co-A dehydrogenase</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4596506</td>
<td>2 G T GAA TAA E * aidB</td>
<td>probable acyl Co-A dehydrogenase</td>
<td>1.A.1 Degradation of carbohydrates</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4600253</td>
<td>N C T GGC AGC G S yjiR</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4621008</td>
<td>2 C T TGG TAG W * ytfJ</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>2</td>
<td>2.08333333</td>
</tr>
<tr>
<td>4627402</td>
<td>N C T CCG TCG P S ytfN</td>
<td>putative exported protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4650675</td>
<td>N C T GGA AGA G R nrdD</td>
<td>anaerobic ribonucleoside-triphosphate reductase</td>
<td>1.F.3 2'-Deoxyribonucleotide metabolism</td>
<td>2</td>
<td>2.08333333</td>
</tr>
<tr>
<td>4669035</td>
<td>N C T ACG ATG T M STY4808</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>2</td>
<td>2.08333333</td>
</tr>
<tr>
<td>4681321</td>
<td>N C T GAT AAT D N STY4819</td>
<td>regulatory protein</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4716667</td>
<td>N C T CGC TGC R C STY4853</td>
<td>hypothetical protein</td>
<td>SPI-10</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4724452</td>
<td>N C G GCC CCC A P STY4861</td>
<td>hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4728683</td>
<td>N C T TCG TTG S L STY4866</td>
<td>probable aspartate racemase</td>
<td>1.C. Central Intermediary metabolism</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4759459</td>
<td>N C T ATG ATA M I dnaC</td>
<td>probable DNA replication protein</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4773929</td>
<td>N C T ACA ATA T I STY4914</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4786646</td>
<td>N A T AGC TGC S C radA</td>
<td>putative DNA repair protein</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
<td>1.041666667</td>
</tr>
<tr>
<td>4788023</td>
<td>N G A GCC ACC A T nadR</td>
<td>conserved hypothetical transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>4</td>
<td>4.166666667</td>
</tr>
<tr>
<td>4798548</td>
<td>N A G AGA GGA R G STY4937</td>
<td>inner membrane protein CreD</td>
<td>5.B Colicin-related function</td>
<td>1</td>
<td>1.041666667</td>
</tr>
</tbody>
</table>
## Appendix F  
Carrier-specific nonsynonymous mutations and their functional classes

<table>
<thead>
<tr>
<th>SNP_Sit e</th>
<th>SNP_Sit e</th>
<th>S/ NS</th>
<th>Ref_ base</th>
<th>SNP_ base</th>
<th>ancestral Codon</th>
<th>derived Codon</th>
<th>ancestral AA</th>
<th>derived AA</th>
<th>CDS_na me</th>
<th>product</th>
<th>Functional_classes</th>
<th>Number_carri er_isolate</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2725</td>
<td>2725</td>
<td>N</td>
<td>G</td>
<td>T</td>
<td>GCA</td>
<td>TCA</td>
<td>A</td>
<td>S</td>
<td>STY0002</td>
<td>aspartokinase F:homoserine dehydrogenase I</td>
<td>1.D.2 Aspartate Family</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>122281</td>
<td>122281</td>
<td>N</td>
<td>C</td>
<td>A</td>
<td>CGC</td>
<td>CTC</td>
<td>R</td>
<td>L</td>
<td>STY0120</td>
<td>L-ribulokinase</td>
<td>1.A.1 Degradation of carbohydrate</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>135985</td>
<td>135985</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>ATG</td>
<td>ATA</td>
<td>M</td>
<td>I</td>
<td>STY0134</td>
<td>probable activator protein in lexABC operon</td>
<td>1.D.6 Pyruvate Family</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>204001</td>
<td>204001</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GCT</td>
<td>GTT</td>
<td>A</td>
<td>V</td>
<td>STY0193</td>
<td>carbonic anhydrase</td>
<td>5.I Unknown</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>241338</td>
<td>241338</td>
<td>2</td>
<td>G</td>
<td>T</td>
<td>GAA</td>
<td>TAA</td>
<td>E</td>
<td>*</td>
<td>STY0230</td>
<td>deoxyguanosine triphosphatase triphosphohydrolase</td>
<td>1.C Central intermediary metabolism</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>266603</td>
<td>266603</td>
<td>N</td>
<td>A</td>
<td>T</td>
<td>CAC</td>
<td>CTC</td>
<td>H</td>
<td>L</td>
<td>STY0253</td>
<td>ribonuclease III</td>
<td>3.B.1 Degradation of RNA</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>275632</td>
<td>275632</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GAT</td>
<td>AAT</td>
<td>D</td>
<td>N</td>
<td>STY0259</td>
<td>lysine decarboxylase</td>
<td>1.B.2 Pyruvate dehydrogenase</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>281721</td>
<td>281721</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GTC</td>
<td>ATC</td>
<td>V</td>
<td>I</td>
<td>STY0269</td>
<td>prolyl-tRNA synthetase</td>
<td>3.A.5 Amino acyl-tRNA synthesis; tRNA modification</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>308726</td>
<td>308726</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>ATG</td>
<td>ATA</td>
<td>M</td>
<td>I</td>
<td>STY0293</td>
<td>hypothetical protein</td>
<td>SPI-6</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>341744</td>
<td>341744</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GTG</td>
<td>ATG</td>
<td>V</td>
<td>M</td>
<td>safA</td>
<td>probable lipoprotein</td>
<td>SPI-6</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>343595</td>
<td>343595</td>
<td>N</td>
<td>T</td>
<td>C</td>
<td>TAC</td>
<td>CAC</td>
<td>Y</td>
<td>H</td>
<td>safB</td>
<td>periplasmic fimbral chaperone protein</td>
<td>SPI-6</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>355324</td>
<td>355324</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GGG</td>
<td>AGG</td>
<td>G</td>
<td>R</td>
<td>tsfC</td>
<td>outer membrane fimbral usher protein</td>
<td>SPI-6</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>378535</td>
<td>378535</td>
<td>2</td>
<td>C</td>
<td>A</td>
<td>TGC</td>
<td>TGA</td>
<td>C</td>
<td>*</td>
<td>STY0368</td>
<td>probable secreted protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>438939</td>
<td>438939</td>
<td>N</td>
<td>A</td>
<td>G</td>
<td>CTG</td>
<td>CCG</td>
<td>L</td>
<td>P</td>
<td>STY0429</td>
<td>exonuclease SbcC</td>
<td>3.B.2 Degradation of DNA</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>439695</td>
<td>439695</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>CGT</td>
<td>CAT</td>
<td>R</td>
<td>H</td>
<td>STY0429</td>
<td>exonuclease SbcC</td>
<td>3.B.2 Degradation of DNA</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>464238</td>
<td>464238</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCA</td>
<td>GTA</td>
<td>A</td>
<td>V</td>
<td>STY0458</td>
<td>thiamine-monomophosphate kinase</td>
<td>1.G.8 Thiamine</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>526144</td>
<td>526144</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GAT</td>
<td>AAT</td>
<td>D</td>
<td>N</td>
<td>STY0520</td>
<td>acrilavin resistance protein A precursor</td>
<td>5.D Drug/Analogue sensitivity</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>531394</td>
<td>531394</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GCC</td>
<td>ACC</td>
<td>A</td>
<td>T</td>
<td>STY0523</td>
<td>conserved hypothetical protein</td>
<td>5.II.b Hypothetical protein</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>582703</td>
<td>582703</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>GGC</td>
<td>AGC</td>
<td>G</td>
<td>S</td>
<td>allC</td>
<td>allantoate amidohydrolase</td>
<td>1.A.1 Degradation of carbohydrate</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>588547</td>
<td>588547</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GTG</td>
<td>ATG</td>
<td>V</td>
<td>M</td>
<td>arcC</td>
<td>carbamate kinase</td>
<td>1.A.1 Degradation of carbohydrate</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>599161</td>
<td>599161</td>
<td>2</td>
<td>C</td>
<td>T</td>
<td>CAG</td>
<td>TAG</td>
<td>Q</td>
<td>*</td>
<td>fimD</td>
<td>outer membrane usher protein Fim3 precursor</td>
<td>3.C.3 Surface structure</td>
<td>1</td>
<td>4.17</td>
</tr>
<tr>
<td>ID</td>
<td>Gene</td>
<td>Protein Description</td>
<td>Location</td>
<td>Type</td>
<td>Function</td>
<td>ID</td>
<td>Gene</td>
<td>Protein Description</td>
<td>Location</td>
<td>Type</td>
<td>Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>--------</td>
<td>-----------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>---------------------------</td>
<td>------</td>
<td>--------</td>
<td>-----------------------------------------------------------</td>
<td>----------</td>
<td>-------------</td>
<td>---------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>247</td>
<td>STY0661</td>
<td>molybdopterin-containing oxidoreductase membrane anchor subunit</td>
<td>5.I Unknown</td>
<td>1</td>
<td>4.17</td>
<td>661378</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>GAA AAA E K STY0661</td>
<td>molybdopterin-containing oxidoreductase membrane anchor subunit</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0701</td>
<td>putative hydroxase C-terminus</td>
<td>5.I Unknown</td>
<td>1</td>
<td>4.17</td>
<td>697519</td>
<td>N</td>
<td>G</td>
<td>T</td>
<td>GGC TGC G C STY0701</td>
<td>putative hydroxase C-terminus</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0703</td>
<td>putative sigma-54 dependent transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>4.17</td>
<td>699681</td>
<td>N</td>
<td>A</td>
<td>G</td>
<td>ACC GCC T A STY0703</td>
<td>putative sigma-54 dependent transcriptional regulator</td>
<td>2 Broad regulatory function</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0744</td>
<td>sensor protein KdpD</td>
<td>2 Broad regulatory function</td>
<td>1</td>
<td>4.17</td>
<td>745075</td>
<td>N</td>
<td>T</td>
<td>G</td>
<td>GAT GCT D A STY0744</td>
<td>sensor protein KdpD</td>
<td>2 Broad regulatory function</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0760</td>
<td>putative glycosyl transferase</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>1</td>
<td>4.17</td>
<td>760747</td>
<td>N</td>
<td>T</td>
<td>C</td>
<td>TCA CCA S P STY0760</td>
<td>putative glycosyl transferase</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0775</td>
<td>succinate dehydrogenase cytochrome b-556 subunit</td>
<td>1.B.3 Tricarboxylic acid cycle</td>
<td>1</td>
<td>4.17</td>
<td>773094</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>CGC CAC R H STY0775</td>
<td>succinate dehydrogenase cytochrome b-556 subunit</td>
<td>1.B.3 Tricarboxylic acid cycle</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0827</td>
<td>biotin synthetase</td>
<td>1.G.1 Biotin</td>
<td>1</td>
<td>4.17</td>
<td>796312</td>
<td>N</td>
<td>C</td>
<td>T</td>
<td>CGC CAC R H ybgR</td>
<td>biotin synthetase</td>
<td>1.G.1 Biotin</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY0929</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
<td>4.17</td>
<td>919324</td>
<td>2</td>
<td>C</td>
<td>T</td>
<td>CTA TAA * * STY0929</td>
<td>conserved hypothetical protein</td>
<td>5.H.b Hypothetical protein</td>
<td>1</td>
</tr>
<tr>
<td>247</td>
<td>STY1200</td>
<td>damage-inducible protein</td>
<td>5.I Unknown</td>
<td>1</td>
<td>4.17</td>
<td>1158627</td>
<td>N</td>
<td>G</td>
<td>A</td>
<td>CAT TAT H Y STY1200</td>
<td>damage-inducible protein</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>Chromosome Position</td>
<td>Gene Product</td>
<td>Function</td>
<td>Gene Description</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------------</td>
<td>--------------</td>
<td>----------</td>
<td>------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>248</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chromosome Position</th>
<th>Gene Product</th>
<th>Function</th>
<th>Gene Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1196367</td>
<td>N A G CAC CGC H R STY1242</td>
<td>PTS system, glucose-specific IIBC component</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
</tr>
<tr>
<td>1196978</td>
<td>N G A GCG ACG A T STY1242</td>
<td>PTS system, glucose-specific IIBC component</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
</tr>
<tr>
<td>1196990</td>
<td>N G A GGT AGT G S STY1242</td>
<td>PTS system, glucose-specific IIBC component</td>
<td>4.A.3 Transport Carbohydrates, organic acids and alcohols</td>
</tr>
<tr>
<td>1208535</td>
<td>N A T TGG AGG W R STY1256</td>
<td>transcription-repair coupling factor (TrpF)</td>
<td>3.A.7 DNA replication, repair, restriction/modification</td>
</tr>
<tr>
<td>1241292</td>
<td>N A G CTG CCG L P STY1286</td>
<td>nitrate/nitrite sensor protein NarX</td>
<td>1.B.7.b Anaerobic Respiration</td>
</tr>
<tr>
<td>1278803</td>
<td>N T G AAA CAA K Q STY1321</td>
<td>conserved hypothetical protein</td>
<td>5.Ia Hypothetical protein</td>
</tr>
<tr>
<td>1283955</td>
<td>N A T G TGA CGA/A GA R/R STY1326</td>
<td>indole-3-glycerolphosphate synthase</td>
<td>1.D.4 Aromatic Amino Acid</td>
</tr>
<tr>
<td>1289493</td>
<td>N T G GAT GAG D E STY1331</td>
<td>putative pseudouridine synthase</td>
<td>3.A.3 Ribosomes - maturation and modification</td>
</tr>
<tr>
<td>1296666</td>
<td>N G A GGC AGC G S STY1337</td>
<td>cys regulon transcriptional activator</td>
<td>1.D.3 Serine Family</td>
</tr>
<tr>
<td>1308498</td>
<td>N C T AGG AAG R K STY1349</td>
<td>conserved hypothetical protein</td>
<td>5.I Unknown</td>
</tr>
<tr>
<td>1313439</td>
<td>N C T GGC GAC G S STY1352</td>
<td>enoyl-[acyl-carrier-protein] reductase (NADH)</td>
<td>1.H Fatty acid biosynthesis</td>
</tr>
<tr>
<td>1318260</td>
<td>N G T CCG CAG P Q STY1356</td>
<td>peptide transport system ATP-binding protein SapD</td>
<td>4.A.6 Transport Other</td>
</tr>
<tr>
<td>1399300</td>
<td>N C T GAG AAG E K STY1447</td>
<td>putative ribulose-5-phosphate 3-epimerase</td>
<td>5.I Unknown</td>
</tr>
<tr>
<td>1399483</td>
<td>N T C ATT GTT I V STY1447</td>
<td>putative ribulose-5-phosphate 3-epimerase</td>
<td>5.I Unknown</td>
</tr>
<tr>
<td>1425078</td>
<td>N G A CCT TCT P S srFA</td>
<td>putative virulence effector protein</td>
<td>4.I Pathogenicity</td>
</tr>
<tr>
<td>1445458</td>
<td>N G A GCG GTG A V STY1489</td>
<td>nitrite extrusion protein</td>
<td>4.A Transport/binding proteins</td>
</tr>
<tr>
<td>1458466</td>
<td>N T C CTG CCG L P hyaF2</td>
<td>hydrogenase-1 operon protein HydroF2</td>
<td>1.B.7.a Aerobic Respiration</td>
</tr>
<tr>
<td>1536085</td>
<td>N C T GCG ACG A T STY1589</td>
<td>pyridine nucleotide transhydrogenase subunit-alpha</td>
<td>1.C Central intermediary metabolism</td>
</tr>
<tr>
<td>1589539</td>
<td>N C T GCG GTG A V STY1658</td>
<td>adenosine deaminase</td>
<td>1.F.4 Salvage of nucleotides and nucleotides</td>
</tr>
<tr>
<td>GenBank ID</td>
<td>Accession</td>
<td>Feature</td>
<td>Location</td>
</tr>
<tr>
<td>------------</td>
<td>-----------</td>
<td>---------</td>
<td>----------</td>
</tr>
<tr>
<td>1601143</td>
<td>1601143</td>
<td>NCT</td>
<td>CATTAT</td>
</tr>
<tr>
<td>1606270</td>
<td>1606270</td>
<td>NGAT</td>
<td>ATAPA</td>
</tr>
<tr>
<td>1656792</td>
<td>1656792</td>
<td>NGT</td>
<td>GATC</td>
</tr>
<tr>
<td>1701598</td>
<td>1701598</td>
<td>AAGAT</td>
<td>ATGAM</td>
</tr>
<tr>
<td>1721795</td>
<td>1721795</td>
<td>2GAC</td>
<td>TAGQ*</td>
</tr>
<tr>
<td>1725197</td>
<td>1725197</td>
<td>NCAT</td>
<td>ATGAM</td>
</tr>
<tr>
<td>1756933</td>
<td>1756933</td>
<td>NGTC</td>
<td>ATCA</td>
</tr>
<tr>
<td>1770310</td>
<td>1770310</td>
<td>NTCA</td>
<td>GAGGGG</td>
</tr>
<tr>
<td>1772818</td>
<td>1772818</td>
<td>NCAT</td>
<td>GATATT</td>
</tr>
<tr>
<td>1773651</td>
<td>1773651</td>
<td>NCTGCC</td>
<td>ACCAT</td>
</tr>
<tr>
<td>1813911</td>
<td>1813911</td>
<td>NAGACC</td>
<td>GCCAT</td>
</tr>
<tr>
<td>1861881</td>
<td>1861881</td>
<td>NCTC</td>
<td>TACG</td>
</tr>
<tr>
<td>1876919</td>
<td>1876919</td>
<td>NCCTGG</td>
<td>GAGGGG</td>
</tr>
<tr>
<td>1988113</td>
<td>1988113</td>
<td>NTGCGT</td>
<td>GCGAV</td>
</tr>
<tr>
<td>1995782</td>
<td>1995782</td>
<td>NGAGTC</td>
<td>ATCV</td>
</tr>
<tr>
<td>2006311</td>
<td>2006311</td>
<td>NGAGCG</td>
<td>GTGAV</td>
</tr>
<tr>
<td>2051229</td>
<td>2051229</td>
<td>NCTATG</td>
<td>ATAM</td>
</tr>
<tr>
<td>2054384</td>
<td>2054384</td>
<td>NACTTT</td>
<td>GTTFV</td>
</tr>
<tr>
<td>2066384</td>
<td>2066384</td>
<td>NCCTGG</td>
<td>CAGQR</td>
</tr>
<tr>
<td>2075855</td>
<td>2075855</td>
<td>NTCTAT</td>
<td>CATYH</td>
</tr>
<tr>
<td>2079993</td>
<td>2079993</td>
<td>NGAGCC</td>
<td>ACCAT</td>
</tr>
<tr>
<td>2120553</td>
<td>2120553</td>
<td>NGCTT</td>
<td>CATPH</td>
</tr>
<tr>
<td>2127235</td>
<td>2127235</td>
<td>NAGGTC</td>
<td>GCCAV</td>
</tr>
<tr>
<td>Accession</td>
<td>Start</td>
<td>End</td>
<td>ID</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>-----</td>
<td>----</td>
</tr>
<tr>
<td>2145889</td>
<td>2145</td>
<td>889</td>
<td>N</td>
</tr>
<tr>
<td>2151043</td>
<td>2151</td>
<td>043</td>
<td>N</td>
</tr>
<tr>
<td>2162476</td>
<td>2162</td>
<td>476</td>
<td>N</td>
</tr>
<tr>
<td>2162537</td>
<td>2162</td>
<td>537</td>
<td>N</td>
</tr>
<tr>
<td>2167085</td>
<td>2167</td>
<td>085</td>
<td>N</td>
</tr>
<tr>
<td>2176236</td>
<td>2176</td>
<td>236</td>
<td>N</td>
</tr>
<tr>
<td>2185609</td>
<td>2185</td>
<td>609</td>
<td>N</td>
</tr>
<tr>
<td>2191798</td>
<td>2191</td>
<td>798</td>
<td>N</td>
</tr>
<tr>
<td>220061</td>
<td>2200</td>
<td>61</td>
<td>N</td>
</tr>
<tr>
<td>2239555</td>
<td>2239</td>
<td>555</td>
<td>N</td>
</tr>
<tr>
<td>2268919</td>
<td>2268</td>
<td>919</td>
<td>N</td>
</tr>
<tr>
<td>2331540</td>
<td>2331</td>
<td>540</td>
<td>N</td>
</tr>
<tr>
<td>2333751</td>
<td>2333</td>
<td>751</td>
<td>N</td>
</tr>
<tr>
<td>2343416</td>
<td>2343</td>
<td>416</td>
<td>N</td>
</tr>
<tr>
<td>2350279</td>
<td>2350</td>
<td>279</td>
<td>N</td>
</tr>
<tr>
<td>2366071</td>
<td>2366</td>
<td>071</td>
<td>N</td>
</tr>
<tr>
<td>2400798</td>
<td>2400</td>
<td>798</td>
<td>N</td>
</tr>
<tr>
<td>2419570</td>
<td>2419</td>
<td>570</td>
<td>N</td>
</tr>
<tr>
<td>2463468</td>
<td>2463</td>
<td>468</td>
<td>N</td>
</tr>
<tr>
<td>2493604</td>
<td>2493</td>
<td>604</td>
<td>N</td>
</tr>
<tr>
<td>2493630</td>
<td>2493</td>
<td>630</td>
<td>N</td>
</tr>
<tr>
<td>2514750</td>
<td>2514</td>
<td>750</td>
<td>2</td>
</tr>
<tr>
<td>ID</td>
<td>StartPosition</td>
<td>EndPosition</td>
<td>Accession</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>-------------</td>
<td>-----------</td>
</tr>
<tr>
<td>251</td>
<td>2515747</td>
<td>2515747</td>
<td>STY2681</td>
</tr>
<tr>
<td>251</td>
<td>2533579</td>
<td>2533579</td>
<td>STY2699</td>
</tr>
<tr>
<td>251</td>
<td>2540803</td>
<td>2540803</td>
<td>STY2731</td>
</tr>
<tr>
<td>251</td>
<td>2567455</td>
<td>2567455</td>
<td>STY2735</td>
</tr>
<tr>
<td>251</td>
<td>2590406</td>
<td>2590406</td>
<td>STY2752</td>
</tr>
<tr>
<td>251</td>
<td>2609489</td>
<td>2609489</td>
<td>STY2760</td>
</tr>
<tr>
<td>251</td>
<td>2611167</td>
<td>2611167</td>
<td>STY2760</td>
</tr>
<tr>
<td>251</td>
<td>2611377</td>
<td>2611377</td>
<td>STY2760</td>
</tr>
<tr>
<td>251</td>
<td>2626146</td>
<td>2626146</td>
<td>STY2773</td>
</tr>
<tr>
<td>251</td>
<td>2630309</td>
<td>2630309</td>
<td>STY2800</td>
</tr>
<tr>
<td>251</td>
<td>2655783</td>
<td>2655783</td>
<td>STY2800</td>
</tr>
<tr>
<td>251</td>
<td>267249</td>
<td>267249</td>
<td>STY2813</td>
</tr>
<tr>
<td>251</td>
<td>2694734</td>
<td>2694734</td>
<td>STY2832</td>
</tr>
<tr>
<td>251</td>
<td>2704993</td>
<td>2704993</td>
<td>STY2843</td>
</tr>
<tr>
<td>251</td>
<td>2725661</td>
<td>2725661</td>
<td>STY2855</td>
</tr>
<tr>
<td>251</td>
<td>2733558</td>
<td>2733558</td>
<td>STY2864</td>
</tr>
<tr>
<td>251</td>
<td>2748017</td>
<td>2748017</td>
<td>STY2875</td>
</tr>
<tr>
<td>251</td>
<td>2805601</td>
<td>2805601</td>
<td>STY2932</td>
</tr>
<tr>
<td>251</td>
<td>2840369</td>
<td>2840369</td>
<td>STY2963</td>
</tr>
<tr>
<td>251</td>
<td>2848073</td>
<td>2848073</td>
<td>STY2971</td>
</tr>
<tr>
<td>251</td>
<td>2854109</td>
<td>2854109</td>
<td>STY2996</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Start Position</td>
<td>End Position</td>
<td>Strand</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>--------------</td>
<td>--------</td>
</tr>
<tr>
<td>2915322</td>
<td>2915322</td>
<td>2915322</td>
<td>N</td>
</tr>
<tr>
<td>2915330</td>
<td>2915330</td>
<td>2915330</td>
<td>N</td>
</tr>
<tr>
<td>2915627</td>
<td>2915627</td>
<td>2915627</td>
<td>2</td>
</tr>
<tr>
<td>2918204</td>
<td>2918204</td>
<td>2918204</td>
<td>N</td>
</tr>
<tr>
<td>2936316</td>
<td>2936316</td>
<td>2936316</td>
<td>N</td>
</tr>
<tr>
<td>2955365</td>
<td>2955365</td>
<td>2955365</td>
<td>N</td>
</tr>
<tr>
<td>2992175</td>
<td>2992175</td>
<td>2992175</td>
<td>N</td>
</tr>
<tr>
<td>2994859</td>
<td>2994859</td>
<td>2994859</td>
<td>N</td>
</tr>
<tr>
<td>3051986</td>
<td>3051986</td>
<td>3051986</td>
<td>N</td>
</tr>
<tr>
<td>3053028</td>
<td>3053028</td>
<td>3053028</td>
<td>N</td>
</tr>
<tr>
<td>3109192</td>
<td>3109192</td>
<td>3109192</td>
<td>N</td>
</tr>
<tr>
<td>3117804</td>
<td>3117804</td>
<td>3117804</td>
<td>N</td>
</tr>
<tr>
<td>3128447</td>
<td>3128447</td>
<td>3128447</td>
<td>N</td>
</tr>
<tr>
<td>3239753</td>
<td>3239753</td>
<td>3239753</td>
<td>N</td>
</tr>
<tr>
<td>3311882</td>
<td>3311882</td>
<td>3311882</td>
<td>N</td>
</tr>
<tr>
<td>3323883</td>
<td>3323883</td>
<td>3323883</td>
<td>N</td>
</tr>
<tr>
<td>3326564</td>
<td>3326564</td>
<td>3326564</td>
<td>N</td>
</tr>
<tr>
<td>3329720</td>
<td>3329720</td>
<td>3329720</td>
<td>N</td>
</tr>
<tr>
<td>3364345</td>
<td>3364345</td>
<td>3364345</td>
<td>N</td>
</tr>
<tr>
<td>3370871</td>
<td>3370871</td>
<td>3370871</td>
<td>N</td>
</tr>
<tr>
<td>3371437</td>
<td>3371437</td>
<td>3371437</td>
<td>N</td>
</tr>
<tr>
<td>3378518</td>
<td>3378518</td>
<td>3378518</td>
<td>N</td>
</tr>
<tr>
<td>3379980</td>
<td>3379980</td>
<td>3379980</td>
<td>N</td>
</tr>
<tr>
<td>Accession</td>
<td>Description</td>
<td>Function</td>
<td>Location</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>3437215</td>
<td>putative deoxyribonuclease</td>
<td>3.B.2 Degradation of DNA</td>
<td>1</td>
</tr>
<tr>
<td>3449225</td>
<td>5'-methyltetrahydropteroyl/triglutamate methyltransferase</td>
<td>1.D.2 Aspartate Family</td>
<td>1</td>
</tr>
<tr>
<td>3552039</td>
<td>possible LynR-family transcriptional regulatory protein</td>
<td>2 Broad regulatory function</td>
<td>1</td>
</tr>
<tr>
<td>3602082</td>
<td>vitamin B12 receptor protein</td>
<td>3.C Cell envelope</td>
<td>1</td>
</tr>
<tr>
<td>3652101</td>
<td>putative glycerol metabolic protein</td>
<td>1.C Central intermediary metabolism</td>
<td>1</td>
</tr>
<tr>
<td>3669615</td>
<td>putative membrane protein</td>
<td>3.C Cell envelope</td>
<td>1</td>
</tr>
<tr>
<td>3680371</td>
<td>possible membrane transport protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
</tr>
<tr>
<td>3688431</td>
<td>L-rhamnose isomerase</td>
<td>1.A.1 Degradation of carbohydrate</td>
<td>1</td>
</tr>
<tr>
<td>3716547</td>
<td>putative glycosyl hydrolase</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>3718028</td>
<td>putative membrane permease</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
</tr>
<tr>
<td>3754576</td>
<td>D-ribose-binding periplasmic protein</td>
<td>1.A.1 Degradation of carbohydrate</td>
<td>1</td>
</tr>
<tr>
<td>3795418</td>
<td>putative membrane transport protein</td>
<td>4.A Transport/binding proteins</td>
<td>1</td>
</tr>
<tr>
<td>3808610</td>
<td>recF protein</td>
<td>3.A.7 DNA - replication, repair, restriction</td>
<td>1</td>
</tr>
<tr>
<td>3820902</td>
<td>Two-component sensor protein histidine protein kinase.</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
</tr>
<tr>
<td>3826057</td>
<td>trimethylamine/N-oxide reductase precursor</td>
<td>1.B.7.b Anaerobic Respiration</td>
<td>1</td>
</tr>
<tr>
<td>3828882</td>
<td>probable cytochrome c peroxidase</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>3903736</td>
<td>putative glycosyl hydrolase</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>3926993</td>
<td>formamidopyrimidine-DNA glycosylase</td>
<td>3.A.7 DNA - replication, repair, restriction</td>
<td>1</td>
</tr>
<tr>
<td>3931488</td>
<td>lipopolysaccharide core biosynthesis protein</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>1</td>
</tr>
<tr>
<td>3943300</td>
<td>ADP-L-Glycero-D-mannoheptose-6-spherase</td>
<td>3.C.2 Surface polysaccharides &amp; antigens</td>
<td>1</td>
</tr>
<tr>
<td>3954677</td>
<td>serine acetyltransferase</td>
<td>1.D.3 Serine Family</td>
<td>1</td>
</tr>
<tr>
<td>ID</td>
<td>Start</td>
<td>End</td>
<td>Chromosome</td>
</tr>
<tr>
<td>-----------</td>
<td>-------</td>
<td>------</td>
<td>------------</td>
</tr>
<tr>
<td>4001985</td>
<td>4001985</td>
<td>4001985</td>
<td>NCT</td>
</tr>
<tr>
<td>4038663</td>
<td>4038663</td>
<td>4038663</td>
<td>CTT</td>
</tr>
<tr>
<td>4042168</td>
<td>4042168</td>
<td>4042168</td>
<td>NTG</td>
</tr>
<tr>
<td>4051472</td>
<td>4051472</td>
<td>4051472</td>
<td>NGA</td>
</tr>
<tr>
<td>4117689</td>
<td>4117689</td>
<td>4117689</td>
<td>NCT</td>
</tr>
<tr>
<td>4119751</td>
<td>4119751</td>
<td>4119751</td>
<td>NCT</td>
</tr>
<tr>
<td>4124948</td>
<td>4124948</td>
<td>4124948</td>
<td>CCA</td>
</tr>
<tr>
<td>4186093</td>
<td>4186093</td>
<td>4186093</td>
<td>NGC</td>
</tr>
<tr>
<td>4225021</td>
<td>4225021</td>
<td>4225021</td>
<td>NGA</td>
</tr>
<tr>
<td>4299207</td>
<td>4299207</td>
<td>4299207</td>
<td>NTG</td>
</tr>
<tr>
<td>4304442</td>
<td>4304442</td>
<td>4304442</td>
<td>NGA</td>
</tr>
<tr>
<td>4308440</td>
<td>4308440</td>
<td>4308440</td>
<td>GAG</td>
</tr>
<tr>
<td>4331402</td>
<td>4331402</td>
<td>4331402</td>
<td>NGA</td>
</tr>
<tr>
<td>4356323</td>
<td>4356323</td>
<td>4356323</td>
<td>NAG</td>
</tr>
<tr>
<td>4359035</td>
<td>4359035</td>
<td>4359035</td>
<td>NCT</td>
</tr>
<tr>
<td>4363103</td>
<td>4363103</td>
<td>4363103</td>
<td>NGA</td>
</tr>
<tr>
<td>4369734</td>
<td>4369734</td>
<td>4369734</td>
<td>NCT</td>
</tr>
<tr>
<td>4412203</td>
<td>4412203</td>
<td>4412203</td>
<td>NCT</td>
</tr>
<tr>
<td>4416872</td>
<td>4416872</td>
<td>4416872</td>
<td>NCT</td>
</tr>
<tr>
<td>4428141</td>
<td>4428141</td>
<td>4428141</td>
<td>NCT</td>
</tr>
<tr>
<td>4442998</td>
<td>4442998</td>
<td>4442998</td>
<td>GAG</td>
</tr>
<tr>
<td>4516890</td>
<td>4516890</td>
<td>4516890</td>
<td>NCT</td>
</tr>
<tr>
<td>4518117</td>
<td>4518117</td>
<td>4518117</td>
<td>NGA</td>
</tr>
<tr>
<td>4519552</td>
<td>4519552</td>
<td>4519552</td>
<td>NCT</td>
</tr>
<tr>
<td>Gene ID</td>
<td>Description</td>
<td>Accession</td>
<td>Length</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------</td>
<td>--------</td>
</tr>
<tr>
<td>VTI</td>
<td>Vi polysaccharide biosynthesis protein</td>
<td>VTIB operon</td>
<td>1</td>
</tr>
<tr>
<td>VTIB</td>
<td>Vi polysaccharide biosynthesis protein, UDP-glucose/GDP-mannose dehydrogenase</td>
<td>VTIB operon</td>
<td>1</td>
</tr>
<tr>
<td>STY4665</td>
<td>hypothetical protein</td>
<td>SPI-7</td>
<td>1</td>
</tr>
<tr>
<td>STY4666</td>
<td>probable phage integrase</td>
<td>SPI-7</td>
<td>1</td>
</tr>
<tr>
<td>fnrA</td>
<td>fumarate reductase, flavoprotein subunit</td>
<td>SPI-7</td>
<td>1</td>
</tr>
<tr>
<td>purA</td>
<td>adenyllysaccinate synthetase</td>
<td>1.F.1 Purine ribonucleotide biosynthesis</td>
<td>1</td>
</tr>
<tr>
<td>yijiC</td>
<td>conserved hypothetical protein</td>
<td>5.H.a Hypothetical protein</td>
<td>1</td>
</tr>
<tr>
<td>STY4751</td>
<td>putative membrane protein</td>
<td>3.C.1 Membranes lipoprotein</td>
<td>1</td>
</tr>
<tr>
<td>STY4785</td>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
<tr>
<td>trpS2</td>
<td>probable tryptophanyl-tRNA synthetase</td>
<td>3.A.5 Amino acyl tRNA synthesis; tRNA modification</td>
<td>1</td>
</tr>
<tr>
<td>mrr</td>
<td>mrr restriction system protein</td>
<td>3.A.7 DNA - replication, repair, restriction/modification</td>
<td>1</td>
</tr>
<tr>
<td>STY4912</td>
<td>hypothetical protein</td>
<td>5.I Unknown</td>
<td>1</td>
</tr>
</tbody>
</table>

The table above lists genes and their functions, along with their respective accession numbers and lengths.
Appendix G
Papers published on aspects of this thesis

A novel ciprofloxacin-resistant subclade of H58 Salmonella Typhi is associated with fluoroquinolone treatment failure

Duy Pham Thanh¹, Abhilasha Karkey², Sabina Dongol², Nhan Ho Thi¹, Corinne N Thompson¹,³, Maia A Rabaa¹,³, Amit Arjyal³, Fenway E Holt², Vanessa Wong², Nga Tran Vu Thieu¹, Phat Voong Vinh¹, Tuyen Ha Thanh¹, Ashish Pradhan², Saroj Kumar Shrestha³, Damodar Gajurel⁴, Derek Pickard⁵, Christopher M Parry⁶,⁷, Gordon Dougan⁸, Marcel Wolbers¹,³, Christiane Dolecek¹,³,⁸, Guy E Thwaites⁹,¹⁰, Buddha Basnyat², Stephen Baker¹,³,⁸,¹⁴

¹The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam; ²Oxford University Clinical Research Unit, Patan Academy of Health Sciences, Kathmandu, Nepal; ³Centre for Tropical Medicine and Global Health, Oxford University, Oxford, United Kingdom; ⁴The London School of Hygiene and Tropical Medicine, London, United Kingdom; ⁵Department of Biotechnology and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, The University of Melbourne, Melbourne, Australia; ⁶The Wellcome Trust Sanger Institute, Cambridge, United Kingdom; ⁷Civil Services Hospital, Kathmandu, Nepal; ⁸School of Tropical Medicine and Global Health, Department of Clinical Research, Nagasaki University, Nagasaki, Japan; ⁹Department of Zoology, University of Oxford, Oxford, United Kingdom

Abstract The interplay between bacterial antimicrobial susceptibility, phylogenetics and patient outcome is poorly understood. During a typhoid clinical treatment trial in Nepal, we observed several treatment failures and isolated highly fluoroquinolone-resistant Salmonella Typhi (S. Typhi). Seventy-eight S. Typhi isolates were genome sequenced and clinical observations, treatment failures and fever clearance times (FCTs) were stratified by lineage. Most fluoroquinolone-resistant S. Typhi belonged to a specific H58 subclade. Treatment failure with S. Typhi-H58 was significantly less frequent with ceftriaxone (3/31; 9.7%) than cefotaxim (15/34; 44.1%; Hazard Ratio 0.19, p=0.002). Further, for cefotaxim-treated patients, those infected with fluoroquinolone-resistant organisms had significantly higher median FCTs (8.2 days) than those infected with susceptible (2.9 days) or intermediate-resistant organisms (4.01; p<0.001). H58 is the dominant S. Typhi clade internationally, but there are no data regarding disease outcome with this organism. We report an emergent new subclade of S. Typhi-H58 that is associated with fluoroquinolone treatment failure. Clinical trial registration: ISRCTN53306507.

Introduction Enteric (typhoid) fever, a systemic infection caused predominantly by the bacterium Salmonella enterica subsp. enterica serovar Typhi (S. Typhi), remains one of the principal bacterial causes of febrile disease in low-income countries (Parry et al., 2002). S. Typhi is a distinct, monophyletic lineage of S. enterica that is exquisitely adapted to cause disease only in humans (Koumagné et al., 2006), characterised by a non-specific fever with malaise and asymptomatic convalescent carriage.

Pham Thanh et al. eLife 2016;5:e14003. DOI: 10.7554/eLife.14003

1 of 13

256
**eLife digest** People who ingest a type of bacteria called Salmonella Typhi can develop the symptoms of typhoid fever. This disease is common in low-income settings in Asia and Africa, and causes a high rate of death in people who are not treated with antimicrobial drugs.

During a study in Nepal, Thanh et al. tried to evaluate which of two antimicrobials was better for treating typhoid fever. One of the drugs – called gatifloxacin – did not work in some of the patients. To understand why this treatment failed, Thanh et al. decoded the entire DNA sequences of all the Salmonella Typhi bacteria isolated during the study. Comparing this genetic data to the clinical data of the patients identified a new variant of Salmonella Typhi. These bacteria have a specific combination of genetic mutations that render them resistant to the family of drugs that gatifloxacin belongs to – the fluoroquinolones.

Patients infected with the variant bacteria and treated with gatifloxacin were highly likely to completely fail treatment and have longer-lasting fevers. On further investigation Thanh et al. found these organisms were likely recently introduced into Nepal from India.

Fluoroquinolones are amongst the most effective and common antimicrobials used to treat typhoid fever and other bacterial infections. However, the presence of bacteria that are resistant to these compounds in South Asia means that they should no longer be the first choice of drug to treat typhoid fever in this location.

DOI: 10.7554/eLife.14003.002

(Perry et al., 2002). There are an estimated 20–30 million new cases of enteric fever per year globally (Gramp and Mintz, 2010), with the majority occurring in Asia, but there is an increasingly recognised burden of disease across sub-Saharan Africa.

Antimicrobial resistance is a major global health challenge, and resistance against the most commonly used antimicrobials for treating enteric fever has evolved successively over the last 30 years. Ampicillin, chloramphenicol and trimethoprim-sulfamethoxazole were originally standard-of-care for enteric fever. However, multidrug resistance (MDR) against these agents began to emerge in the 1970s and 1980s (Olarte and Galindo, 1973; Walin et al., 2003). Consequently, third-generation cephalosporins and fluoroquinolones (FQs) became the most clinically reliable drugs for treating enteric fever (Kariuki et al., 2013), and were formally advocated by the World Health Organization (WHO) in 2003 (World Health Organization, 2003). S. Typhi isolates with acquired resistance against third-generation cephalosporins are rare (Hendriksen et al., 2015), but S. Typhi exhibiting reduced susceptibility to FQs, induced by sequential mutations in the gene encoding a target protein (gyrA), now dominate internationally (Emary et al., 2012; Kariuki et al., 2010). The global ascendency of S. Typhi strains with reduced susceptibility to FQs has been partly catalysed by the dissemination of a specific MDR lineage (H58) across Asia and Africa (Wong et al., 2015). These H58 strains are rapidly displacing other lineages, and strains with gyrA mutations may have a fitness advantage, even in the absence of antimicrobial exposure (Biever et al., 2013).

We have previously shown that protracted fever clearance times (FCTs) are associated with organisms with higher Minimum Inhibitory Concentrations (MIC) against FQs in enteric fever patients treated with ciprofloxacin and ofloxacin (Perry et al., 2011). However, whilst the clinical efficacy of the older FQs in enteric fever is contentious, we have shown that the fourth-generation FQ, gatifloxacin, has remained efficacious for uncomplicated disease, even in patients infected with S. Typhi strains with reduced ciprofloxacin susceptibility (MIC >0.125 μg/mL) (Pandit et al., 2007; Keirala et al., 2012; Arjyal et al., 2011).

During a recent randomised controlled trial (RCT) comparing ceftriaxone and gatifloxacin, conducted in Nepal, we observed an increased number of treatment failures associated with FQ-resistant ciprofloxacin MIC >0.125 μg/mL S. Typhi, promoting the data safety and monitoring board to stop the trial (Arjyal et al., 2016). Aiming to assess the molecular epidemiology of the infecting isolates and investigate how genotype may be related to treatment outcome, we performed whole genome sequencing (WGS) on the S. Typhi isolated during this trial, and after stratifying by genotype, we assessed clinical presentation and outcome.
Results

Salmonella Typhi whole genome sequencing

We performed WGS on the 78 available S. Typhi isolates from patients in both RCT treatment arms (gatifloxacin and ceftriaxone) (Supplementary file 1). The resulting phylogeny, which incorporated reference sequence CTT18, indicated that the majority of isolates (65/78; 83.3%) fell within the H58 lineage, while the remaining 13 (16.7%) represented eight different lineages (Figure 1). All but four of the H58 strains contained the common DNA gyrase (gyrA) mutation in codon 83 (S83F), which conveys reduced susceptibility to FQs (ciprofloxacin MIC, 0.125–0.5 μg/ml) (Perry et al., 2019). Nested within the S83F-H58 group, but separated from the rest of the group by a branch defined by 30 SNPs, was an H58 sublineage comprised of 12 isolates containing the S83F gyrA mutation, a mutation in gyra1 at codon 87 (D87N), and an additional mutation in the topoisomerase gene, parC (S80I) (H58 triple mutant). Notably, these H58 triple mutants shared high MICs against ciprofloxacin (≥24 μg/ml). Further, an additional two non-H58 RCT isolates with ciprofloxacin MIC ≥24 μg/ml had the S83F gyrA mutation, an alternative mutation at codon 87 (D87V), the S80I parC mutation, and an A364V mutation in parE (Figure 1, Supplementary file 1). Notably, none of the sequenced isolates harboured plasmid-mediated quinolone resistance genes (PMQR) or contained additional antimicrobial resistance genes within the well-described S. Typhi-associated IncH1 family of plasmids.

Clinical presentation of Salmonella Typhi infections

We stratified clinical data from the RCT by H58 status of the corresponding S. Typhi isolates (H58; N=65, non-H58; N=13) and compared baseline characteristics between these groups. We found no significant differences in demographics and no association between disease severity at presentation

Figure 1. The phylogenetic structure of 78 Nepali Salmonella Typhi isolated during a gatifloxacin versus ceftriaxone randomised controlled trial. Maximum likelihood phylogeny based on core-genome SNPs of 78 Salmonella Typhi RCT isolates with the corresponding metadata, including the presence of mutations (dark grey) in gyrA (S83F, D87V and D87N), parC (S80I) and parE (A364V) and susceptibility to ciprofloxacin (susceptible, light blue; intermediate, mid-blue and non-susceptible, dark blue) by Minimum Inhibitory Concentration (MIC). The reference strain CTT18 was used for context and highlighted by the black boxes. Red lines linking to metadata show isolates belonging to the Salmonella Typhi H58 lineage (with H58 triple mutants highlighted), other lineages (non-H58) are shown with black lines. The scale bar indicates the number of substitutions per variable site (see methods). Asterisks indicate ≥85% bootstrap support at nodes of interest.

DOI: 10.7554/eLife.14003.003
between those infected with an HS8 S. Typhi isolate or a non-HS8 isolate (Supplementary file 2A). Next, we compared the baseline characteristics of patients stratified by ciprofloxacin susceptibility (susceptible, intermediate and resistant), and found no differences in disease severity or demographics on presentation; the only exception being that FO-resistant S. Typhi were more frequently isolated from adults (Supplementary file 2B). A significantly lower proportion of HS8 S. Typhi (4/65; 6.2%) were susceptible to FQs compared to non-HS8 isolates (6/13; 46%) (p=0.001) (Table 1) and, overall, HS8 isolates had significantly higher (but not resistant) MICs against the majority of tested antimicrobials than non-HS8 isolates (Table 1).

### Treatment failure and fever clearance times

The primary endpoint of the RCT in which these data were generated was a composite for treatment failure (see method and previous publication) (Adiyal et al., 2016). Treatment failure with HS8 S. Typhi was significantly less common in the ceftriaxone group (3/31; 9.7%) than the ciprofloxacin group (15/34; 44.1%) (Hazard Ratio (HR) of time to failure 0.19, 95%CI 0.05–0.56, p=0.002) (Table 2). Conversely, there was no significant difference in treatment failure between those infected with non-HS8 isolates treated with ciprofloxacin (0/6; 0%) or ceftriaxone (2/7; 28.6%) (p=0.32). Similarly, time to fever clearance differed significantly between the two treatment groups in HS8 infections, with median FCTs of 5.03 days (Interquartile range (IQR): 3.18–7.21) in the ceftriaxone group and 3.07 days (IQR: 1.89–4.52) in the ceftriaxone group (p<0.0001). Again, this trend was not mirrored in the non-HS8 S. Typhi infections, with FCTs of 2.67 (IQR: 2.08–3.7) and 3.12 (IQR: 2.2–4.12) days for ciprofloxacin and ceftriaxone, respectively (p=0.61) (Table 3). Moreover, in the ceftriaxone arm, HS8 S. Typhi tended to be associated with a higher risk of treatment failure (p<0.06) and a longer fever clearance time (p<0.013) (Figure 2, Table 2 and Supplementary file 2C).

As we identified two non-HS8 isolates that were also FQ-resistant (Figure 1), we additionally stratified outcome for the ceftriaxone arm (N=40 patients) by FO susceptibility of the infecting organism. Those infected with FO-resistant S. Typhi failed ceftriaxoin treatment more frequently (8/10, 80%) than those infected with an intermediate-resistant organism (7/25, 28%) or a susceptible organism (0/5, 0%) (p=0.007) (Figure 2 and Table 2). Furthermore, in the ceftriaxone arm, those infected with FO-resistant organisms had significantly higher median FCTs than those infected with FO-sensitive organisms (10/31; 33.3%).

### Table 1. Comparison of antimicrobial susceptibility by Salmonella Typhi lineage.

<table>
<thead>
<tr>
<th>E test</th>
<th>Non-HS8 (N=13)</th>
<th>HS8 (N=5)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MICs (range)</td>
<td>MICs (range)</td>
<td></td>
</tr>
<tr>
<td>Amoxicillin</td>
<td>0.5 (0.38–1)</td>
<td>0.75 (0.38–1)</td>
<td>0.0342</td>
</tr>
<tr>
<td>Chlortetracyclin</td>
<td>1.3 (1.5–8)</td>
<td>1 (1.0–2)</td>
<td>0.0042</td>
</tr>
<tr>
<td>Ceftriaxone</td>
<td>0.06 (0.01–1)</td>
<td>0.19 (0.03–0.64)</td>
<td>0.0004</td>
</tr>
<tr>
<td>Ciprofloxacin</td>
<td>0.01 (0.01–1)</td>
<td>0.32 (0.01–1)</td>
<td>0.0016</td>
</tr>
<tr>
<td>Ciprofloxacin susceptibility group</td>
<td>0.0008*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Susceptible</td>
<td>6 (46.2%)</td>
<td>4 (6.2%)</td>
<td></td>
</tr>
<tr>
<td>- Intermediate</td>
<td>0.01 (23.1%)</td>
<td>48 (73.8%)</td>
<td></td>
</tr>
<tr>
<td>- Resistant</td>
<td>13 (100%)</td>
<td>13 (100%)</td>
<td></td>
</tr>
</tbody>
</table>

*Comparisons between Salmonella Typhi lineages for MICs and ciprofloxacin susceptibility groups were based on the Wilcoxon rank sum test and Fisher’s exact test, respectively.

MD: minimum inhibitory concentration, measured in μg/ml.

DOI: 10.7554/eLife.14003.004
Table 2. Summary of time to treatment failure by Salmonella Typhi lineage and ciprofloxacin susceptibility.

<table>
<thead>
<tr>
<th>Time to treatment failure</th>
<th>Ceftriaxone (events/ N)</th>
<th>Ceftriaxone (events/ N)</th>
<th>Hazard ratio of time to failure (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS8*</td>
<td>15/34</td>
<td>3/31</td>
<td>0.19 (0.05, 0.59); p = 0.002</td>
<td>0.000</td>
</tr>
<tr>
<td>- Non-HS8</td>
<td>0/6</td>
<td>2/7</td>
<td>3.87 (0.31, 53.24); p = 0.32</td>
<td></td>
</tr>
</tbody>
</table>
| Ciprofloxacin susceptibility group
  | Susceptible             | 0/5                     | 1/5                     | 2.40 (0.13, 35.21); p = 0.57                   | 0.08                       |
  | Intermediate             | 7/25                    | 2/27                    | 0.21 (0.05, 0.99); p = 0.049                    |                            |
  | Resistant                | 8/10                    | 2/6                     | 0.21 (0.05, 1.07); p = 0.032                    |                            |

*Likelihood ratio test p=0.06 and 0.40 for comparison of time to treatment failure between HS8 vs. non-HS8 groups in ceftriaxone arm only and in all patients, respectively.

Likelihood ratio test p=0.0037 for comparison of time to treatment failure between MIC groups in ceftriaxone arm only.

DOI: 10.7554/eLife.14003.005

S. Typhi with alternative FQ susceptibility profiles (median FCTs (days)); susceptible, 2.96 (IQR: 2.13–3.85), intermediate, 4.01 (IQR: 2.76–5.37) and resistant 8.2 (IQR: 5.99–10.5), respectively [p=0.0001] (Table 3 and Supplementary file 2D). Comparatively, the median FCT for those infected with an FQ-resistant organism but randomised to ceftriaxone was 3.83 days (IQR: 2.96–4.7) (p=0.0001 for the between-treatment comparison).

The emergence of fluoroquinolone-resistant Salmonella Typhi

To measure the pattern of emergence of FQ-resistant S. Typhi in Nepal, we compiled FQ susceptibility data from 837 organisms isolated during enteric fever RCTs conducted at Patan Hospital between 2005 and 2014 (Figure 3) (Pandit et al., 2007; Kaikala et al., 2013; Arjyal et al., 2011). MICs against FQs were generally higher for S. Paratyphi A than for S. Typhi. There was a significant temporal increase in S. Typhi MICs against both ciprofloxacin (p<0.0001) and ceftriaxone (p<0.0001), with a sharp increase from 2009. MICs against ceftriaxone in S. Paratyphi A also significantly increased with time (p=0.0001); however, MICs against ciprofloxacin showed only weak evidence of an upward trend over time (p=0.06).

We hypothesised that the HS8 triple mutants represented a contemporary importation into Nepal. To explore this, we compared the genomes of the 78 RCT S. Typhi isolates with those from 58 supplementary S. Typhi isolates from previous studies conducted between 2006 and 2013 in this setting (Figure 4; Supplementary file 1) (Wong et al., 2015). We found that the majority of the local HS8 isolates (84/121; 69.4%) were closely related; these strains represented an endemic Nepali HS8 clade containing a single SS3F gyrA mutation. Additionally, we identified a further five Nepali strains isolated in 2013 that belonged to the HS8 triple mutant group, and had an MIC ≥24 μg/ml against ciprofloxacin. Incorporating additional genome sequences from a recent international study of the HS8 lineage (Wong et al., 2015), we found that all the Nepali HS8 triple mutants were very closely related (5 SNPs; to nearest neighbour) to HS8 triple mutants isolated previously in neighbouring India between 2008 and 2012 (Figure 4).

Discussion

Our study shows that a new FQ-resistant subclade of HS8 S. Typhi has been introduced into Nepal and is associated with a lack of FQ efficacy. This subclade was associated with longer FCTs and treatment failure in patients treated with the FQ, ceftriaxone. For the first time, we can conclusively show how enteric fever patients respond to FQ treatment when infected with a specific subclade of HS8, thereby linking organism genotype with a treatment phenotype. Given the international significance of FQs for the treatment of enteric fever and other bacterial infections, our findings have major global health implications for the long-term use and efficacy of this group of antimicrobials.

Pham Thanh et al. eLife 2016;5:e14003. DOI: 10.7554/eLife.14003

S of 13

260
Our data suggest these FQ-resistant S. Typhi strains circulating in Nepal most likely descended from a single ancestor carrying the triple gyrA/parC mutant, such as that isolated in Nepal in 2011 (Koirala et al., 2012). This isolate was also associated with treatment failure, although this organism was not genome sequenced and was assumed to be an isolated case. More significantly, several very closely related strains were genome sequenced during an international study of H58 S. Typhi (Wong et al., 2016). These organisms had the same combination of triple FQ resistance mutations as those described here; our analysis shows they belong to the same subclade of H58. These strains had equivalently high MICs against ciprofloxacin and were isolated in India between 2006 and 2012. However, there were no associated patient outcome data for these strains and other reports from India have been limited. Our data implies that this lineage was introduced into Nepal from India or elsewhere in South Asia within the last 4-5 years and has subsequently entered in an endemic transmission cycle in Kathmandu. Given the large extent of human movement between India and Nepal, we propose this is the most likely route of introduction. However, there is also a small possibility
Table 3. Summary of fever clearance time by Salmonella Typhi lineages and ciprofloxacin susceptibility.

<table>
<thead>
<tr>
<th>Fever clearance time</th>
<th>Gatifloxacin median (IQR) days</th>
<th>Ceftriaxone median (IQR) days</th>
<th>Acceleration factor (95%CI); p value</th>
<th>Heterogeneity test (p value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HS4</td>
<td>5.03 (3.18, 7.21)</td>
<td>3.07 (1.90, 4.52)</td>
<td>1.59 (1.22, 2.09); p=0.0006</td>
<td>0.07</td>
</tr>
<tr>
<td>- H58</td>
<td>2.87 (2.08, 3.7)</td>
<td>3.12 (2.2, 4.12)</td>
<td>0.90 (0.59, 1.36); p=0.61</td>
<td>0.015</td>
</tr>
</tbody>
</table>

Ciprofloxacin susceptibility group
- Susceptible: 2.96 (2.13, 3.85)
- Intermediate: 4.01 (2.76, 5.32)
- Resistant: 8.2 (5.99, 10.5)

$^a$p=0.013 and p=0.029 for comparison of interval censored time to fever clearance between H58 vs. non-H58 groups in gatifloxacin arm only and in all patients, respectively. $^b$p=0.001 for comparison of interval censored time to fever clearance between MIC groups in gatifloxacin arm only.

DOI: 10.7554/eLife.14003.007

that multiple strains independently gained resistance against FCQs through the same selective pressure.

Figure 3. Minimum Inhibitory Concentrations of Nepali Salmonella Typhi and Salmonella Paratyphi against ciprofloxacin and gatifloxacin over ten years. Minimum Inhibitory Concentrations (μg/ml) for 568 Nepali Salmonella Typhi (Blue) and 269 Nepali Salmonella Paratyphi A (grey) against (A) ciprofloxacin and (B) gatifloxacin collected from four randomised controlled trials conducted between 2006–2014 at Patan Hospital in Kathmandu, Nepal (Pandit et al., 2009; Koirala et al., 2013; Aryal et al., 2011). The smoothed line derived from the generalized additive model showing a non-linear increase in Minimum Inhibitory Concentration over time, with shading representing the 95% confidence interval. Lower and upper horizontal lines represent the current CLSI cut-offs for susceptible/intermediate and intermediate/resistant, respectively (CLSI, 2012).

DOI: 10.7554/eLife.14003.008
Appropriate antimicrobial therapy is critical in the treatment of enteric fever, as effective drugs curtail symptoms and prevent life-threatening complications. Our data has substantial repercussions for enteric fever treatment, and we advocate that FQs should no longer be used for empirical enteric fever therapy on the Indian subcontinent, as we predict these strains are likely to be widespread and are associated with poor outcomes with FQ therapy. Notably, in the RCT from which these data were derived, we used the newer generation FQ, gatifloxacin, which binds to a different location on the DNA gyrase than the older FQs and is not as susceptible to the common resistance mutations (Lu et al., 1999). The isolates in this study were not generally resistant to gatifloxacin according to the current CLSI guidelines for Enterobacteriaceae (CLSI, 2012); we suggest that these guidelines be modified specifically for S. Typhi to reflect these new clinical data. We additionally propose that S. Typhi genotyping, mapping and susceptibility testing is performed routinely and rapidly in reference laboratories inside and outside of South Asia to monitor the international spread of these strains and ensure the provision of alternative efficacious therapies to returning travellers (Lee et al., 2013; Garcia-Fernández et al., 2015). In cases of infection with these FQ-resistant isolates, we suggest that ceftriaxone and azithromycin be used as alternatives, and do not currently recommend a return to the use of first-line drugs without contemporary data on treatment outcome. While none of the isolates in this study were MDR, we predict a rapid return of MDR strains if there is a hasty return to older first-line alternatives.

This study has limitations. First, the clinical data was collected from one study in a single location, thus limiting utility outside this setting. Second, the overall sample size (and the gatifloxacin group subsampling) of those with culture-positive S. Typhi-associated enteric fever was relatively small, and the analysis presented here was performed in a post hoc manner. Notwithstanding these limitations, we were able to show a highly significant association between disease outcome and susceptibility.

DOI: 10.7554/eLife.14003
profile of the infecting organism. Further, by using WGS, we were able to pinpoint causative mutations, identify the subclade responsible for treatment failure and relate these strains to other isolates circulating outside Nepal in other parts of South Asia. The methodologies presented here, in which clinical outcome data are combined with genome sequences and antimicrobial susceptibility data, should become the gold standard for informing empiric treatment for all invasive bacterial infections and understanding the role of bacterial genotype and resistance profile on disease outcome for other bacterial infections. No other combination of methodologies would provide the granularity of data required to understand the epidemiology and clinical impact of this emergent strain in detail.

In conclusion, our data, for the first time, show a significant association between S. Typhi genotype, antimicrobial susceptibility and disease outcome for those treated with ceftriaxone in a cohort of Nepali enteric fever patients. A PO-resistant variant of Typhi HS8 has emerged in Nepal and is associated with the clinical failure of FQs. Our data suggest these isolates are likely widespread in the subcontinent and FQs should not be recommended for empirical enteric fever therapy in this setting.

Materials and methods

Study design and setting

The RCT from which the organisms and corresponding clinical data originated for these analyses was conducted at Patan Hospital and the Civil Hospital in the Lalitpur area of Kathmandu, Nepal, between 2011 and 2014, as described previously (Arjyal et al., 2016). The trial was registered at www.clinicaltrials.gov (ISRCTN63006567). Briefly, patients were randomly assigned to seven days of treatment with either oral ceftriaxone (400 mg tablets, Square Pharmaceuticals Limited, Bangladesh) at a dose of 10 mg/kg once daily or intravenous ceftriaxone (Powecef, 1000 mg injection vial, Wockhardt Ltd, India), injected over 10 min at a dose of 60 mg/kg up to a maximum of two grams (aged 2 to 13 years) or two grams (≥14 years) once daily.

A detailed description of the RCT from which these data were generated has been previously published (Arjyal et al., 2016). The primary endpoint was a composite of treatment failure, defined as the occurrence of at least one of the following events: fever clearance time (FCT) (time from the first dose of a study drug until the temperature dropped to ≤37.5°C and remained there for at least two days) more than seven days post-treatment initiation; requirement for rescue treatment as judged by the treating physician; blood culture positivity for S. Typhi or S. Paratyphi on day eight of treatment (microbiological failure); culture-confirmed or syndromic enteric fever relapse within 28 days of initiation of treatment; and the development of any enteric fever-related complication (e.g., clinically significant bleeding, fall in the Glasgow Coma Score, perforation of the gastrointestinal tract and hospital admission) within 28 days after the initiation of treatment. Time to treatment failure was defined as the time from the first dose of treatment until the date of the earliest failure event. FCTs were calculated electronically using twice-daily recorded temperatures and treated as interval-censored outcomes. Patients without fever clearance or relapse, respectively, were censored at the time of their last follow-up visit (additional details regarding study procedures can be found in Arjyal et al., 2016 (Arjyal et al., 2016)).

Blood (3 ml if aged <14 years; 8 ml if aged ≥14 years) was taken from all patients for bacterial culture on enrolment. Adult blood samples were inoculated into media containing tryptone soya broth and sodium polyanethol sulphonate, up to a total volume of 50 mL. Bectec Peds Plus culture bottles (Becton Dickinson, New Jersey, USA) were used for paediatric blood samples. Culture results were reported for up to seven days, positive bottles were subcultured onto blood, chocolate and MacConkey agar and presumptive Salmonella colonies were identified using standard biochemical tests and serotype-specific antisera (Murex Biotech, Dartford, England). Antimicrobial susceptibility testing was performed by the modified Bauer-Kirby disc diffusion method with zone size interpretation based on CLSI guidelines (CLSI, 2012). E-tests were used to determine MICs, following the manufacturer's recommendations (bioMérieux, France). Ciprofloxacin MICs were used to categorise S. Typhi isolates as susceptible (≤0.06 μg/mL), intermediate (0.12–0.5 μg/mL) and resistant (≥1 μg/mL) following CLSI guidelines (CLSI, 2012).

Pham Thanh et al. eLife 2016;5:e14003. DOI: 10.7554/eLife.14003

9 of 13

264
Whole genome sequencing and analysis
Genomic DNA from Nepali S. Typhi organisms originating from this RCT (78 isolates) was extracted using the Wizard Genomic DNA Extraction Kit (Promega, Wisconsin, USA) (Supplementary file 1) (Karkey et al., 2013). Two µg of genomic DNA was subjected to WGS on an Illumina MiSeq platform, following the manufacturer’s recommendations to generate 250bp/100bp paired-end reads. All reads were mapped to the reference sequence of S. Typhi CT18 (accession no: AL315582) using SMALT (version 0.7.4). Candidate single nucleotide polymorphisms (SNPs) were called against the reference sequence using SAMtools (Li et al., 2009) and filtered with a minimal phred quality of 30 and a quality cut-off of 0.75. The allele at each locus in each isolate was determined by reference to the consensus base in that genome, using samtools mpileup and removing low confidence alleles with consensus base quality ≤20, read depth ≤5 or a heterozygous base call. SNPs called in phage regions, repetitive sequences or recombinant regions were excluded. (Wong et al., 2015) resulting in a final set of 1,607 chromosomal SNPs. Strains belonging to haplotype HS8 were defined by the SNP gpfA-C1047T (position 2348902 in S. Typhi CT18, Bip33) (Emary et al., 2012; Holt et al., 2008; Parkhill et al., 2001).
A maximum likelihood (ML) phylogeny was estimated using a 1440 SNP alignment of the 78 RCT isolates in RAxML (version 7.8.6) with the generalized time-reversible substitution model (GTR) and a gamma distribution, with support for the phylogeny assessed via 1000 bootstrap replicates. The alignment was then compared to a global S. Typhi sequence database, with a particular focus on identifying sequences with a mutational profile suggestive of shared ancestry with a divergent HS8 clade identified in the previous phylogeny. A secondary ML phylogenetic tree was then inferred from the SNP alignment of the 136 Nepali Typhi along with 19 recently described Typhi HS8 with the aforementioned mutational profile, using the same parameters as above (1642 SNPs; Supplementary file 1) (Wong et al., 2015). Raw sequence data are available in the European Nucleotide Archive (ENA) (Supplementary file 1).

Statistical analysis
Comparison of baseline characteristics within patient groups, stratified by the HS8 status or susceptibility category of their corresponding S. Typhi isolates was performed using the Kruskal Wallis test for continuous variables and Fisher’s exact test for categorical variables. Time to treatment failure was analysed using Firth’s penalized maximum likelihood bias reduction method for Cox regression as a solution for the non-convergence of likelihood function in the case of zero event counts in subgroups (Firth, 1993). For comparisons between treatment arms, HS8 status, or ciprofloxacin susceptibility group, the model included treatment arm, HS8 status, or susceptibility group as a single covariate. Confidence intervals (CI) and p-values were calculated by profile-penalized likelihood. FCT was analysed as an interval-censored outcome, i.e. as the time interval from the last febrile temperature assessment until the first afebrile assessment, using parametric Weibull accelerated failure time models (Kalbfleisch and Prentice, 2002). Median and inter-quartile range (IQR) FCT calculations for subgroups were based on models for each subgroup separately. Acceleration factors were based on models that included treatment arm as the only covariate. The non-parametric maximum likelihood estimator (NPML) was used to visualize the distribution of FCT between groups. Heterogeneity between subgroups was tested with models that included an interaction between treatment arm and the sub-grouping variable. To study the emergence of FQ resistance, data from previous enteric fever trials from 2003–2014 (Pandit et al., 2007; Kolasa et al., 2013; Arjyal et al., 2011) was pooled and generalized additive models (GAM) were used to examine potential non-linear trends of ciprofloxacin and gatifloxacin MICs over time. All analyses were performed using R software version 3.2.2 (Team, 2012).

Acknowledgements
This project was funded by the Wellcome Trust of Great Britain (106158/2/14/Z), SB is a Sir Henry Dale Fellow, jointly funded by the Wellcome Trust and the Royal Society (100387/Z/12/Z). KEH is supported by fellowship #1061409 from the NHMRC of Australia. CD was funded by the Li Ka Shing Foundation Global Health Programme at the University of Oxford. DTP and AK are funded as leadership fellows through the Oak Foundation. The funders had no role in study design, data collection
and analysis, decision to publish, or preparation of the manuscript. We wish to acknowledge Mangal Rawal, Sumit Munarkamri, Bilbek Karki, Radheshyam KC, Sudeep Dhioj Thapa, Rabin Gautam, Priyanka Tiwari, Manisha Risal, Surendra Shrestha, Balmukunda Neupane, Nabraj Regmi, Krishna Prajapati, Bimal Thapa and the trial monitors Nguyen Thi Phuong Dung and Nguyen Thi Thanh Thuy for their assistance in conducting this trial.

Additional information

<table>
<thead>
<tr>
<th>Funding</th>
<th>Grant reference number</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wellcome Trust</td>
<td>100007/Z/12/Z</td>
<td>Stephen Baker</td>
</tr>
<tr>
<td>Royal Society</td>
<td>100007/Z/12/Z</td>
<td>Stephen Baker</td>
</tr>
<tr>
<td>National Health and Medical Research Council</td>
<td>1061409</td>
<td>Kathryn E Holt</td>
</tr>
<tr>
<td>Li Ka Shing Foundation</td>
<td></td>
<td>Christane Delceck</td>
</tr>
<tr>
<td>Wellcome Trust</td>
<td>106150/Z/14/Z</td>
<td>Kathryn E Holt, Gordon Douglas, Christane Delceck, Buddhi Basnyat, Stephen Baker</td>
</tr>
<tr>
<td>The Oak Foundation</td>
<td>OCAY-15-547</td>
<td>Duy Pham Thanh, Abhijit Kashyap, Stephen Baker</td>
</tr>
</tbody>
</table>

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

DPT, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article; AK, SO, VW, NYT, PW, TH, DP, Acquisition of data, Contributed unpublished essential data or reagents; NHT, CNT, MAR, KEH, Analysis and interpretation of data, Drafting or revising the article; AA, BB, Conception and design, Acquisition of data; AP, SKS, DG, Recruited patients into study and collected clinical data, Acquisition of data; CMP, GD, GET, Conception and design, Drafting or revising the article; MW, Analysis and interpretation of data, Contributed unpublished essential data or reagents; CD, Conception and design, Acquisition of data, Contributed unpublished essential data or reagents; SB, Conception and design, Acquisition of data, Analysis and interpretation of data, Drafting or revising the article.

Author ORCID

Maia A Rabaa http://orcid.org/0000-0003-0529-2282

Ethics

Clinical trial Registry: ISRCTN. Registration ID: ISRCTN63066567.

Human subjects: This study was performed following the principles of the declaration of Helsinki. Written informed consent to participate in all studies from Nepal contributing data for this analysis was required from all patients. For those aged <18 years, written informed consent was obtained from a parent or an adult guardian. The protocol was reviewed and approved by the Ethics Committee of the Nepal Health Research Council (NHRC) and the Oxford Tropical Research Ethics Committee (OxTREC) UK.

Additional files

Supplementary files

• Supplementary file 1. Table of Salmonella Typhi isolates and their corresponding sequencing metadata used in this study.
Major datasets
The following dataset was generated:

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Dataset title</th>
<th>Dataset URL</th>
<th>Database, license, and accessibility information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duong Pham Thanh, Stephen Baker, Kathryn E. Holt, Gordon Dougan, Vanessa Wong</td>
<td>2015</td>
<td>Sequence data</td>
<td><a href="http://www.ebi.ac.uk/ena/data/view/PRJEB10959">http://www.ebi.ac.uk/ena/data/view/PRJEB10959</a></td>
<td>PRJEB10959</td>
</tr>
</tbody>
</table>

References


Pham Thanh et al. eLife 2016;5:e14003. DOI: 10.7554/eLife.14003

Epidemiology and global health | Microbiology and infectious disease

268
The Molecular and Spatial Epidemiology of Typhoid Fever in Rural Cambodia

Duy Pham Thanh 1, Corinne N. Thompson 1,2,4, Mala A. Rabaa 1,2,3, Seong Sone 4, Sun Sophary 5, Varun Kumar 6, Catherine Moore 4, Nga Tran Vu Thieu 1, Lalith Wijedenu 1, Kathryn E. Holt 1, Vanessa Wong 6, Derek Pickard 6, Guy E. Thwaites 1,4, Nicholas Day 2,4, Gordon Dougan 4,4, Paul Turner 4,4, Christopher M. Parry 4,4,4,4, Stephen Baker 1,2,4,4,4

1 The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam, 2 Centre for Tropical Medicine, Oxford University, Oxford, United Kingdom, 3 The London School of Hygiene and Tropical Medicine, London, United Kingdom, 4 Cambridge-Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia, 5 Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, Australia, 6 The Wellcome Trust Sanger Institute, Hinxton, Cambridge, United Kingdom, 7 The Department of Medicine, University of Cambridge, Cambridge, United Kingdom, 8 School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan

Abstract

Typhoid fever, caused by the bacterium Salmonella Typhi, is an endemic cause of febrile disease in Cambodia. The aim of this study was to better understand the epidemiology of pediatric typhoid fever in Cambodia. We assessed routine blood culture data from Angkor Hospital for Children (AHC) in Siem Reap province between 2007 and 2014, and performed whole genome sequencing (WGS) on the isolated bacteria to characterize the S. Typhi population. The resulting phylogenetic information was combined with conventional epidemiological approaches to investigate the spatiotemporal distribution of S. Typhi and population-level risk factors for reported disease. During the study period, there were 262 cases of typhoid within a 10 km radius of AHC, with a median patient age of 8.2 years (IQR: 5.1–11.5 years). The majority of infections occurred during the rainy season, and commune incidences as high as 11.36/1,000 in children aged <15 years were observed over the study period. A population-based risk factor analysis found that access to water within households and increasing distance from Tonle Sap Lake were protective. Spatial mapping and WGS provided additional resolution for these findings, and confirmed that proximity to the lake was associated with discrete spatiotemporal disease clusters. We confirmed the dominance of MDR-H69 S. Typhi in this population, and found substantial evidence of diversification (at least seven sublineages) within this single lineage. We conclude that there is a substantial burden of pediatric typhoid fever in rural communes in Cambodia. Our data provide a platform for additional population-based typhoid fever studies in this location, and suggest that this would be a suitable setting in which to introduce a school-based vaccination programme with Vi conjugate vaccines.
Author Summary

Typhoid fever is an infectious disease caused by the bacterium Salmonella Typhi. The disease is generally restricted to those living in low-income settings with poor sanitation. Typhoid fever is a common cause of fever requiring hospital treatment in Cambodia, but limited data is available on the epidemiology of the disease. To better understand typhoid fever in Cambodia, we accessed routine hospital data for typhoid fever from a single healthcare facility treating sick children in Siem Reap in the central Cambodia between 2007 and 2014. We mapped the location of these cases and examined population-based risk factors for reported disease. Additionally, we decoded the genomes of the S. Typhi isolated from children attending the hospital to understand how the organism has evolved and spread throughout the population. We found a large burden of typhoid fever in children in this largely rural setting in central Cambodia. We also found that disease was associated with the rainy season and that living close to Tonle Sap Lake increased the risk of disease. The genomes of the sequenced bacteria showed that a diverse range of strains were circulating during the study, and allowed us to identify signatures of location- and time-specific outbreaks. Our work provides baseline data for additional typhoid fever studies in the population living in this location and findings suggest that rural Cambodia would be a suitable setting in which to introduce a school-based vaccination program with new typhoid vaccines.

Introduction

The bacterium Salmonella enterica serovar Typhi (S. Typhi) is the cause of the human infection typhoid fever, a systemic disease predominantly diagnosed in children and young adults in low-income settings [1]. S. Typhi is primarily contracted via ingestion of food or water contaminated with human feces from patients excreting the organism, and typhoid fever remains a major public health issue in areas with poor sanitation and limited access to safe water [2]. The control of typhoid fever is largely dependent on improving the availability of clean water, hygienic food preparation and access to adequate sanitation, but such interventions are substantial challenges in many locations where typhoid remains endemic [3]. As a result, active case detection and appropriate antimicrobial therapy are currently the principal methods for controlling this disease in endemic locations. The lack of rapid and reliable diagnostics and the emergence of antimicrobial resistance (AMR) reduce the effectiveness of these strategies [4,5].

The World Health Organization (WHO) currently recommends the use of licensed typhoid vaccines in areas where the burden of typhoid fever is high and AMR organisms are prevalent, though limited programmatic use of vaccines has occurred in endemic countries [6]. Identifying areas that have a burden of typhoid fever warranting immunization programs can be challenging. Much of the current focus for vaccine implementation is on highly populated urban slums with poor infrastructure [7-9]. However, it is unclear if the epidemiology of typhoid is similar between urban and rural regions in the developing world. Understanding the dominant modes of transmission and epidemiological risk factors for typhoid fever in both urban and rural endemic regions is vital for controlling and preventing the disease [3].

New molecular tools now permit an unprecedented insight into how S. Typhi may be circulating locally and internationally [10,11,12]. Comparing the composition of phylogenetically informative Single Nucleotide Polymorphisms (SNPs) across the S. Typhi genome allows the subtyping of populations of S. Typhi and inference of evolutionary relationships between isolates [12,13]. SNP-based typing methods, and now whole genome sequencing (WGS), have
been successfully used to study the molecular epidemiology of typhoid in different settings, revealing the importance of environmental transmission and the diversity of commonly co-circulating haplotypes of S. Typhi within localized human populations [13–16]. Such molecular approaches play an important role in identifying dominant transmission pathways and can also capture both the emergence of AMR and the dynamics of the bacterial population. We now know that the current population of S. Typhi has been driven by a clonal expansion and international dispersal of a specific haplotype (H58) in Asia and Africa. This H58 haplotype now dominates internationally and is associated with a multidrug-resistant (MDR) phenotype (non-susceptibility to ampicillin, chloramphenicol and trimethoprim-sulphamethoxazole) and reduced susceptibility to fluoroquinolones [17].

Typhoid fever is endemic in Cambodia, although only limited data regarding the morbidity, mortality and risk factors for the disease are available in published literature. A community-based study conducted near Phnom Penh (the capital city) between 2006 and 2009 reported routine isolation of MDR S. Typhi from the blood of febrile patients [17]. Additionally, a hospital-based cross-sectional study of pediatric bloodstream infections in a children’s hospital in Siem Reap (in northwest Cambodia, near the World Heritage Site of Angkor Wat) between 2007 and 2011 found that S. Typhi was the most commonly isolated pathogen in this setting and confirmed the presence and dominance of H58 S. Typhi (98/102; 96% of S. Typhi isolates) and exhibited reduced susceptibility to ciprofloxacin [18,19]. In the present study, we aimed to utilize the precision of WGS to characterize the H58 S. Typhi population in Siem Reap, Cambodia. Further, we combined the resulting phylogenetic information with additional epidemiological approaches to investigate the spatiotemporal distribution of S. Typhi and population-level risk factors for typhoid fever infection in this location.

Methods

Ethics statement

The study involved characterization of stored bacterial isolates cultured from specimens taken for routine clinical care. Therefore, it was not possible to obtain consent from the patient or their parent/guardian for participation in this retrospective study, but all patient data was anonymized. The study protocol was reviewed and approved by both the Angkor Hospital for Children Institutional Review Board (AHC IRB, reference 423/13) and the Oxford Tropical Research Ethics Committee (OxTREC, reference 512–13).

Study site and setting

This study was conducted at Angkor Hospital for Children (AHC) in Siem Reap City in Cambodia between January 2007 and December 2014. AHC is one of two pediatric hospitals in Siem Reap City and has approximately 125,000 attendances and 4,000 admissions per year. The patients attending AHC are <16 years of age and come from a wide geographical radius and attend the hospital for various conditions. The majority of patients reside in the province of Siem Reap, which is located in northwest Cambodia and is bordered in the south by the Tonle Sap Lake, the largest freshwater lake in Southeast Asia. According to available census data, the province had a population of 896,443 people living in an area of 10,299 km² in 2008; the province is subdivided administratively into 12 districts, 100 communes (which are within districts) and 907 villages. [20]. Cambodia has a tropical climate with a dry and wet season each year. During the wet season (April–October) the area of the Tonle Sap Lake can expand dramatically, increasing from 3,500 km² up to approximately 14,500 km², with the depth increasing from 0.5 m up to 6–9 m [21].
Definition of the case population and the control population

Case and control populations were identified from the electronic hospital and laboratory information system of AHC. For the purposes of this study, the case population was defined as the population of hospital inpatients from whom S. Typhi was isolated from a blood culture. The control population was defined as the patient population admitted to AHC who did not have typhoid fever based on the recorded discharge diagnosis (International Classification of Disease (ICD)-10 code). Patients with a discharge diagnosis of typhoid fever but without blood culture confirmation (n = 410) were not included in the risk factor analysis. Additionally, for the mapping and population risk factor analyses, cases that lived outside of a 100km radius from AHC were excluded. Data on age, sex, home location (commune level), admission and discharge dates for cases and controls were extracted from the electronic hospital information system. If a case or control was readmitted to the hospital with the same discharge diagnosis within a seven-day period, only the initial admission was included in the analysis.

Data sources

Commune-level census data were obtained from the Cambodian National Report on General Population Census of 2008 [20]. The extracted information included details regarding demographic indicators, age structure, literacy and education, housing and household characteristics, and access to toilet facilities and drinking water. Based on this report, a commune was classified as urban if the population density exceeded 200/km², less than half of men were employed in agriculture and the total population exceeded 2,000. Monthly average precipitation was collected from Siem Reap Weather Station and MROCS (Mekong River Commission Secretariat) [21]. Shuttle Radar Topography Mission (SRTM) elevation data were obtained from the CGIAR Consortium for Spatial Information (CGIAR-CSI) [22]. Shapefile layers containing 2008 commune-level population census data were accessed from Open Development Cambodia, an open-access data website providing data on Cambodia and its economic and social development (http://www.opendecvelopmentcambodia.net).

Typhoid diagnosis and bacterial identification

Routine diagnosis of typhoid fever was performed by blood culture. Blood (1–4 ml) was taken for bacterial culture from all patients with fever including those with a clinical suspicion of typhoid fever. Blood was inoculated into media containing tryptone soya broth and sodium polyanethol sulphonate, up to a total volume of 25ml. Blood culture bottles were incubated for up to seven days, with blind sub-cultures at 24 hours, 48 hours and 7 days or if the broth was cloudy. Positive bottles were subcultured onto sheep blood, chocolate and MacConkey agar and presumptive Salmonella colonies were identified using standard biochemical tests and serotype-specific antisera (Murex Biotech, Dartford, England). Antimicrobial susceptibility testing was performed by the modified Bauer-Kirby disc diffusion method with zone size interpretation based on CLSI guidelines [23]. Tests were used to determine MICS, following the manufacturer’s recommendations (bioMérieux, France). Ciprofloxacin MICS were used to categorise S. Typhi isolates as susceptible (<0.06 µg/mL), intermediate (0.12–0.5 µg/mL) and resistant (≥1 µg/mL) following CLSI guidelines [24].

Statistical analysis

Rates of hospitalized typhoid fever were calculated at the commune level using the population under the age of 15 years from 2008. Multivariable negative binomial regression was used to identify commune-level risk factors associated with the rate of cases per 1,000 population.
under the age of 15 years. Interaction between commune level factors was evaluated using the likelihood ratio test. Variables included in the evaluation of the final model included those with significant associations ($p<0.10$) in the univariate analysis and a priori sanitation and water source variables. Variables that did not add significantly to the fit of the final model (determined by the likelihood ratio test) were not included. All analyses were performed in STATA (v13, College Station, TX, USA) and plots were created in R v3.1.1 (R Foundation for Statistical Computing, Vienna, Austria, https://cran.r-project.org/) using ggplot2 [25].

Spatial-temporal clustering detection

Spatial-temporal clustering analysis was performed using Moran’s $I$ and SaTScan methodologies. First, Moran’s $I$ test was used to evaluate global autocorrelation amongst communes that reported at least one case ($n=78$) of typhoid fever in GeoDa software (v1.6.7, https://geospatialcenter.asu.edu/). This test statistic provides an evaluation of whether the rates across the area of interest are spatially random (Moran’s $I=0$), over-dispersed (Moran’s $I>0$) or clustered (Moran’s $I<0$) [26]. Next, Kulldorff’s scan statistic in SaTScan (v9.1.1, http://www.satscan.org/) was used to identify the location of clusters of communes with high rates of typhoid fever over space and time [27,28]. A cylindrical window was used to scan the area for clusters, with the size of the circle corresponding to the spatial scan and the height of the cylinder corresponding to time. The significance of the detected clusters was assessed by a likelihood ratio test, with a $p$-value obtained by 999 Monte Carlo simulations generated under the null hypothesis of random spatiotemporal distribution. In this analysis, scan windows were used to fit discrete Poisson models. For the sublineage-specific analyses, all case communes were included and those without cases of a specific sublineage were classified as having 0 cases. The upper limit for cluster detection was specified as 25% of the study population over each year. All maps were created in ArcGIS 10.2 (ESRI, Redlands, CA, USA).

Whole-genome sequencing and phylogenetic analysis

Of the 284 S. Typhi isolates collected between 2007 and 2014, a total of 209 (74%) collected between 2007 and 2012, were subjected to genomic DNA extraction using the Wizard Genomic DNA Extraction Kit (Promega, Wisconsin, USA) (SI Table). Two micrograms of genomic DNA was subjected to WGS on an Illumina HiSeq2000 platform following the manufacturer’s recommendations to generate 100bp paired-end reads. All reads were mapped to the reference sequence of S. Typhi strain CT18 (Accession no: AL513382), plasmid pHC11 (AL513383) and pHC42 (AL513384) using SMALT (version 0.7.4). Candidate SNPs were called against the reference sequence using SAMtools and filtered with a minimal mapping quality of 30 and a quality ratio cut-off of 0.75. The allele at each locus in each isolate was determined by reference to the consensus base in that genome, using samtools mpileup and removing low-confidence alleles with consensus base quality $\leq 20$, read depth $\leq 5$ or a heterozygous base call. SNPs in phage regions, repetitive sequences or recombinant regions were excluded, resulting in a final set of 750 chromosomal SNPs. Strains belonging to haplotype H58 were defined by the SNP g6A-C1047T (position 2,348,902 in S. Typhi CT18_ BIP33, as previously described [12,13]).

A maximum likelihood phylogenetic tree was constructed from a 188-chromosomal SNP alignment of H58 isolates with RAxML (version 7.8.6) using the generalised time-reversible model (GTR) and a gamma distribution to model site-specific rate variation (GTR+Γ) nucleotide substitution model in RAxML. Support for the ML phylogeny was assessed via 1,000 bootstrap pseudo-analyses of the alignment data. Phylogenetic subgrouping was defined based on monophyletic groups (lineages) with well-supported bootstrap value (≥85%).
To investigate the short-term divergence within the bacterial population and the transmission within the local population, a minimum spanning tree was reconstructed from the SNP alignment of lineage III and lineage IV identified in the ML tree (accounting for 95% of isolates) using the goeBURST algorithm in Phylowiz software (version 1.1) [29]. This algorithm identified seven sublineages based on similarity among allelic profiles and frequency of isolation within the population. Sequences with identical SNP profiles and isolated at the highest frequency within each sublineage were assigned as founder genotypes (viewed as the central nodes within each of the sublineages), with descendant genotypes (represented by terminal nodes surrounding the founder genotype) assigned based on similarity to founder SNP profiles. These descendant genotypes can differ from the parental genotype by a single or multiple SNPs. The raw sequence data for this study are available in the European Nucleotide Archive (ENA) under the accession numbers described in S1 Table.

Results
Baseline characteristics
Between 2007 and 2014, there were 284 microbiologically confirmed cases of typhoid fever caused by S. Typhi at AHC in Siem Reap. S. Paratyphi A was uncommon, with only three cases in 2008 followed by an isolated outbreak in 2013–2014 (38 cases). A total of 262/284 (93%) of the confirmed S. Typhi cases lived within a 100 km radius of AHC and spanned 78 communes; these 78 communes were selected for the spatial comparison and the typhoid fever population level risk factor analyses. During this same period there were 19,857 admissions with an ICD-10 discharge diagnosis other than typhoid fever originating from the same geographic area. The baseline characteristics of all communes and those with at least one case of typhoid fever are shown in Table 1.

Of the 262 cases of typhoid fever living within a 100 km radius of AHC, the median age was 8.2 years (interquartile range (IQR): 5.1–11.5 years). Additionally, 62/262 (24%) of the cases were less than five years of age and 142/262 (54%) were female. As shown in Fig 1a, the absolute number of confirmed cases of typhoid fever increased dramatically (from 12 cases per year to 71 cases per year) between 2009 and 2012, but then declined in 2013 and 2014 (28 and 45 cases in 2013 and 2014; respectively); data from our non-confirmed typhoid cases also reflected this trend. Over this same time period (2009 to 2014) the number of patients attending AHC for other conditions (control population) mirrored the distribution of the cases (Fig 1b). There was seasonal variation in the number of typhoid cases, with the majority of the cases (178/262; 68%) occurring during the early monsoon months (April, May, June and July) (Fig 1c & 1d). In late monsoon months (August to October), the number of cases declined to less than two cases per month and generally remained below this threshold in the dry season (November to March) (Fig 1c & 1d).

Spatiotemporal clustering of typhoid fever cases
The majority of S. Typhi cases (241/284; 85%) originated from communes located within Siem Reap province (Fig 2). The median population density in communes with at least one case of typhoid fever was 119 people/km² (IQR: 60–212), and 70/78 (90%) of communes with a typhoid fever case were classified as rural. Compared to typhoid cases, the non-typhoid fever population controls came from a larger area (243 communes), the median population density of which was lower at 106 people/km² (IQR: 53–210); however, a similar proportion of these communes (220/243; 91%) was also classified as rural (Fig 2).

The estimated median commune level minimum incidence of reported cases of typhoid fever over the study period was 0.62/1,000 children aged <15 years (IQR: 0.37–1.02; range: ...
Table 1. Baseline characteristics of all communes and those with at least one case of typhoid fever.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All communes</th>
<th>Typhoid communes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>median</td>
<td>IQR</td>
</tr>
<tr>
<td>Population density/km²</td>
<td>105.7</td>
<td>53–219</td>
</tr>
<tr>
<td>Elevation, m</td>
<td>17</td>
<td>12–28</td>
</tr>
<tr>
<td>Distance to lake, km</td>
<td>45.3</td>
<td>24–83</td>
</tr>
<tr>
<td>Average household size</td>
<td>4.8</td>
<td>4.6–5.0</td>
</tr>
<tr>
<td>Percent of population &lt;15 yr</td>
<td>36.4%</td>
<td>34–39%</td>
</tr>
<tr>
<td>Median age of population, yr</td>
<td>19.5</td>
<td>18–21</td>
</tr>
<tr>
<td>Adult literacy</td>
<td>72.8%</td>
<td>59–82%</td>
</tr>
<tr>
<td>Female adult literacy</td>
<td>65.3%</td>
<td>50–75%</td>
</tr>
<tr>
<td>Total attending school</td>
<td>25.8%</td>
<td>25–31%</td>
</tr>
<tr>
<td>Female attending school</td>
<td>25.9%</td>
<td>24–25%</td>
</tr>
<tr>
<td>Female education &gt;25 years /1,000 population</td>
<td>85.6</td>
<td>63–101</td>
</tr>
<tr>
<td>Primary not completed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary/upper secondary</td>
<td>27.8</td>
<td>16–55</td>
</tr>
<tr>
<td>Secondary or above</td>
<td>0.51</td>
<td>0.1–1.6</td>
</tr>
<tr>
<td>Toilet, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>83.1%</td>
<td>63–92%</td>
</tr>
<tr>
<td>Sewage</td>
<td>5.2%</td>
<td>2–14%</td>
</tr>
<tr>
<td>Septic tank</td>
<td>3.9%</td>
<td>1–16%</td>
</tr>
<tr>
<td>Pit latrine</td>
<td>2.0%</td>
<td>1–5%</td>
</tr>
<tr>
<td>Drinking water, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fiped</td>
<td>1.6%</td>
<td>1–4%</td>
</tr>
<tr>
<td>Tube/tube well</td>
<td>10.2%</td>
<td>3–28%</td>
</tr>
<tr>
<td>Dug well</td>
<td>29.9%</td>
<td>11–56%</td>
</tr>
<tr>
<td>Spring river</td>
<td>23.9%</td>
<td>4–64%</td>
</tr>
<tr>
<td>Drinking water location, % of households</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within premises</td>
<td>19.3%</td>
<td>10–55%</td>
</tr>
<tr>
<td>Near premises</td>
<td>31.1%</td>
<td>22–45%</td>
</tr>
<tr>
<td>Away premises</td>
<td>41.8%</td>
<td>23–56%</td>
</tr>
</tbody>
</table>

IQR: interquartile range

0.5–11.36). The reported incidence varied significantly across the 78 communes. Kampong Kleang commune (Sotr Nikom district, Siem Reap) showed the highest incidence of typhoid fever over the study period with 11.36 cases of typhoid fever /1,000 population of children aged <15 years (Fig 2c). This area is renowned for its floating villages and is situated on the edge of Tonle Sap Lake, approximately 35 km southeast of Siem Reap City. The second highest incidence was identified in Kaol Chveang commune (Aek Phnum district, Battambang, 33 km southwest of Siem Reap City) with 4.1 cases/1000 people aged <15 years over the study period (Fig 2c). Both of these areas experience heavy flooding when the Tonle Sap Lake expands during the rainy season.

Overall, there was some evidence of positive spatial autocorrelation (case clustering) across the 78 communes that had at least one case of typhoid fever between 2007 and 2014 (Moran’s I = 0.11, p<0.056). The magnitude of this autocorrelation varied over time, and was the most significant in 2013 (Moran’s I = 0.19, p<0.019) but was non-significant in other years. We
were able to identify three significant spatiotemporal clusters associated with high rates of typhoid fever. The first occurred in 2008 toward the west of the study area and had a radius of 23.6 km; this cluster had 1.27 predicted cases and 10 observed cases (relative risk [RR] = 8.17, $p = 0.002$). The second cluster occurred in 2012 in the central northern area and had a radius of 10.8 km, with 1.67 predicted cases and 12 observed cases (RR = 7.47, $p < 0.001$). The final cluster occurred in 2013 in the southeastern area and had a radius of 13.5 km, with 0.88 predicted cases and 14 observed cases (RR = 16.8, $p < 0.0001$) (Fig. 2d).

The population structure of *Salmonella* Typhi in Siem Reap province, Cambodia

The resulting WGS data demonstrated that 97% (203/209) of the sequenced Cambodian isolates could be attributed to haplotype H58. The majority (159/203, 98%) of the H58 isolates
Fig 2. The spatial distribution of typhoid fever cases in Siem Reap province, Cambodia. a) North oriented map of Cambodia, the black cross shows the location of AHC. b) Map showing the population density (people/km², color-coded in key) of the 76 communes within the typhoid study area. AHC is shown by the black cross, the black border denotes Siem Reap province and the left and right asterisks mark the locations of the communes with highest incidence of typhoid fever; Kach Chilveang and Kampong Kleang, respectively. c) Map of the study area showing the rate of reported typhoid cases per 1,000 population under the age of 15 years (color-coded in key). d) Map of the study area showing significant spatiotemporal clusters of typhoid during the study period, the size of the grey circles corresponds to the radius of the cluster and the years of the clusters are denoted.

doi:10.1371/journal.pntd.0004785.g002
exhibited intermediate susceptibility against fluoroquinolones (0.12–0.5 μg/mL) via the common amino acid substitution of serine to phenylalanine at codon 83 (S83F) in the DNA gyrase protein encoded by gyrA. There was a strong association between haplotype H58 and an IncHI1 plasmid, which confers an MDR phenotype, with 89% (180/203) of the H58 isolates harboring the common IncHI1 plasmid and the corresponding antimicrobial resistance phenotype. For the six non-H58 isolates, no mutations were observed in the gyrA gene, while two (33%) carried the same IncHI1 plasmid as found in the H58 isolates. We identified 188 SNPs across the H58 population and, from a SNP-based phylogeny, identified the circulation of at least four lineages of H58 circulating in the selected area of Cambodia between 2007 and 2012 (Fig 3). These lineages, designated here as I-IV, differed from each other by as little as three to
five SNPs and were phylogenetically well-supported (bootstrap values ≥ 87%). The majority of the H58 isolates fell into lineage IV (152/203, 75%) and lineage III (41/203, 20%).

The spatiotemporal distribution of Salmonella Typhi genotypes

To investigate short-term evolutionary traits within the identified lineages, we constructed a SNP-based minimum spanning tree (Fig 5b). Using these data, we were able to investigate the local population dynamics and detect several clonal clusters emerging from lineage III (IIIa-IIIc) and lineage IV (IVa-IVc). SNPs defining these sublineages are shown in Table S2. Our data show a complex temporal distribution of S. Typhi H58 sublineages circulating in this location between 2007 and 2012 (Fig 4a). The distribution of these various strains was highly dynamic, with strain replacements, potential extinctions and the specific microevolution and expansion of H58-IVc (Fig 4a). In 2011 and 2012, H58-IVc became the dominant genotype, accounting for 44% (184/42) and 85% (61/72) of all S. Typhi isolates in these years, respectively.

We next aimed to identify spatiotemporal clustering of the various S. Typhi H58 sublineages, and found that IIIc, IVc, IVb and IVc all displayed significant evidence of clustering over space and time. Notably, the locations of these clusters were generally different between sublineages, signifying some degree of geographical variation of the circulating S. Typhi strains. For example, we identified significant clustering of H58-IIIc in the western part of the study area in 2011 (p<0.001, RR 26.7, radius 36km) (Fig 4b) and clustering of the emergent

![Graph and maps showing spatiotemporal distribution of Salmonella Typhi genotypes in Siem Reap province, Cambodia.](https://example.com/graph_image)

**Fig 4.** The spatiotemporal distribution of the various Salmonella Typhi lineages/sublineages in Siem Reap province, Cambodia. (a) Bar chart showing the annual distribution of the various S. Typhi lineages/sublineages from 2007 to 2012; sublineages are color-coded as in Fig 5b. (b) Maps showing significant spatiotemporal clusters identified for sublineages IIIc, IVa, IVb and IVc. The timing of each cluster is shown by the year in black text and the dotted circle represents the radius of the detected cluster. Background colors represent the rate of each sublineage per 1,000 population aged under 15 years. The incidence rates vary between sublineages, ranging from 0 to a maximum of 9.8 (IIIc), 3.12 (IVa), 2.86 (IVb) and 9.84 (IVc) 8.84 cases/1,000 population aged under 15 years.

*doi:10.1371/journal.pntd.0004785.g004*
H58-IVc strain in both 2011 (Kampong Khleang commune, p<0.001, RR: 39.4, radius: <1km) and in two locations in 2012 (smaller cluster, p = 0.017, RR: 5.17, radius: 6.2km; larger cluster, p<0.001, RR: 5.87, radius: 31.9km).

Population risk factors for typhoid fever

Finally, we investigated associations between rates of typhoid in children and demographic and sanitation variables at the commune level. We found a number of significant risk factors (e.g. low female education level and collection of drinking water near the household premises) and protective factors (e.g. higher population density, elevation from lake and attendance at school) associated with the rate of typhoid hospitalizations in the univariate analysis (Table 2). However, after controlling for confounders, we found that the distance of the centroid of the commune to the perimeter of the lake was strongly and significantly associated with rate of typhoid cases (10km increase in distance from the lake, incidence rate ratio (IRR): 0.38, 95%CI 0.26–0.55, p<0.001) (Table 3). Furthermore, the relative numbers of households within the commune connected to public sewage services and households using a sunken well were also strongly protective, however these associations were reversed through interaction with increasing number of households using wells and distance from the lake, respectively.

Table 2. Regression results highlighting factors associated with typhoid cases.

<table>
<thead>
<tr>
<th>Commune characteristic</th>
<th>Univariable</th>
<th>Multivariable</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IRR (95% CI)</td>
<td>p</td>
</tr>
<tr>
<td>Population density*</td>
<td>0.81 (0.72–0.95)</td>
<td>0.006</td>
</tr>
<tr>
<td>Elevation, 10m</td>
<td>0.89 (0.81–0.98)</td>
<td>0.026</td>
</tr>
<tr>
<td>Distance to lake, 10km</td>
<td>0.81 (0.74–0.89)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Average household size</td>
<td>1.54 (0.66–3.57)</td>
<td>0.317</td>
</tr>
<tr>
<td>Total attending school/1,000*</td>
<td>0.11 (0.04–0.33)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Female education &gt;25 years /1000 population*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Primary not completed</td>
<td>2.59 (1.53–4.30)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Primary/1ower secondary</td>
<td>0.94 (0.70–1.26)</td>
<td>0.654</td>
</tr>
<tr>
<td>Secondary or above</td>
<td>0.96 (0.80–1.17)</td>
<td>0.714</td>
</tr>
<tr>
<td>Telets per 1000 people*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>1.24 (0.84–1.79)</td>
<td>0.246</td>
</tr>
<tr>
<td>Sewage</td>
<td>0.96 (0.81–1.15)</td>
<td>0.676</td>
</tr>
<tr>
<td>x households with wells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Septic tank</td>
<td>0.89 (0.75–1.09)</td>
<td>0.050</td>
</tr>
<tr>
<td>Pit latrine</td>
<td>0.95 (0.73–1.29)</td>
<td>0.621</td>
</tr>
<tr>
<td>Dring water, hh/1000 hh*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Piped</td>
<td>0.87 (0.74–1.03)</td>
<td>0.104</td>
</tr>
<tr>
<td>Tucer/pipe well</td>
<td>0.62 (0.73–0.90)</td>
<td>0.001</td>
</tr>
<tr>
<td>Dog well</td>
<td>0.83 (0.64–1.03)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>x distance to lake</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spring/silver</td>
<td>1.15 (1.05–1.25)</td>
<td>0.003</td>
</tr>
<tr>
<td>Dring water location, hh/1000 hh*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Within premises</td>
<td>0.71 (0.55–0.93)</td>
<td>0.013</td>
</tr>
<tr>
<td>Near premises</td>
<td>3.38 (2.24–5.16)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Away premises</td>
<td>0.88 (0.73–1.06)</td>
<td>0.177</td>
</tr>
</tbody>
</table>

*lag of the variable was included; hh: household

doi:10.1371/journal.pntd.0004785.s002
(Table 2). Finally, a high number of households reporting drinking water retrieval from 'within the household premises' was also associated with a significant protective effect (log households/1,000 households, IDR: 0.65, 95% CI: 0.49–0.86, \( p = 0.003 \)).

Discussion

In this study we combined conventional epidemiological methods, current genome sequencing tools and geospatial mapping to add insight into the epidemiology of typhoid fever in pediatric patients attending a single healthcare facility in central Cambodia. The majority of recent typhoid fever studies originate from urban locations in low-income countries. This study provides a new perspective into this important community-acquired infection from a predominantly rural setting. The primary finding of this study is that there is a considerable and widespread burden of pediatric typhoid fever in rural Cambodia, thus questioning the dogma that typhoid fever is predominantly geographically restricted to urban populations with poor sanitation systems [9,27]. Our data are consistent with findings from a recent study conducted across sub-Saharan Africa [30]. The Typhoid Surveillance in Africa Programme (TSAP) found a large burden of typhoid fever in younger children and almost equivalent population incidences between urban and rural settings. This distribution was most apparent in West Africa (Burkina Faso and Ghana) and was similarly restricted to children aged less than 15 years [30]. Therefore, we infer that the epidemiology of typhoid fever in Cambodia may be more similar to contemporary observations from sub-Saharan Africa, as opposed to the urban distribution that has commonly been observed across much of Asia [7,31].

The impending availability of Vi-conjugate vaccine raises the question of who should be given this vaccine and when it should be given to obtain maximum benefit in the control of typhoid fever [32,33]. This issue is complicated by a lack of population-based incidence data and a poor understanding of the burden of disease in school and preschool aged children, for whom the conjugated form of the Vi polysaccharide vaccine would be particularly beneficial [34]. Our data indicate a substantial burden of typhoid fever in school and preschool aged children in this area, with a hospital-based incidence (i.e. a minimum population incidence) of 11.36 cases of typhoid fever/1,000 population in children aged <15 years over the study period. The overall burden of typhoid fever in this population is likely to be greater than we have estimated due to poor sensitivity of blood culture and restriction of the study to a single healthcare center. Siem Reap province could be a suitable location in which to trial, or even introduce, the next generation typhoid vaccines in Cambodia that have been tested elsewhere [32]. Further, we suggest that immunizing school-aged children in the period prior to the wet season may provide the most economic and prudent approach for vaccine introduction.

Between 2007 and 2012, we observed a sharp increase in the number of typhoid cases concurrent with an increasing geographic expansion. We also observed that typhoid fever in this population followed a seasonal pattern, suggesting an association with rainfall and potentially with localized flooding and the contamination of water sources. The population-based risk factors support these hypotheses, as living further away from Tonle Sap Lake and access to water within the household were highly protective. Additionally, we found that two communes located next to the lake (Koh Kong and Kampot) had the highest incidence of typhoid fever and had large clusters of cases in 2008 and 2013. This case clustering in specific locations warrants further investigation at the household level to understand specific sanitation-associated risk factors and likely exposures to S. Typhi in this setting [35]. It appears that access to lake water in some of these communes, such as Koh Kong, is vital for the household water supply and we hypothesize that the lake water is more prone to localized fecal contamination at specific times throughout the year.
Using targeted SNP-specific PCR, we have previously shown that MDR H58 S. Typhi strains dominate in this population [18]. Our WGS investigation confirmed these findings and identified additional diversification in this population. We were able to separate these H58 strains into seven (IIIa, IIIb, IIIC, IVC, IVA, IVb, IVc) major sublineages. These discrete groups varied in size and were segregated by only limited numbers of SNPs. We did observe some evidence of expansion of sublineage IVc between 2009 and 2012; this correlated with several spatiotemporal clusters suggesting small disease outbreaks. We currently cannot explain the expansion of this group and our strain selection for sequencing was limited by the availability of strains isolated only up to 2012. Despite some clustering of closely related strains, the overall temporal and spatial distribution of strains was random, with a range of S. Typhi H58 sublineages circulating throughout the study period, which is similar to patterns described in urban settings in Asia [14,36].

This study has some limitations. First the data originated from patients attending a single healthcare facility, without the added support of healthcare utilization data. This approach, while cost-effective, induces bias in the spatial and risk factor analyses. Furthermore, while the associations identified in the regression analysis are plausible and provide direction for future investigations, they should be viewed with caution. The population level census data does not allow examination of exposures at an individual or household-level and provides only broad epidemiological evidence. However, the association with distance to the lake and water and sanitation variables suggests these factors should be examined more rigorously in the future with respect to the dynamics of typhoid fever outbreaks. Similarly, the identification and location of the spatiotemporal clusters should be interpreted with some degree of caution. Communities without cases were not included in the cluster analyses due to a lack of data as to whether these regions truly lacked typhoid cases. A dataset with more complete spatial information on presence and absence of typhoid would permit a more reliable analysis.

In conclusion, we find a large burden of typhoid fever in children in rural Cambodia. Our conventional population-based risk factor analysis identified access to water in the household and increasing distance from Tonle Sap Lake as protective against typhoid fever in communities. Spatial mapping and WGS provided additional resolution to investigate these findings and confirmed that proximity to that lake was associated with discrete disease clusters. We confirmed the dominance of MDR H58 S. Typhi in this location and found a substantial amount of diversification within this lineage. Our data provide a platform for additional studies in the Cambodian population and suggest that this is a suitable location in which to introduce Vi conjugate vaccines for school children.

Supporting Information
S1 Table. Strain list and accession numbers for organisms used in this study. (XLS)
S2 Table. SNPs defining H58 sublineages. (XLS)
S1 Checklist. STROBE checklist. (DOC)

Acknowledgments
We wish to acknowledge the staff of Ankor Hospital for Children for their assistance with this study and their on-going commitment in providing healthcare to the local community.
Author Contributions
Conceived and designed the experiments: DPT CNT MAR PT CMP SB. Performed the experiments: DPT CNT MAR SIm SoP VR GM NTVT LW KEH VW DP PT CMP. Analyzed the data: DPT CNT MAR NTVT KEH VW DP GET PT CMP. Contributed reagents/materials/analysis tools: DPT CNT MAR LW KEH VW DP GET ND GD PT CMP SB. Wrote the paper: DPT CNT MAR VR GM LW KEH GET ND GD PT CMP SB.

References
PLOS Neglected Tropical Diseases | DOI:10.1371/journal.pntd.0004785  June 22, 2016  16 / 16

284
Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events

The emergence of multidrug-resistant (MDR) typhoid is a major global health threat affecting many countries where the disease is endemic. Here whole-genome sequence analysis of 1,832 *Salmonella enterica* serovar Typhi (S. Typhi) identifies a single dominant MDR lineage, H58, that has emerged and spread throughout Asia and Africa over the last 30 years. Our analysis identifies numerous transmissions of H58, including multiple transfers from Asia to Africa and an ongoing, unrecognized MDR epidemic within Africa itself. Notably, our analysis indicates that H58 lineages are displacing antibiotic-sensitive isolates, transforming the global population structure of this pathogen. H58 isolates can harbor a complex MDR element residing either on transmissible IncHI1 plasmids or within multiple chromosomal integration sites. We also identify new mutations that define the H58 lineage.

This phylogeographical analysis provides a framework to facilitate global management of MDR typhoid and is applicable to similar MDR lineages emerging in other bacterial species.

S. Typhi, the primary global cause of human typhoid (enteric fever), is a monophyletic serovar of *S. enterica*. Unlike many *Salmonella*, S. Typhi are highly restricted to infection of humans and are associated with systemic infection, prolonged fever and an asymptomatic carrier state. Typhoid is still a common disease in many regions of the world with poor infrastructure and limited economic development; and is also a risk for travelers who visit such regions. It is estimated that 20–30 million cases of typhoid occur annually, although deaths are less frequently reported than before the availability of effective antimicrobials.

In addition to improvements in access to clean water and sanitation, typhoid can potentially be controlled by other interventions such as vaccination and antimicrobial therapy. Chloramphenicol, ampicillin and trimethoprim-sulfamethoxazole were traditional first-line drugs commonly used to treat acute typhoid, and these agents continue to be used in areas of the world where S. Typhi are deemed susceptible. However, since the 1970s, S. Typhi have emerged that display multidrug resistance, defined as resistance to the above antimicrobials, compromising treatment.

The early emergence of MDR S. Typhi was driven in large part by the acquisition of IncHI1 plasmids carrying antibiotic resistance genes and, more recently, by chromosomal mutations associated with resistance to fluoroquinolones, and MDR strains have been reported across Asia and Africa.

Phylogenetic analysis, initially based on subgenomic DNA sequences but later on whole-genome DNA sequences, showed that the global S. Typhi population is highly clonal and likely originated from a common ancestor that moved into the human population several thousand years ago. It also indicated that the population is relatively small and that recombination between S. Typhi and other *Salmonella* is rare. Simple SNP-based typing schemes have been developed that stratify the S. Typhi population into haplotypes, and these schemes are now used to unequivocally map new isolates to the phylogeny. Notably, this approach identified a single emerging, highly clonal MDR haplotype of S. Typhi, H58, which is being reported with increasing frequency in many countries in Africa and Asia.

Within the H58 lineage, IncHI1 MDR plasmids of the restricted subtype PSTe (ref. 23) and chromosomal point mutations conferring quinolone resistance are common. However, relatively little is known about the emergence and evolutionary history of the H58 lineage or how it is moving across endemic regions. Here we have used phylogenetic analysis based on the whole-genome sequences of a global collection of S. Typhi from 63 countries to investigate the genomic architecture of this highly successful S. Typhi lineage.

**RESULTS**

**Phylogeny of H58**

Of a global collection of 1,832 sequenced S. Typhi (listed in Supplementary Tables 1 and 2), 853 (47%) belonged to haplotype H58, initially defined by the SNP g/jra-C1047T (position 2,348,902 in S. Typhi CT18, RP33 in ref. 17). The earliest H58 isolates in our collection were from 1992 (Fiji) and 1993 (Fiji and Vietnam), and H58 isolates were represented every year from 1992 to 2013, at a mean rate of 40% per year (Fig. 1a). 1158 isolates formed a tight cluster within the whole genome maximum-likelihood phylogeny (Fig. 1b), forming a unique lineage separated by 151 SNP differences from the nearest neighboring non-H58 cluster, which consisted exclusively of isolates from Fiji.
Figure 1. Population structure of the S. Typhi isolates analysed in this study. (a) Temporal distribution of the S. Typhi isolates included in the study. (b) Rooted maximum likelihood tree of S. Typhi inferred from 22,145 SNPs, rooted using an outgroup (S. enterica serovar Paratyphi A, isolate 9903). A, 4Outgroup; Paratyphi_ATCC_2010. The colored ring indicates the geographical origin of the isolates. Red arc, H58 lineage; blue dashed lines, public reference genomes reported in Holt et al.7, including the CH38 (K) reference genome (ALU122382); black dashed lines, other publicly available genomes. Branch lengths are indicative of the estimated substitution rate per variable site.

(n = 140) and Tonga (n = 2). Individual H58 isolates differed from the most recent common ancestor (MRCA) of the H58 lineage by a median of just two SNPs, and the median distance between pairs of H58 isolates was six SNPs, strongly indicative of recent clonal expansion. Nearly all the H58 isolates (93%; 797/853) had ≤5 isolate-specific SNPs, consistent with frequent transmission relative to substitution mutations. This finding was in contrast to that for the rest of the S. Typhi tree, which included a wide diversity of isolates (Fig. 1b), among which only 66 (642/970) had ≤5 isolate-specific SNPs (P < 0.0001; Fisher’s exact test) (Supplementary Fig. 1).

The population structure within H58 was consistent with our previous work defining two major sublineages of H58 (Lineages I and II)2,12,27 but provided much greater resolution of substructure with a strong phylogeographical signal (Fig. 2). Our H58 S. Typhi isolates were collected from 21 countries across Asia, Africa and Oceania. We observed strong phylogeographical clustering within 13 countries (Fig. 2), indicating transmission of H58 within these locations. Although our sample spans distinct time periods in different locations, in most cases, the localized subclades were isolated from the same country over 24 years, indicating the establishment of long-term local reservoirs (Fig. 3a,b).

These data demonstrate that H58 is now widely disseminated across distinct geographical areas, and the phylogeny provides several insights into the spatial patterns of its spread. There were numerous instances of very closely related isolates from different countries (Fig. 2), which indicate likely transfer events or regional outbreaks and identify routes for geographical dissemination (Fig. 3a,b). Maximum-likelihood analysis of inter-region transfers based on the maximum-likelihood phylogenetic and locations

Figure 2. Population structure of the S. Typhi H58 lineage. Rooted maximum likelihood phylogeny inferred from 1,534 SNPs identified in the H58 H58 isolates, rooted using an S. Typhi isolate from the nearest neighbor cluster of non-H58 isolates as an outgroup (black-filled circle, isolate 100060, 5, 62). The colored ring indicates the countries of isolation; countries discussed in the text are labeled around the tree. Branch lengths are indicative of the estimated substitution rate per variable site.
Figure 2 Geographical persistence and routes for dissemination of S. Typhi H58.
(a) Maximum-likelihood tree for the H58 lineage (I and II), with clades containing isolates from a single country collapsed into nodes (circles), sized to indicate the number of isolates in the clade and colored by country of isolation. Branches are colored to indicate the country of origin of descendant nodes. (b) Years of isolation for each phylogenographical cluster in the tree, indicated by lines spanning the earliest and latest years of isolation for each cluster and colored to indicate the country. Four regions with extensive local clonal expansion are highlighted by shaded boxes, spanning the phylogenetic (x-axis) and temporal (y-axis) extent of the expansion. Locations from which singleton isolates were clustered within the phylogenetical clusters are shown to the right, indicative of further onward transmission.

where isolates were collected highlighted several candidate intercontinental transfers (Figs. 4). These data suggest that South Asia was an early hub for H58, from which it was propagated to many locations around the world, including countries in Southeast Asia, western Asia and East Africa, as well as Fiji (Figs. 2–4). Most of the diversity in lineage I was present among Indian isolates, with unique local subclusters detected in neighboring countries (Nepal and Pakistan) and in Africa, indicative of occasional transfers out of Asia. In contrast, lineage II was associated mainly with Southeast Asia (Vietnam, Cambodia and Laos), with evidence of transmission to Thailand, Pakistan, Fiji and Africa.

There have been sporadic reports of the emergence of S. Typhi in Africa14–29. Indeed, H58 isolates were predominant among the eastern and southern African S. Typhi isolates (63%) (Supplementary Fig. 2); in contrast, the H58 lineage was relatively rare in northern, western and central Africa. H58 lineages I and II were detected in Kenya, Tanzania, Malawi and South Africa, providing compelling evidence for multiple introductions of H58 S. Typhi from South Asia into the continent (Figs. 2–4). In addition, we uncovered evidence of an unreported recent wave of transmission of H58, based on 138 isolates, from Kenya to Tanzania and on to Malawi and South Africa (Figs. 2 and 3). These isolates differed from one another by an average of 10 (range of 0–30) SNPs, consistent with a recent clonal expansion. Therefore, this analysis demonstrates an ongoing epidemic of H58 typhoid across countries in eastern and southern Africa.

Dating the emergence of H58

Estimating mutation rates and divergence dates within the S. Typhi population has been challenging, as S. Typhi is known to establish persistent asymptomatic carriage, during which time it likely evolves at a different rate than during acute infection, disrupting the molecular clock37. Here a temporal signal was barely detectable across the full S. Typhi maximum-likelihood tree, assessed via linear regression of root-to-tip branch lengths on the basis of year of isolation (correlation coefficient (R) = 0.09 (95% confidence interval (CI) = 0.64–0.11); P = 0.0002, Fisher’s exact test). However, a moderate signal was evident within the H58 subtree (R = 0.60 (95% CI = 0.56–0.64); P < 1 × 10⁻⁷, Fisher’s exact test; Supplementary Fig. 2). This temporal signal was entirely destroyed by randomization of isolation dates (mean R = 0.01), indicating that uneven sampling of the H58 lineage across space and time was not solely responsible for the observed association. We propose that a temporal signal was detectable within the H58 tree because these data capture epidemic spread over a relatively short time span (2–3 decades), whereas the wider S. Typhi tree represents much more variable population dynamics over thousands of years of evolution, including recent periods of endemic transmission, that differ greatly from the clonal expansion of H58.

We therefore proceeded to estimate the divergence date of the H58 lineage via Bayesian phylogenographical modeling of the H58 population, implemented in Bayesian Evolutionary Analysis Sampling Trees (BEAST)40. To limit potential bias due to highly variable sampling intensities in different geographical locations, we performed BEAST analyses on several cross-sections of 114 H58 isolates (13% of the total), each sampled from 21 countries and spanning the years 1992 to 2013. The combined estimate for the median substitution rate within the H58 population was 1.42 × 10⁻² substitutions per site per year (95% highest posterior density (HPD) = 1.0 × 10⁻² to 1.8 × 10⁻²), equivalent to the accumulation of 0.63 SNPs per genome per year (95% HPD = 0.99 to 0.67). The analyses predicted that the MICA of all extant H58 strains existed ~25 years ago (median calendar year for divergence, 1989; 95% HPD = 1985–1992) and that the effective population size of H58 increased dramatically after 1993 (Supplementary Fig. 2). Although the BEAST analysis was limited
because of the moderate strength of the temporal signal, these results are consistent with the very low numbers of SNPs within the H58 lineage, the clear evidence of clonal expansion from the maximum-likelihood tree (Figs. 1b and 3a,b) and epidemiological data reporting increasing rates of multidrug resistance in Asia in the early 1990s (refs. 13,32).

Multidrug resistance in H58

The H58 lineage is associated with high levels of multidrug resistance and reduced susceptibility to fluoroquinolones.12,13,15

Acquired resistance genes were identified in 671 of the 1,832 S. Typhi isolates, including 68% of the H58 isolates in comparison to just 9% of non-H58 isolates (P < 10−10, Fisher's exact test). In our collection, which included 15 countries with ≥2 MDR isolates, H58 was significantly associated (P < 0.01) with multidrug resistance in nearly all these locations, with the exception of central and western Africa, where multidrug resistance was detected but H58 was not (Table 3). The most common resistance genes detected were blaTEM-1 (ampicillin resistance), dfrA7, sul1 and sul2 (resistance to trimethoprim and sulfonamides, respectively, and to trimethoprim-sulfamethoxazole collectively), catA1 (chloramphenicol resistance) and strA/B (streptomycin resistance). These genes were each found in >540 H58 isolates, including in 525 isolates that carried all 7 genes. These genes are encoded within a Tn2670-like complex transposable element comprising transposon Tn2670, which carries blaTEM-1, strA/B and sul2, inserted into transposon Tn2670, which itself comprises Tn21 carrying a class I integron (including sul2, with dfrA7 in the gene cassette (refs. 12,33) (Fig. 5). In addition, 405 H58 isolates harbored the tcrA gene located in Tn7/6. Other acquired resistance genes were rare, identified in <1% of the H58 isolates.

Previous reports linked multidrug resistance in H58 S. Typhi to the Tn2670-like element described above encoded on IncH1 plasmids of the PST6 genotype.12,21,23 Here we identified IncH1-PST6 plasmids in 74% of the H58 isolates harboring the MDR element, including isolates from Southeast Asia, East Africa and South Asia (Supplementary Table 3), indicating intercontinental transmission of the IncH1-PST6 MDR plasmid with its H58 S. Typhi host (Fig. 6). However, the remaining MDR H58 isolates lacked IncH1 plasmid sequences, indicating that the resistance-confering
genes must be located elsewhere in the genomes of these 139 isolates. We screened all H58 isolates for known plasmid replicons and identified non-IncHI MDR plasmids in 23 isolates but not among isolates carrying the Tn2670-like MDR element (Table 3). We carried out a copy of insertion sequence IS231 among the chromosomal genes STY3616 and STY3619, near cyca (Fig. 6). We propose that this insertion was acquired early in the H58 lineage, originating from the IncHI1-PT16 plasmid. Consistent with this hypothesis is the present analysis showing that IS231 was absent from only one of the few other isolates that were tested (Fig. 6). Long-read sequencing of 2012 Indian isolate BRH12960 (Supplementary Table 4) confirmed that the entire MDR locus, flanked by copies of IS231, was integrated at this location near cyca (Fig. 5), in agreement with other recent reports. The nucleotide sequence was identical to that in the IncHI1-PT16 plasmid of H58 isolate 1H24_1_48, Vetyer3_193_1997, which we also sequenced for comparison. Further analysis of our data identified new insertion sequences in MDR isolates that lacked the IncHI1 plasmid, including (i) 25 phyllogenetically related isolates mainly from Bangladesh, Pakistan and Iraq with IS231 in the ydaM gene, (ii) 9 related isolates from Fiji with IS231 in STY4438 and (iii) 1 isolate from India with IS231 in the yfaA gene (Fig. 6). Long-read sequencing of the 2013 isolate 124438 from India confirmed integration of the MDR locus in the ydaM gene (Fig. 5), which was confirmed integration at the other two sites by PCR. The distribution of isolates in the H58 tree indicates single integration events at each of the ydaM, STY4438 and yfaA loci but numerous independent integrations at the cyca site (Fig. 6). The latter suggests that the Tn2670-like element, which is flanked by IS231 sequences, may target existing IS231 sequences during its mobilization.

In addition, we found that four isolates harboring genes associated with aminoglycoside resistance. These isolates were 10393_2_14_Alg14_8842_2005 from Algeria carrying evA and three isolates from Indonesia carrying either movA (10349_1_50_Indo403_MovA_1983) or marR (9953_5_22_Indo340_2010 and 9953_5_42_Indo327_2010).

Quinolone resistance in H58

The primary targets of the fluoroquinolones are the DNA gyrase subunits (gyra and gyrB) and the topoisomerase IV components (parC and parE).Nonsynergistic mutations in the quinolone-resistance-determining regions (QRDRs) of each gene can decrease susceptibility to fluoroquinolones such as ciprofloxacin, which is commonly used in the treatment of typhoid.

We found that nonsynergistic changes in the QRDR of these four genes were far more common in the H58 isolates (59%) than in other S. Typhi (15%; P < 1 × 10⁻³). Fisher's exact test (Supplementary Table 5). The most frequent QRDR mutations were changes in codon 83 of gyrA encoding pSer83Phe (49% of H58 isolates) and p.Ser837Phe (9% of H58 isolates) substitutions (Supplementary Table 5). The distribution of gyrA substitutions within the H58 phylogeny indicates that these mutations have arisen independently on multiple occasions, consistent with our previous observations of convergent evolution and confirming that this region of gyrA is under strong positive selection (Supplementary Fig. 4). The accumulation of multiple mutations within the gyr and par genes can result in a higher minimum inhibitory concentration (MIC) for fluoroquinolones. Additionally, we detected multiple mutations in these genes in 199 isolates from the three regions: 190 H58 isolates (predominantly from Cambodia and India) and 9 non-H58 isolates (mainly from India) (Supplementary Fig. 5 and Table 3).

Transmissible fluoroquinolone resistance can occur in Salmonella via plasmid-mediated acquisition of tetracycline (tetM) genes. Here we identified such genes in seven H58 isolates, which carried both gyrA mutations and the gyrS gene that confer high-level resistance. The gyrS gene was present within a mobile element also containing blaTEM_1, mlt2, and catB4, in association with and possibly mobilized by IS256, an IncFIB(K) plasmid (Fig. 6 and Table 3).

Trends in antimicrobial resistance in H58

The data show some geographical differences in patterns of antimicrobial resistance within the H58 lineage (Supplementary Fig. 5). Multidrug resistance was common among H58 isolates from Southeast Asia in the 1990s, and in recent years, gyrA mutations have arisen on this background, resulting in high rates of MDR H58 with reduced susceptibility to fluoroquinolones. This observation likely reflects the therapeutic use of fluoroquinolones to treat typhoid over this period. Although we have few examples of H58 isolates from South Asia before 2000, the pattern is clearly different in this region, with the majority of isolates from recent years almost all harboring gyrA mutations but with low rates of multidrug resistance.
The situation appears to be different yet again in Africa, where the majority of recent isolates (mainly from Malawi, Kenya and South Africa) were identified as MDR but without gyrA mutations, potentially reflecting the continued use of traditional antimicrobial agents.

**Genomic signatures of the H58 lineage**

Because the S. Typhi H58 lineage emerged rapidly over the past 30 years, we searched for any distinctive genetic signatures, other than those for multidrug resistance, that might be facilitating its dissemination. Prophage-like elements are known hotspots for variations within S. Typhi and other S. enterica; however, H58 genomes shared five of the seven prophage-like elements previously identified in the S. Typhi reference isolate CT18 (haplotype H1), and only rare acquisitions of new phages were found (five phages, affecting 19% of the H58 isolates; **Supplementary Fig. 6 and Supplementary Table 6**). The SNPs that define the H58 lineage include several nonsynonymous changes in genes associated with pathogenicity, adaptation and chaperones (**Supplementary Table 7**). The affected genes consist of saaP; encoding a Salmonella pathogenicity island-2 (SPI-2)- associated protein involved in intracellular survival and persistence29, and the regulatory genes strA and carbB, which have been implicated in Salmonella virulence26,27. Additionally, H58-associated SNPs also included changes in genes involved in central or intermediary metabolism, including tyrP, tprB, betC, trnM, melH, yfbP, tbr, muoG and iaaA, and genes involved in membranes or structures, for example, iplP, yidA, ksgA, aceA, ycgT, etcG, ycgT and SBOV18161 (dyneE)31. All S. Typhi display substantial genome degradation, via the accumulation of deletions and inactivating mutations within coding sequences, to form pseudogenes; each S. Typhi genome has >200 pseudogenes, reflecting a loss of ~4% of protein-coding capacity22,28. Here we found that all H58 isolates additionally harbored a point mutation in the spgP gene (STY3001) resulting in a premature stop codon at position 155, effectively rendering H58 strains deficient for SpgP protein. SpgP is an SPI-1 effector protein known to have a role in modulating the host cell actin cytoskeleton via its GAP domain that targets CDC42 and RAC-1 (ref. 42). We previously identified this nonsense mutation in seven sequenced H58 isolates and a second distinct nonsense mutation in the same gene in the H58 isolate E98-3139. Convergent loss-of-function mutations such as this are quite rare in S. Typhi and may reflect a selective advantage for inactivation of this gene29.

**DISCUSSION**

Here, we provide the first comprehensive global phylogenographical analysis of the emerging MDR-associated S. Typhi clade known as H58, covering many of the key geographical regions where typhoid remains endemic. This analysis indicates a major ongoing clonal replacement of resident non-H58 S. Typhi haplotypes by this clade and identifies previously unappreciated inter- and intrainterventional transmission events. Smaller regional studies performed in different Asian countries and in Kenya have described the emergence of the H58 haplotype at local and country levels12,14,15,17,18,22. However, here we show the true global impact of H58, which is transforming the S. Typhi population structure across the world. Indeed, we show definitively that H58 has expanded dramatically since the early 1990s and that the MDR phenotype of H58 strains is largely determined by different regional antibiotic usage31. Further, our analyses show an emerging epidemic of MDR typhoid moving across Africa, potentially driven by this antimicrobial usage. The existence of this epidemic is supported by the available epidemiological data, which include increasing numbers of reports of MDR typhoid in Africa, and by observations from members of this consortium12,29,30. The H58 lineage has previously been associated with multidrug resistance, which may be a key factor driving its current expansion31. Here we show that this association holds across numerous countries in Asia and Africa, such that the majority of the global burden of MDR typhoid can be attributed to the H58 lineage. Intriguingly, our data indicate that multidrug resistance in H58 is tightly linked to the presence of a single Tn3670-like element that was probably first introduced via the IncHI1-PST6 plasmid but has since transferred to the S. Typhi chromosome in numerous distinct integration events, each affecting different subregions of H58. Such integration has recently been noted in isolates from Bangladesh (ycaJ and ydaA sites)31,14, and Zambia (cysA site)30; however, our data provide important context for these observations, showing that integrations are relatively frequent and have been occurring since the emergence of H58. Integration of the MDR locus into the chromosome may facilitate loss of the large IncHI1 plasmid, thus moderating any potential fitness burden while maintaining the MDR phenotype. A similar phenomenon has been observed in S. enterica, where chromosomal integration of an MDR transposon in the late 1970s appears to be associated with global dissemination of a single successful sublineage41, and in Salmonella that have acquired the salmonella Genomic island31. H58 strains also harbor a higher frequency of quinolone-resistance-associated mutations than other S. Typhi, potentially owing to enhanced exposure of the large, diversifying and frequently MDR H58 population to fluoroquinolones, coupled with the lack of a fitness cost associated with these mutations39.

These results provide, to our knowledge, the first global, large-scale, genome-based study of an MDR clade of S. Typhi. The global dissemination of H58 requires urgent international attention. Indeed, the arrival of S. Typhi H58 in Africa appears to be transforming the epidemiology of the disease, with MDR outbreaks of typhoid being reported where the disease was previously unappreciated or absent. It will be particularly important to control antimicrobial prescribing practices, including the use of prophylactic antimicrobials such as trimethoprim-sulfamethoxazole42 in this region, as such use likely promotes multidrug resistance. This study highlights the need for longstanding routine surveillance to capture epidemics and monitor changes in bacterial populations as a means to facilitate public health measures, such as the use of effective antimicrobials and the introduction of vaccine programs, to reduce the vast and neglected morbidity and mortality caused by typhoid.

**METHODS**

Methods and any associated references are available in the online version of the paper.

**Accession codes.** Raw sequence data have been submitted to the European Nucleotide Archive (ENA) under accession EPI001718. Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

**ACKNOWLEDGMENTS**

This work was supported by the Wellcome Trust. We would like to thank the members of the Pathogen Informatics team and the core sequencing teams at the Wellcome Trust Sanger Institute (Cambridge, UK). We are grateful to D. Harris...
ONLINE METHODS
Backbone isolates and sequencing. A total of 1,832 S. Typhii isolates were included in the study. These were isolated between 1905 and 2013 and originated from 63 countries spanning six continents (Asia, Africa, North and South America, Europe, Oceania) (Supplementary Table 1). This collection included 14 of the 19 S. Typhii isolates previously sequenced by Heit et al. 19, including the 2 isolates with finished reference genomes, CT18 (AABB 16735) and TV2 (~AABB 16735) (Supplementary Table 2). Seven other public S. Typhii references were downloaded from public databases and included in the analysis (Supplementary Table 3). Information on the source was available for 1,523 of the 1,832 S. Typhii isolates. The isolates included in this study were supplied by numerous contributing laboratories and were cultivated from a wide range of clinical specimens, including blood (90.9%), stool or rectal swab (9.3%), 127(1.5%), urine (0.4%), 61(0.3%), pus (0.3%), 56(0.6%), gallbladder fluid (0.1%), bile (0.1%), pleural fluid (0.1%), cerebral fluid (0.1%), and cerebrospinal fluid (0.1%). There were data on the age of the patient for 405 S. Typhii isolates, and these comprised 330 (82%) isolates originating from children (0-5 years old) and 75 (18%) isolates cultured from adults (>16 years old).

Each of the collaborating laboratories used their own individual methodologies for isolation of whole-genome DNA. For 231 isolates, DNA was prepared using the Wizard Genomic DNA Kit (Promega) according to the manufacturer’s instructions. Index-tagged paired-end Illumina sequencing libraries were prepared as previously described. 20 These were then combined into pools, each containing 96 uniquely tagged libraries and sequenced on the Illumina HiSeq 2000 or HiSeq 2500 platform according to the manufacturer’s protocols to generate tagged 200 bp paired-end reads.

In addition, the genomes of five H58 isolates were sequenced on the PacBio RS II platform (Pacific Biosciences) for better resolution of the integration of the large composite transposon element into the chromosome (Supplementary Table 4). Genomic DNA (3 μg) was sheared using the HydroShear Plus (DigitalBio), and a library was prepared using DNA Template Prep Kit 2.0 (Pacific Biosciences) according to the manufacturer’s instructions. Sequencing was performed on SMRT cells with XL polymerase and DNA Sequencing Kit C2 (Pacific Biosciences). De novo assembly was performed with Spade v1.0 (see URLs) and HGAP v1.0 (ref. 30) with default parameters. The contigs from Spade were visualized with the SPAdes package when the script detected a significant overlap between the beginning and end of contigs.

Read alignment and SNP detection. For analysis of SNPs, paired-end Illumina reads were mapped to the CTB reference genome of S. Typhii, including the chromosome and phtaC1 and pthaC2 plasmids, using SMALT (version 0.7.4) (see URLs) as previously described. 31 Candidate SNPs were identified as previously described. 31,32,33 Using SAMTools command mpigrep -d h1000 -l1000 -f > results.bed, we determined that a total of 125,673 SNPs with quality scores above 20 were collated into a single list of variant sites, and the allele at each SNP site in each isolate was determined by reference to the consensus base call for that genome (using SAMTools and removing low-confidence alleles with consensus base quality <50). SNPs contained in >5% of reads with mapping quality >30, depth >2, and reads per strand, strand bias P < 0.001, mapping bias P < 0.001 were called, and SNPs called in phage regions, repetitive sequences (354 kb; 7.4% of bases) in the CT18 reference chromosome, and defined previously 34 or recombinant regions (-180 kb) of the CT18 reference chromosome, identified using an approach described previously 35 were excluded. Results in a final set of 22,143 chromosomal SNPs identified in an alignment of length 4,275,037 bp.

Phylogenetic analysis. The maximum-likelihood phylogenetic tree shown in Figure 1B was built from the 22,143-SNP alignment of 1,832 isolates, plus a 5. Paratyphi A strain included as an outgroup for tree rooting, using RAxML (version 7.8.27) with the generalized time-reversible model and a Gamma distribution to model site-specific rate variation (the GTR+I+T substitution model). RAxML and using maximum-likelihood phylogenies was assessed via 100 bootstrap pseudoreplicates of the alignment data. A maximum-likelihood phylogenetic tree was also inferred separately from the SNP alignment of 853 H58 S. Typhii isolates using the same parameters as above (Fig. 2). The H58 phylogenetic tree was rooted using an S. Typhii isolate from the nearest neighboring cluster of non-H58 isolates (Supplementary Table 1). The evolutionary relationships were inferred from this collapsed H58 tree using discrete trait transition modeling, implemented in the mle traitmap function in the phytools R package. Briefly, each phylogeny was assigned to a geographical region on the basis of country of isolation or geographical region of isolation for travel-associated isolates, these regions were treated as discrete tip states on the H58 maximum-likelihood tree and a Markov model for the evolution of this state (i.e., geographical region) was fitted to the tree, as proposed in ref. 56. Entries in the resulting transition matrix were interpreted as likely geographical transfers between regions, drawn as arrows on the map in Figure 4 (directionality inferred from visual inspection of the phylogeny).

Temporal analysis. To investigate temporal signal in the maximum-likelihood phylogeny for S. Typhii, we used Path-O-Gen (see URLs) to extract two-time slices and analyzed their linear relationship with year of isolation using R. To assess the robustness of the H58 temporal signal, analysis of the H58 genome was repeated 100 times with randomly permuted tip dates. The evolutionary dynamics of the H58 lineage were investigated via Bayesian analysis with BEAST (v1.6). 35 Initial analyses were conducted on 114 isolates from across the H58 maximum-likelihood tree, covering the full temporal and geographical range of H58. In isozyme distribution was reduced by selecting a maximum of eight isolates from each geographical location, with as much temporal diversity as possible within that location. For countries represented by fewer than eight H58 isolates, all isolates at that location were included in the BEAST analysis. The concatenated SNP alignments of these 114 isolates were subjected to multiple BEAST analyses using both coalescent and skyline models of changes in population size, in combination with either a strict molecular clock or relaxed clock (uncorrelated lognormal distribution), to identify the model that best fit the data. The Skygrid model for the GTR+I substitution model was selected, and tip dates were defined as the year of isolation. For all model combinations, 3 independent chains of 10 million generations each were run to ensure convergence, with sampling every 1,200 iterations. The 3 runs were combined with a script in the Splits package when the script detected a significant overlap between the beginning and end of contigs.

Gene content analysis. The reads for each isolate were assembled de novo using the short-read assembler Velvet34 with parameters optimized with Velvet Optimizer (see URLs) to produce the highest N50 value. Contigs that were less than 300 bp long were excluded from further analysis. The assemblies were concatenated and annotated with Prokka 1.0 (the Pathogen Informatics team at the Wellcome Trust Sanger Institute, Cambridge, UK) using an automated pipeline.

The de novo-assembled contig sets were mapped iteratively to the penta-genome reference set (initially as the core genome of the 5. Typhii CT8).
Statistical methods. Simple, descriptive statistics were used to compare the geographical distributions of strains, prevalence of plasmids, and resistance genes and mutations and to calculate 95% confidence intervals. The significance of differences between studied groups of variables was calculated using Fisher’s exact test. All statistical tests were two-sided at α = 0.05, and analyses were performed using STAT (version 12.1, StatSoft Corp) and R.

74. Lertomi, R. & Bierne, N. Interactive Tree Of Life (iTOOL) is an online tool for phylogenetic tree display and annotation. Bioinformatics 23, 1271-1272 (2007).
An extended genotyping framework for *Salmonella enterica* serovar Typhi, the cause of human typhoid

Vanessa K. Wong1,2, Stephen Baker3,4,5, Thomas R. Connor5, Derek Pickard1, Andrew J. Page1, Jayshree Dave6, Niamh Murphy7, Richard Holliman7, Armine Setfot8, Michael Millar8, Zoe A. Dyson9,10, Gordon Dougan1, Kathryn E. Holt3,10, International Typhoid Consortium7

The population of *Salmonella enterica* serovar Typhi (S. Typhi), the causative agent of typhoid fever, exhibits limited DNA sequence variation, which complicates efforts to rationally discriminate individual isolates. Here we utilize data from whole-genome sequences (WGS) of nearly 2,000 isolates sourced from over 60 countries to generate a robust genotyping scheme that is phylogenetically informative and compatible with a range of assays. These data show that, with the exception of the rapidly disseminating H58 subclade (now designated genotype 4.3.1), the global S. Typhi population is highly structured and includes dozens of subclades that display geographical restriction. The genotyping approach presented here can be used to interrogate local S. Typhi populations and help identify recent introductions of S. Typhi into new or previously endemic locations, providing information on their likely geographical source. This approach can be used to classify clinical isolates and provides a universal framework for further experimental investigations.

1 The Wellcome Trust Sanger Institute, Hinxton, Cambridge CB10 1SA, UK. 2 Addenbrooke’s Hospital, Cambridge University Hospitals NHS Foundation Trust, Cambridge Biomedical Campus, Hills Road, Cambridge CB2 0QQ, UK. 3 The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam. 4 Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, Oxford University, Oxford, UK. 5 Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London WC1E 7HT, UK. 6 Cardiff University School of Biosciences, Cardiff University, Cardiff, UK. 7 Public Health Laboratory London, Public Health England, London, UK. 8 Division of Infection, Barts Health NHS Trust, London, UK. 9 Centre for Systems Genomics, University of Melbourne, Parkville, Victoria 3010, Australia. 10 Department of Biochemistry and Molecular Biology, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia. Correspondence and requests for materials should be addressed to VKW (email: vkw@sanger.ac.uk).

1A full list of consortium members appears at the end of the paper.
Typhoid fever (typhoid), caused by *Salmonella enterica* serovar Typhi (S. Typhi) bacteria, is a systemic human infection that affects an estimated 20.6 million people globally each year, causing an estimated 223,000 deaths. Typhoid remains endemic in populations with limited access to sanitation and safe water, and in a notifiable or reportable infection in many industrialized countries, where it is generally associated with travel to endemic areas. Public health laboratories have relied on techniques such as plasmotype typing or pulsed-field gel electrophoresis, which are phylogenetically naive and have limited discriminatory power to support epidemiological investigations and surveillance.

A genotyping scheme based on 88 single-nucleotide polymorphism (SNPs) identified within a limited set of genes was previously developed for *S. Typhi*. This enabled the classification of the *S. Typhi* population into 85 haplotypes (haplotype genotypes) based on biallelic profiles and provided the first phylogenetic framework for epidemiological studies. Subsequently, whole-genome sequencing (WGS) has been used to identify many more SNPs and other phylogenetically informative markers for discriminating within *S. Typhi*, which has limited genetic variation. Similar progress has been made in other monophyletic clades of bacterial pathogens, such as *Mycobacterium tuberculosis* and *Yersinia pestis*.

We have recently reported the WGS of almost 2,000 *S. Typhi* isolates sourced from 63 countries. This study identified >22,000 chromosomal SNPs in the core genome, which were used to build a comprehensive phylogenetic tree. Notably, the analysis confirmed the emerging dominance of the multidrug resistance-associated H58 clade, including the recent spread of H58 *S. Typhi* into Africa, confirming the value of SNP-based WGS analysis of *S. Typhi* to understand contemporary typhoid epidemiology. Here we utilize these WGS data to define a global population framework for *S. Typhi* and to define a new genotyping scheme comprising 68 SNPs that provides extensive coverage of typhoid-causing bacteria circulating globally. Given the increasing widespread adoption of WGS by public health laboratories for the tracking of bacterial pathogens, we further aimed to explore the utility of *S. Typhi* WGS data, analysed via genotyping, to predict the geographical source of travel-associated *S. Typhi* isolated in the United Kingdom. This approach gives greater discriminatory power and improved phylogenetic information than the earlier scheme, and forms a robust framework for public health surveillance, epidemiological investigations and laboratory experiments of typhoid.

Results

Defining phylogenetically informative genotypes for *S. Typhi*. In order to develop a comprehensive genotyping system, we used WGS data from >1,800 globally representative *S. Typhi* to identify phylogenetically informative clades and subclades based on SNP architecture. A summary of the isolates is shown in Table 1 and full details are provided in Supplementary Data 1 and Supplementary Table 1. Using a combination of phylogenetic tree topology and population genetic methods, we identified 16 *S. Typhi* clades that could be further divided into 49 subclades (Fig. 1, see Methods). Most of the clades could be grouped into four nested clusters (1–4), each with >100 bootstrap support and defined by >20 SNPs (coloured branches in Fig. 1a). The median pairwise distances between isolates were as follows: 25 SNPs within subclades, 169 SNPs within clades and 243 SNPs between clades. We labelled these primary clusters, clades and subclades using a structured hierarchical nomenclature system similar to that used for *M. tuberculosis*, whereby cluster 1 is subdivided into clades 1.1 and 1.2; clade 1.1 is further subdivided into subclades 1.1.1, 1.1.2, 1.1.3 and so on (see Fig. 1b, Methods).

Under the new genotype nomenclature, the globally disseminated multidrug-resistant clone commonly referred to as H58 (which actually includes haplotype H58 and eight other H58-derived haplotypes under the original Roumagruca et al. scheme), constitutes a single subclade (4.3.1). No other subclades were identified within clade 4.4. The CT18 reference genome (H1 in Roumagruca et al. scheme) belonged to subclade 3.2.1, while the laboratory strain Ty2 and its attenuated mutant BD9948 (H10 under the Roumagruca scheme) belonged to clade 4.11 (with no further differentiation to subclade level by BAPS). The backbone of the minimum spanning tree of Roumagruca et al. haplotypes was broadly consistent with the backbone structure of the whole-genome phylogeny (Supplementary Fig. 1a), however, mapping the Roumagruca haplotypes to the whole-genome phylogeny showed that the older scheme provides highly uneven resolution across the *S. Typhi* phylogeny (Supplementary Fig. 1b), with a lack of resolution in some cases (11 Roumagruca haplotypes span two or more distinct subclades each; for example, H52 comprises clades 3.4, 3.5, 4.1 and 4.2) and excessive resolution in others (24 subclades are further divided into two or more haplotypes in the Roumagruca scheme).

A new SNP-based genotyping framework for *S. Typhi*. We identified a minimum set of 68 SNPs that can be used to genotype *S. Typhi* into the four primary clusters, 16 clades and 49 subclades. For each of these groups, we identified all SNPs that were unique to members of the group, and selected one such SNP to be used for genotyping. We prioritized the inclusion of synonymous intragenic SNPs (that is, located within a protein-coding sequence, but with no change to the encoded amino acid), within genes that showed evidence of genetic stability within the *S. Typhi* population (that is, nucleotide diversity <1% and dN/dS <0.7 across the global data set, with no inactivating mutations identified). Details of the genotyping SNPs are given in Supplementary Table 2. This genotyping scheme has greater discriminatory power than the original Roumagruca haplotyping scheme (D = 0.26 versus 0.58), is phylogenetically informative by design and the hierarchical nomenclature of genotypes is taxonomically informative with respect to phylogenetic relationships between clades and subclades.

Geographical distribution of *S. Typhi* clades and subclades. Next, we examined the geographical distribution of *S. Typhi* genotypes. For these analyses, isolates of the same subclade, country and year were collapsed to a single representative to reduce the impact of localized outbreaks on our collection; this resulted in 541 unique isolates for analysis. Primary clusters 2, 3, and 4 were broadly distributed across continents (greens, blues and reds, respectively, in Fig. 1c), likely reflecting the relatively ancient spread of *S. Typhi* across the globe. Isolates outside these clusters, which result from deep branching closer to the root of the *S. Typhi* whole-genome tree, were rare in our collection (n = 24 unique isolates) and mostly found in Africa (n = 16). While the three common clusters (2–4) were present in most regions we analysed, cluster 2 predominated among American isolates (n = 18/23 unique isolates, 78%). Most clades were detected on multiple continents (n = 10/16) and included isolates from Asia (n = 10/16) and/or Africa (n = 10/10), which together
<table>
<thead>
<tr>
<th>Continent (region)</th>
<th>Country of origin (n≥5)</th>
<th>Range of isolation dates (years)</th>
<th>Number of isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asia</td>
<td>Cambodia</td>
<td>1976-2012</td>
<td>106 (54.3%)</td>
</tr>
<tr>
<td></td>
<td>Indonesia</td>
<td>2007-2012</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>Laos</td>
<td>1976-2012</td>
<td>129</td>
</tr>
<tr>
<td></td>
<td>Malaysia</td>
<td>1972-2011</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td>Vietnam</td>
<td>2000-2010</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>Other (Philippines, Thailand, East Timor, Myanmar)</td>
<td>2005-2011</td>
<td>21</td>
</tr>
<tr>
<td>South Asia</td>
<td>Bangladesh</td>
<td>1977-2012</td>
<td>321</td>
</tr>
<tr>
<td></td>
<td>India</td>
<td>1977-2012</td>
<td>318</td>
</tr>
<tr>
<td></td>
<td>Nepal</td>
<td>1999-2012</td>
<td>211</td>
</tr>
<tr>
<td></td>
<td>Pakistan</td>
<td>2006-2012</td>
<td>46</td>
</tr>
<tr>
<td></td>
<td>Other (Sri Lanka, Afghanistan)</td>
<td>2005-2012</td>
<td>3</td>
</tr>
<tr>
<td>Western Asia</td>
<td>Iraq</td>
<td>1997-2011</td>
<td>25</td>
</tr>
<tr>
<td></td>
<td>Lebanon</td>
<td>2009-2011</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Other (Armenia, Palestine, Turkey, Western Asia)</td>
<td>2001-2011</td>
<td>7</td>
</tr>
<tr>
<td>Eastern Asia</td>
<td>Other (China)</td>
<td>2002-2011</td>
<td>3</td>
</tr>
<tr>
<td>Africa</td>
<td></td>
<td></td>
<td>374 (20.4%)</td>
</tr>
<tr>
<td>North Africa</td>
<td>Algeria</td>
<td>1961-2009</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Morocco</td>
<td>1969-2009</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Other (Sudan, Egypt, Tunisia)</td>
<td>1961-2008</td>
<td>9</td>
</tr>
<tr>
<td>East Africa</td>
<td>Kenya</td>
<td>1980-2010</td>
<td>115</td>
</tr>
<tr>
<td></td>
<td>Tanzania</td>
<td>2004-2010</td>
<td>52</td>
</tr>
<tr>
<td></td>
<td>Other (Comoros, Madagascar)</td>
<td>1960-2000</td>
<td>7</td>
</tr>
<tr>
<td>Central Africa</td>
<td>Cameroon</td>
<td>1959-2011</td>
<td>49</td>
</tr>
<tr>
<td></td>
<td>DRC</td>
<td>1979-2011</td>
<td>27</td>
</tr>
<tr>
<td></td>
<td>Other (Angola, Central African Republic)</td>
<td>2001-2009</td>
<td>5</td>
</tr>
<tr>
<td>West Africa</td>
<td>Other (Burkina Faso, Cape Verde, Benin, Guinea, Ivory Coast, Gabon, Liberia, Mali, Niger, Nigeria, Mauritania, Senegal, Togo)</td>
<td>1999-2009</td>
<td>30</td>
</tr>
<tr>
<td>Southern Africa</td>
<td>Malawi</td>
<td>2004-2013</td>
<td>153</td>
</tr>
<tr>
<td></td>
<td>South Africa</td>
<td>2004-2012</td>
<td>112</td>
</tr>
<tr>
<td></td>
<td>Other (Gabon, Democratic Republic of Congo)</td>
<td>2009-2012</td>
<td>41</td>
</tr>
<tr>
<td>Africa</td>
<td>Unknown</td>
<td>2009-2012</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>302 (16.2%)</td>
</tr>
<tr>
<td>Europe</td>
<td>Other (Russia)</td>
<td>1961-1986</td>
<td>7 (0.4%)</td>
</tr>
<tr>
<td></td>
<td>Western Europe</td>
<td>1961-1996</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Other (France (suspected African origin of infection))</td>
<td>2009</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Other (Malta)</td>
<td>2009</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>342 (18.7%)</td>
</tr>
<tr>
<td>Australia and Oceania</td>
<td>Australia</td>
<td>2010-2012</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Oceania</td>
<td>2010-2012</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1980-2012</td>
<td>339</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1980-2012</td>
<td>170</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1980-2012</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1980-2012</td>
<td>47</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1980-2012</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>North America</td>
<td>Other (USA, Mexico)</td>
<td>1959-2011</td>
</tr>
<tr>
<td></td>
<td>Other (El Salvador)</td>
<td>1959-2011</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1959-2011</td>
<td>1 (0.05%)</td>
</tr>
<tr>
<td></td>
<td>Central America</td>
<td>Other (El Salvador)</td>
<td>2012</td>
</tr>
<tr>
<td></td>
<td>South America</td>
<td>Other (French Guiana, Peru, South America)</td>
<td>2003-2012</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1939-2012</td>
<td>10 (0.9%)</td>
</tr>
<tr>
<td>Unknown origin</td>
<td></td>
<td></td>
<td>1939-2012</td>
</tr>
</tbody>
</table>

5. Typh, Salmonella enterica serovar Typhi.
6. Typhi isolates from the global collection, which were used to define genotypes. Countries with fewer than five isolates were grouped into the category “Other” and indicate the number of such countries in each region.

Nature Communications | DOI: 10.1038/ncomms12827 | www.nature.com/naturecommunications

297
Figure 1 | Population structure of S. Typhi based on genome-wide SNPs. (a) Whole-genome tree of 1,631 global S. Typhi isolates. Primary clusters 1-4 are indicated in the outer coloured ring; branches defining these groups are coloured in the tree. These groups are further divided into clades, which are shaded and labelled. The location of S. Typhi reference genomes CT18 (accession number A1531382) and Ty2 (accession number AE016623) are indicated on the tree. Subclade 4,3,1 (H58, marked in red), which comprises half of the global collection, is represented by 850 (60%) randomly selected isolates out of the total 1,631 belonging to this subclade, so that the relationships between other clades can be visualised. (b) Tree backbone showing further division of 16 S. Typhi clades (shaded) into 49 subclades (labelled; note 12 undifferentiated clade groups shown in brackets). Branches are coloured by primary clade. (c) Map of the world showing subclade diversity of S. Typhi isolates in the global collection, by subregion. Where groups of isolates from the same country and year belonged to the same subclade, this was classified as an ‘outbreak’ and the group is only represented once in the pie graphs. Pie slices are sized to indicate number of isolates; slices are coloured by clade; multiple slices of the same colour indicate multiple subclades belonging to the same clade.
Figure 2 | Geographical clustering of S. Typhi subclades. Heatmap shows, for each subclade, the percentage of unique isolates originating from each of the geographical regions. Where groups of isolates from the same country and year belonged to the same subclade, this was classified as an ‘outbreak’ and the group is only represented once. The same data are represented as a scaled bar graph to the right. The full list of isolates by country and subclade is provided in Supplementary Data 1.

made up 78% of our isolate collection (Table 1). However, there were differences in the geographic distributions of clades, with most clades being dominated by unique isolates from a single continent (Asia, Africa or Oceania; see Supplementary Fig. 2).

In contrast, at the subclade level, only 22% of subclades (n = 11) were found on more than one continent, and most were dominated by unique isolates from a single country or region: 40 subclades (82%) had ≥50% of non-outbreak isolates from a single country (Fig. 2) and 44 subclades (90%) had ≥50% of non-outbreak isolates from a single region (Fig. 2). A total of 28 subclades comprised five or more non-outbreak isolates each, and of these common subclades, 12 (43%) were detected in a single region only (six in Oceania, five in Southeast Asia and one in South Asia; Fig. 2). In total, 16 common subclades (57%) were highly restricted to a region (>90% of isolates drawn from a single region) and 20 (71%) were generally associated with one region (>70% of isolates drawn from a single region; Fig. 2). These data suggest that most S. Typhi subclades represent localized bacterial subpopulations with barriers to geographical dispersion, and that transfers to new locations rarely result in long-term establishment of local populations. In contrast with this general pattern, subclade 4.3.1 (previously H58) was found in nine different regions across Africa, Asia and Oceania. Only 10 other subclades (20%) were found on more than one continent, and the majority of these were dominated either by Asian, African or Oceanian isolates (Fig. 2). Thus, the recent global dissemination of subclade 4.3.1, which spread out of South Asia ~30 years ago and has established successful local clonal
expansions in dozens of countries, likely represents a comparatively rare event in the evolutionary history of S. Typhi.

Genomic prediction of the geographical origins of S. Typhi by comparison with the global framework. Since most S. Typhi subclades were associated with a narrow geographical source, we hypothesized that genotyping of S. Typhi isolates could be used to predict the likely geographical origins of typhoid cases. As this is clearly challenging for the more widely distributed subclades, we also sought to examine whether specific SNPs could be used to predict origins down to the country level. For 1,201 out of 1,183 (99%) isolates in our global collection, the genetically closest isolate was from the same country. Where the closest isolate was zero SNPs away, this frequency was 99% and for <10 SNPs, 90% (Supplementary Fig. 3).

Since our current global genome collection includes groups of isolates that were frequently collected from the same time and place, this should not be taken as a reliable measure of the general predictive power of SNP distance for S. Typhi. In order to further explore the power of our global genomic framework to predict geographic origins of travel-associated typhoid, we sequenced and genotyped 99 novel S. Typhi that were isolated from patients attending a hospital in East London, United Kingdom between 2005 and 2010 (Table 2). A total of 13 genotypes were identified. Epidemiological interviews were able to link 81 of these cases with travel to a specific country; the remaining 18 cases were not associated with travel. The median SNP distance between these novel isolates and genomes in our global collection was 21 SNPs (interquartile range 18–25 SNPs), posing a challenge for prediction of their geographical origin. Among the 81 travel-associated UK isolates, 53 were genotyped as 4.3.1; these were all linked to travel to countries within South Asia (Fig. 3), and clustered along with South Asian isolates from the global collection (Fig. 3 and Supplementary Fig. 4n). For the 28 non-4.3.1 travel-associated UK isolates, the location of travel generally matched the geographical origin of the closest isolate (in terms of number of SNPs) in the global collection: travel location and closest global isolate source matched at the region level in all cases, and the country level in most cases (n=26, 71%). This is prediction of geographical origin based on the

### Table 2 | Summary of 99 East London travel-associated S. Typhi isolates used in the study.

<table>
<thead>
<tr>
<th>Country of origin</th>
<th>Range of isolation dates (year)</th>
<th>Number of isolates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>2006–2012</td>
<td>38</td>
</tr>
<tr>
<td>India</td>
<td>2006–2012</td>
<td>22</td>
</tr>
<tr>
<td>Pakistan</td>
<td>2006–2012</td>
<td>13</td>
</tr>
<tr>
<td>Nepal</td>
<td>2007</td>
<td>1</td>
</tr>
<tr>
<td>India/Pakistan</td>
<td>2008</td>
<td>1</td>
</tr>
<tr>
<td>India/Kashmir</td>
<td>2008</td>
<td>2</td>
</tr>
<tr>
<td>Bangladesh/India</td>
<td>2010</td>
<td>2</td>
</tr>
<tr>
<td>Nigeria</td>
<td>2009</td>
<td>1</td>
</tr>
<tr>
<td>Ghana</td>
<td>2007</td>
<td>1</td>
</tr>
<tr>
<td>No known travel</td>
<td>2005–2011</td>
<td>18</td>
</tr>
<tr>
<td>Total</td>
<td>2005–2012</td>
<td>99</td>
</tr>
</tbody>
</table>

S. Typhi, Salmonella enterica serovar Typhi.

The county of origin and range of isolation dates (year) for the isolates are described. The five isolates, matching countries of travel were resequenced, and it was not possible to confirm in which country the S. Typhi infection occurred.

### Table 3 | Summary of genotyping and SNP results for travel-associated S. Typhi isolates with known country of travel.

<table>
<thead>
<tr>
<th>Subclone</th>
<th>Country of travel</th>
<th>N (travel)</th>
<th>Closest genome country match</th>
<th>N (global)</th>
<th>Region frequencies (excluding match)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.0.1</td>
<td>Bangladesh</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>South Asia (100%)</td>
</tr>
<tr>
<td>2.1.7</td>
<td>India</td>
<td>2</td>
<td>2</td>
<td>49</td>
<td>Oceania (92%)</td>
</tr>
<tr>
<td>2.2.0</td>
<td>Pakistan</td>
<td>1</td>
<td>1</td>
<td>17</td>
<td>Southern Africa (24%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>East Africa (24%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>South Asia (18%)</td>
</tr>
<tr>
<td>2.2.2</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>9</td>
<td>South Asia (33%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Western Asia (33%)</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Bangladesh</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>South America (55%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>South Asia (33%)</td>
</tr>
<tr>
<td>3.1.1</td>
<td>Nigeria, Ghana</td>
<td>2</td>
<td>0</td>
<td>25</td>
<td>West Africa (68%)</td>
</tr>
<tr>
<td>3.2.2</td>
<td>Bangladesh, Pakistan</td>
<td>3</td>
<td>2</td>
<td>12</td>
<td>South Asia (83%)</td>
</tr>
<tr>
<td>3.3.0</td>
<td>Bangladesh, Pakistan, India</td>
<td>14</td>
<td>13</td>
<td>10</td>
<td>South Asia (83%)</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Pakistan</td>
<td>1</td>
<td>1</td>
<td>32</td>
<td>East Africa (44%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>South Asia (19%)</td>
</tr>
<tr>
<td>4.1.0</td>
<td>India</td>
<td>1</td>
<td>1</td>
<td>78</td>
<td>Southeast Asia (71%)</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Bangladesh, India, Pakistan, Nepal, Kuwait</td>
<td>53</td>
<td>—</td>
<td>853</td>
<td>Southeast Asia (50%)</td>
</tr>
</tbody>
</table>

Closest genome country match, number of travel-associated isolates whose country of travel matched that of the closest genome in the global collection (i.e. lowest number of SNPs). N (global), number of isolates in the global collection that were assigned to this subclone. N (travel), number of travel-associated isolates that were assigned to this subclone. Region frequencies, frequency of each geographic region assignment of the subclone from the global collection (note groups of isolates from the same subclone, country and year were classified as isolates and represented differently per group in the frequency calculations). N (global) = 99. *Author enterica serovar Typhi.

**Highlights the most frequent region for this subclone among the global collection, where this matches the region of travel.**
closest strain of known location in the current global framework would have yielded the correct region of origin in all cases, and the correct country of origin in 71% of cases (95% confidence interval (CI), 66–76%). Furthermore, for non-4.3.1 subclades, genotyping alone was predictive of geographical origin at the regional level for the same proportion of isolates (71%).

It is likely that power to predict the geographical sources of UK isolates would be improved by wider geographical coverage in the reference genome collection. Two of these isolates were genotyped as subclade 3.1.1 and linked with travel to Ghana and Nigeria; the closest isolates in our global collection were 16–17 SNPs away and were not from these precise locations, but likely originated from bordering countries in West Africa (Supplementary Fig. 4c). It is likely that a deeper coverage of West African isolates in our global framework would provide greater power to resolve geographic associations within this region, which comprised less than 2% of our current global collection (n = 39 isolates). Similarly, for the other travel-associated isolates for which the recorded country of travel did not match the closest genome in the global tree, the closest genome was also from a neighbouring country (for example, Pakistan, India, Bangladesh, see Supplementary Fig. 4).

Genomic predictions of the geographical origins of 18 non-travel-associated UK isolates are shown in Table 4 and Supplementary Fig. 4. Thirteen isolates were 4.3.1 and clustered together with travel-associated isolates from South Asia, within a broader group of South Asian 4.3.1 isolates (Fig. 3). This suggests that S. Typhi imported into the United Kingdom from these regions have likely been transmitted onwards within the United Kingdom to individuals with no recent travel history (Supplementary Fig. 4n). Two additional isolates were from subclades that were dominated by a single region in our global collection—3.3.1 (60% West Africa) and 3.3.0 (83% South Asia). Notably, while the 4.3.1 isolates were closely related to travel-associated isolates recently obtained in London, they were ≥17 SNPs away from any isolates in the global collection. Thus, the diversity captured by the global collection does not provide the resolution to precisely identify the origin of these isolates23.

Discussion
Our data show that the global S. Typhi population consists of 49 distinct subclades that are strongly geographically clustered, with many locations harbouring subpopulations of S. Typhi established over long periods of time. We show how these
Table 4 | Summary of genotyping and SNP results for travel-associated S. Typhi isolates of unknown origin.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Country of closest SNP match</th>
<th>Distance (SNPs)</th>
<th>Subclade</th>
<th>Subclade distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td>H001424126</td>
<td>Mexico</td>
<td>146</td>
<td>2.02</td>
<td>North America (50%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>North Africa (50%)</td>
</tr>
<tr>
<td>H00156550</td>
<td>Pakistan</td>
<td>92</td>
<td>3.01</td>
<td>South Asia (50%)</td>
</tr>
<tr>
<td>H00272442</td>
<td>Ghana</td>
<td>17</td>
<td>3.11</td>
<td>*West Africa (66%)</td>
</tr>
<tr>
<td>H00716223</td>
<td>Bangladesh</td>
<td>11</td>
<td>3.0</td>
<td>*South Asia (83%)</td>
</tr>
<tr>
<td>H0108335</td>
<td>India</td>
<td>18</td>
<td>3.31</td>
<td>East Africa (44%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>South Asia (19%)</td>
</tr>
<tr>
<td>H01046338</td>
<td>Bangladesh1</td>
<td>13</td>
<td>4.31</td>
<td>Southeast Asia (50%)</td>
</tr>
<tr>
<td>H05006403</td>
<td>Bangladesh2</td>
<td>4</td>
<td></td>
<td>South Asia (26%)</td>
</tr>
<tr>
<td>H06136379</td>
<td>Bangladesh3</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0603830</td>
<td>Bangladesh4</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H05182620</td>
<td>Bangladesh5; India1</td>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0103404</td>
<td>India1</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0194694</td>
<td>India2</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0507484</td>
<td>India3</td>
<td>9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H0519507</td>
<td>India4</td>
<td>14</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H05212220</td>
<td>India5</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H03129553</td>
<td>Pakistan1</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H02020316</td>
<td>Pakistan2</td>
<td>10</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SNP, single nucleotide polymorphism; S. Typhi, Salmonella enterica serovar Typhi.
For each isolation index, the closest isolate in the global collection was determined using a smallest SNP distance, that is, smallest number of core genome SNPs; the country and SNP distance are reported. *West Africa includes the travel-frequent regions for this subclade. **South Asia includes the travel-frequent regions for this subclade. ***Southeast Asia includes the travel-frequent regions for this subclade.

subclades can be identified through a simple genotyping scheme consisting of 68 SNPs. Importantly, while we show that this scheme is highly phylogenetically informative, it can be readily inferred from raw sequence data without the need for multiple genome comparisons, phylogenetic analysis or any other complex or computationally intensive steps. Such properties make this universal SNP-based system a valuable tool upon which researchers can develop future studies. The S. Typhi genome is highly stable and exhibits minimal genetic variation and virtually no recombination12,14, and we recently estimated the substitution rate to be lower than one SNP per genome per year15; therefore, the genotyping framework is expected to be robust to future evolution. According to the strong geographical clustering of the various subclades, whole-genome comparison of novel S. Typhi isolates to the existing global population framework is strongly predictive of geographic origin at the regional level and has the potential to accurately predict origins to the country level. This has important public health implications for typhoid surveillance and control in endemic and non-endemic areas; however, ongoing updates to the global genomic framework will be important to ensure the utility of genomic surveillance for typhoid. For example, we found that the origin of travel-associated 4.31 isolates could not be resolved using the prior global framework alone, but benefited from updated information provided by other recent travel-associated isolates of known geographical origin. This illustrates the importance of expanding and updating the global genomic framework through sequencing of novel isolates and suggests that, while ongoing surveillance in endemic areas is undoubtedly important, the use of clinically well-characterized travel-associated strains isolated in non-endemic countries may also provide a valuable source for improving the granularity of data in the framework for genome-based surveillance of S. Typhi12. In addition, it will be important to expand the current global framework to include more recent isolates (the most recent in our current collection was from 2013) as well as isolates from regions that are currently underrepresented (including Africa, the Americas and northeast Asia).

WGS-equipped reference laboratories provide a highly accessible source to expand the global genomic framework for typhoid, with potential benefit to local but also global typhoid control. For example, in England, Wales and Northern Ireland ~520 typhoid cases are reported annually to the national reference laboratory (Public Health England). These cases are investigated in order to determine whether they are associated with travel to typhoid endemic regions15. However, approximately one-fifth of typhoid cases in the United Kingdom cannot be traced to a country of origin. At present, Public Health England provides molecular typing, which since April 2015 includes WGS as well as antimicrobial susceptibility profiling, for S. Typhi isolates from such cases. The resulting data are considered important for local epidemiology. However, we propose that this could also serve as a proxy for informal surveillance of typhoid molecular epidemiology in endemic regions. This may prove particularly valuable when supported by our genotyping framework for simplified attribution.

Methods

Bacterial isolates and WGS. A total of 1,830 S. Typhi isolates were analysed in this study (Supplementary Data 1), including a collection of 1,693 globally distributed isolates contributed by members of the International Typhoid Consortium10 and 99 novel S. Typhi isolates from East London, UK. S. Typhi comprising the global collection were isolated between 1965 and 2013 and originate from 65 countries spanning six continents (Asia, Africa, North and South America, Europe, and Australia and Oceania) as previously described12. An additional 99 novel S. Typhi isolates were obtained from returning travellers with a definite illness who presented at The Royal London Hospital, Barts Health NHS Trust in East London, UK, between 2005 and 2013. Travel history, available for all of the isolates, included visits to seven countries within the continents of Asia and Africa. DNA was extracted using the Wizard Genomic DNA Kit (Promega, Madison, WI, USA) as per the manufacturer’s instructions. Index-tagged paired-end Illumina sequencing libraries were prepared as previously described12. These were combined into pools, each containing 40 uniquely tagged libraries, and were sequenced on the Illumina HiSeq2500 platform (Illumina, San
Subclade 4.3.3, which is the only subclade of Clade 4.3, corresponds to the group referred to as 4.3B, based on the haplotype scheme of Rosengrant et al. in which it is defined by the presence of a single SNP gylt-C160CT (position 24,880,62) in S. Typhimurium BF05 (ref. 73). BAPS clustering at any level could not further subdivide subclade 4.3.3 (1012).

SNP-based genotyping. We identified a minimum set of 68 SNPs with which to rapidly genotype S. Typhi from the 14 clades and 48 subclades, as described above (Supplementary Table 2). Short read alignment (BAM) files, generated by mapping Illumina reads to the CT18 reference genome (accession AE533392), were used to assign genotypes for each novel read set using a custom Python script (available at https://github.com/katohgenotyping). Briefly, the script uses samtools mpileup to extract from each BAM the consensus base calls at the SNP loci. The resulting variant call format file is then processed to identify the presence of clades, clade and subclade-defining SNP alleles (defined in Supplementary Table 2) that pass a minimum quality threshold (default consensus base Phred score ≥ 20) and uses these to assign the read to a specific clade and subclade. Discriminatory power was calculated using the method outlined in ref 35.

Data availability. Raw sequence data are available in the European Nucleotide Archive under accession ERP001718. Supplementary Data 1 lists accession numbers for each isolate. The software for Microtree interactive tree viewer is available at https://microtree.org/project/microtree/. SMALT is available at https://www.angeron.com/resources/software/smalt/. Python script to visualize and annotate trees is available at https://github.com/katohgenotyping. Python script to call SNPs is downloadable at https://github.com/katohgenotyping.

References
Members of the International Typhoid Consortium

Julian Parkhill1, Nicholas A Reaney2, Robert A Kingsley1,2, Nicholas R Thomson1,3,4, Jacqueline A Keane1, François-Xavier Weill5,6, Simon Le Hello7,8, Jane Hawkey9,10,14, David J Edwards6,9,10, Simon R Harris1, Amy K Cain1, James Hadfield10, Peter J Hart5, Nga Tran Vu Thieu11, Elizabeth J Klemm1, Robert F Breiman12,13,18, Conall H Watson12, W John Edmunds12, Samuel Kariuki10,16, Melita A Gordon2,21, Robert S Heyderman2,23,26, Chinyere Okoro1,2, Jan Jacobs23,24, Octavie Lunguya25,26, Chisomo Mselulu21,27, Jose A Chabalgoity28, Mike Kama29, Kylie Jenkins30, Shanta Dutta30, Florian Marks32, Josifina Campos33, Corinne Thompson3,4.

1. Wellcome Trust Sanger Institute, Cambridge, UK.
3. Cluster for Research Excellence in Infectious Diseases, Department of Medicine, University of Oxford, Oxford, UK.
5. Gastroenterology and Hepatology, Institut National de Recherches Médicales, Villejuif, France.
7. Department of Laboratory Medicine and Infectious Diseases, University of Oxford, Oxford, UK.
8. Service de Pathologie Infectieuse et Parasitaire, Université Libre de Bruxelles, Brussels, Belgium.
9. Department of Infectious and Parasitic Diseases, Faculty of Veterinary Medicine, University of Liège, Liège, Belgium.
10. Department of Infectious Diseases, Royal Free Hospital, London, UK.
11. Department of Tropical Medicine, Mc Gill University, Montreal, Canada.
12. Department of Medical Microbiology, University of Cambridge, Cambridge, UK.
13. Department of Pathology, University of Cambridge, Cambridge, UK.
15. Department of Microbiology and Immunology, University of Oxford, Oxford, UK.
17. Department of Microbiology, University of Oxford, Oxford, UK.
18. Department of Microbiology, University of Oxford, Oxford, UK.
19. Department of Microbiology, University of Oxford, Oxford, UK.
20. Department of Microbiology, University of Oxford, Oxford, UK.
22. Department of Microbiology, University of Oxford, Oxford, UK.
23. Department of Microbiology, University of Oxford, Oxford, UK.
24. Department of Microbiology, University of Oxford, Oxford, UK.
25. Department of Microbiology, University of Oxford, Oxford, UK.
27. Department of Microbiology, University of Oxford, Oxford, UK.
29. Department of Microbiology, University of Oxford, Oxford, UK.
30. Department of Microbiology, University of Oxford, Oxford, UK.
31. Department of Microbiology, University of Oxford, Oxford, UK.
32. Department of Microbiology, University of Oxford, Oxford, UK.
33. Department of Microbiology, University of Oxford, Oxford, UK.
34. Department of Microbiology, University of Oxford, Oxford, UK.
35. Department of Microbiology, University of Oxford, Oxford, UK.
36. Department of Microbiology, University of Oxford, Oxford, UK.
37. Department of Microbiology, University of Oxford, Oxford, UK.

1Liverpool School of Tropical Medicine, Pembroke Place, Liverpool L3 5QA, UK; 2Institute of Food Research, Norwich Research Park, Colney, Norwich NR4 7UA, UK; 3Institut Pasteur, Unite des Bacteries Pathogènes Entérales, Paris, France; 4Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville, Victoria 3052, Australia; 5Institute of Biomedical Research, School of Immunity and Infection, College of Medicine and Dental Sciences, University of Birmingham, Birmingham, UK; 6Kenya Medical Research Institute (KEMRI), PO Box 43640-00100, Nairobi, Kenya; 7Centers for Disease Control and Prevention, 1600 Clifton Road, Atlanta, Georgia 30329-4027, USA; 8Emory Global Health Institute, 1599 Clifton Road, NE, 1999-001-1AH, Atlanta, Georgia 30322, USA; 9Centre for the Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; 10Institute of Infection and Global Health, University of Liverpool L69 7BE, UK; 11Malawi-Liverpool Wellcome Trust Clinical Research Programme, College of Medicine, University of Malawi, PO Box 30096, Chichiri, Blantyre 3, Malawi; 12Division of Infection and Immunity, University College London, London, UK; 13Department of Clinical Sciences, Institute of Tropical Medicine, Antwerp, Belgium; 14KU Leuven, University of Leuven, Department of Microbiology and Immunology, Belgium; 15National Institute for Biomedical Research, Kishanpur, Democratic Republic of the Congo; 16University Hospital of Kinshasa, Kinshasa, Democratic Republic of the Congo; 17Microbiology Department, College of Medicine, University of Malawi, Malawi; 18Instituto de Desarrollo Biotecnologico, Instituto de Higiene, Facultad de Medicina, Avenida Navarro 1301, Montevideo, CP 11600, Uruguay; 19Ministry of Health, Toorak, PO Box 2223, Suva, Fiji; 20Health Sector Support Program, PO Box 52986, Suva, Fiji; 21National Institute of Cholera and Enteric Diseases, 7-33 CIT Road, Scheme XIB, Bellaghat, Kolkata 700 010, India; 22International Vaccine Institute, Department of Epidemiology, Immunovac PO Box 14, Seoul 151-600, Republic of Korea; 23Enteric Pathogen Division, ANLS-Coles G Mulford Institute, CASA, Argentina; 24Division of Pediatric Infectious Diseases, University of Nebraska Medical Center, Omaha, NE 68198-3897, USA; 25University of Abuja Teaching Hospital, Gwagwalada, FCT, Nigeria; 26Bingham University, Karu, Nassarawa State, Nigeria; 27The Jenner Institute, Nuffield Department of Medicine, University of Oxford, Oxford, UK; 28Centre for Enteric Diseases, National Institute for Communicable Diseases, Division in the National Health Laboratory Service and Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; 29Department of Clinical Research, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; 30Graduate School of Tropical Medicine and Global Health, Nagasaki University, Nagasaki, Japan; 31School of Health Sciences, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Kathmandu, Nepal; 32Murdoch Childrens Research Institute, Melbourne, Australia; 33Enteric and Lymphopenia Reference Laboratory, Institute of Environmental Science and Research Limited (ESR), New Zealand; 34National Centre for Biosecurity and Infectious Disease, Institute of Environmental Science and Research, Pori, New Zealand; 35Semarang Ministry of Health, Aja, Semarang, Indonesia; 36National Influenza Centre, World Health Organization, Centre for Communicable Disease Control, Yoyogi, Fukuoka, Japan; 37Department of Microbiology, Hasanuddin University, Makassar, Indonesia; 38Oxford-AstraZeneca Hospital-Welwyn Trust Research Unit, Microbiology Laboratory, Welshed Hospital, Vientiane, Lao People's Democratic Republic; 39National Health Services, Tsukasa Tianumaei Hospital, Semarang, Indonesia; 40Mahidol-Oxford Tropical Medicine Research Unit, Faculty of Tropical Medicine, Mahidol University, Bangkok, Thailand; 41Paediatric Emergency Medicine, Chelsea and Westminster Hospital, London, UK; 42Centre for International Health, University of Oslo, PO Box 56, Blindern 0316, Norway; 43Sambodrila Reference Service, Public Health England, Cirencester, London N9V 9Q UK; 44Emerging Disease Surveillance and Response, Division of Pacific Technical Support, World Health Organization, PO Box 113, Suva, Fiji; 45Cambodia-Oxford Medical Research Unit, Angkor Hospital for Children, Siem Reap, Cambodia; 46Medical Laboratory Diagnostic Unit - Public Health Laboratory, the Peter Ureth Institute for Infection and Immunity, The University of Melbourne, Victoria 3010, Australia; 47Department of Veterinary Medicine, University of Cambridge, Cambridge, UK; 48St. Augustine's Hospital, Muháea, Tanzania.
Whole Genome Sequence Analysis of *Salmonella* Typhi Isolated in Thailand before and after the Introduction of a National Immunization Program

Zoe A. Dyson1,†,‡, Duy Pham Thanh3, Ladaporn Bodhidatta5, Carl Jefries Mason6, Apichai Srijan4, Maia A. Rabaa4,5, Phat Voong Vinh7, Tuyen Ha Thanh7, Guy E. Thrailikes2,6,7, Stephen Baker2,6,8, Kathryn E. Holt1,9,‡

1 Centre for Systems Genomics, University of Melbourne, Parkville, Victoria, Australia; 2 Department of Biochemistry and Molecular Biology, Bca9 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria, Australia; 3 The Hospital for Tropical Diseases, Wellcome Trust Major Overseas Programme, Oxford University Clinical Research Unit, Ho Chi Minh City, Vietnam; 4 Department of Enteric Diseases, Armed Forces Research Institute of Medical Sciences, Bangkok, Thailand; 5 Centre for Tropical Medicine and Global Health, Oxford University, Oxford, United Kingdom; 6 The London School of Hygiene and Tropical Medicine, London, United Kingdom.

† These authors are joint senior authors on this work.

* khol@unimelb.edu.au (KEH), zoe.dyson@unimelb.edu.au (ZAD)

Abstract

Vaccines against *Salmonella* Typhi, the causative agent of typhoid fever, are commonly used by travellers, however, there are few examples of national immunization programs in endemic areas. There is therefore a paucity of data on the impact of typhoid immunization programs on localised populations of *S. Typhi*. Here we have used whole genome sequencing (WGS) to characterise 44 historical bacterial isolates collected before and after a national typhoid immunization program that was implemented in Thailand in 1977 in response to a large outbreak; the program was highly effective in reducing typhoid case numbers. Thai isolates were highly diverse, including 10 distinct phylogenetic lineages or genotypes. Novel prophage and plasmids were also detected, including examples that were previously only reported in Shigella sonnei and *Escherichia coli*. The majority of *S. Typhi* genotypes observed prior to the immunization program were not observed following it. Post-vaccine era isolates were more closely related to *S. Typhi* isolated from neighbouring countries than to earlier Thai isolates, providing evidence for the local persistence of endemic *S. Typhi* following the national immunization program. Rather, later cases of typhoid appeared to be caused by the occasional importation of common genotypes from neighbouring Vietnam, Laos, and Cambodia. These data show the value of WGS in understanding the impacts of vaccination on pathogen populations and provide support for the proposal that large-scale typhoid immunization programs in endemic areas could result in lasting local disease elimination, although larger prospective studies are needed to test this directly.
Author Summary

Typhoid fever is a systemic infection caused by the bacterium Salmonella Typhi. Typhoid fever is associated with inadequate hygiene in low-income settings and a lack of sanitation infrastructure. A sustained outbreak of typhoid fever occurred in Thailand in the 1970s, which peaked in 1975–1976. In response to this typhoid fever outbreak the government of Thailand initiated an immunization program, which resulted in a dramatic reduction in the number of typhoid cases in Thailand. To better understand the population of S. Typhi circulating in Thailand at this time, as well as the impact of the immunization program on the pathogen population, we sequenced the genomes of 44 S. Typhi obtained from hospitals in Thailand before and after the immunization program. The genome sequences showed that isolates of S. Typhi bacteria isolated from post-immunization era typhoid cases were likely imported from neighbouring countries, rather than strains that have persisted in Thailand throughout the immunization period. Our work provides the first historical insights into S. Typhi in Thailand during the 1970s, and provides a model for the impact of immunization on S. Typhi populations.

Introduction

Salmonella enterica subspecies enterica serovar Typhi (S. Typhi) is a human restricted bacterial pathogen and the etiological agent of typhoid fever. S. Typhi is transmitted feco-orally and can establish asymptomatic carriage in a small subset of an exposed population [1]. Recent estimates [2–5] place the global burden of typhoid fever at 25–30 million cases annually, of which 200,000 are associated with deaths. Typhoid fever occurs most commonly in industrialising countries, specifically in locations with limited sanitation and related infrastructure [3]; children and young adults are among the most vulnerable populations in these settings [6–8]. Antimicrobial therapy together with water sanitation and hygiene (WASH) interventions are the major mechanisms by which typhoid fever is controlled [5, 10]. However, none of these approaches is optimal and resistance against antimicrobials has become increasingly common in S. Typhi since the 1970s [11–13]. A number of typhoid vaccines are licensed for use [14–18], however, they are not widely used as a public health tools in endemic areas, with the exception of controlling severe outbreaks such as those following natural disasters [19–22].

A sustained typhoid fever outbreak occurred in Thailand in the 1970s. A sharp increase in cases was observed in 1973–1974, which finally peaked in 1975–1976. In response, the government of Thailand established a national typhoid immunization program, which represented the first programmatic use of a typhoid vaccine in the country [14, 22, 23]. The immunization program targeted over 5 million school aged children (7–12 years) each year in Bangkok between 1977 and 1987 (80% of the eligible population). Thus, Thai school children were eligible to receive a single locally produced heat/phenol-inactivated subcutaneous dose of 2.5 x 10⁹ S. Typhi organisms annually [14, 22, 23], before the program was halted in the early 1990s because of high rates of adverse reactions caused by the vaccine [22]. To our knowledge this is the only such programmatic use of a vaccine for controlling Typhoid fever in children in Thailand. Data from four teaching hospitals in Bangkok showed a 93% reduction in blood culture confirmed infections with S. Typhi between 1976 (n = 2,000) and 1985 (n = 132) [14, 23]. Notably, no significant decline was observed in isolation rates of Salmonella Paratyphi A (S. Paratyphi A), a Salmonella serovar distinct from S. Typhi that causes a clinical syndrome indistinguishable from typhoid fever, but for which S. Typhi vaccines
provide little or no cross-protection [14]. This observation suggests that the reduction in S. Typhi infections was not attributable to improvements in infrastructure and hygiene practices only [5, 15, 20, 32]. While the inactivated S. Typhi vaccine was found to be highly efficacious [22, 23], it is no longer used as a consequence of being overly reactogenic [14, 16, 22, 23, 24]. A Vi capsular polysaccharide vaccine [15] and live-attenuated oral vaccine of strain Ty21a [16] have since replaced this vaccine for travellers to endemic locations [5, 21, 24]. The typhoid immunization program in Thailand provided a unique opportunity to investigate the impact of immunization on S. Typhi populations circulating within an endemic area. Here we present an analysis of a historical collection of 44 S. Typhi isolates obtained from patients in Thailand between 1973 and 1992 (before and during the immunization program). As S. Typhi populations demonstrate little genetic diversity, we used whole genome sequencing (WGS) to characterize these isolates, and core genome phylogenetic approaches to compare the historic isolates from Thailand to a recently published global S. Typhi genomic framework [4].

Materials and Methods

Ethics statement

This is a retrospective study of bacterial isolates unlinked to patient information and was not subject to IRB approval.

Bacterial isolation and antimicrobial susceptibility testing

Forty-four S. Typhi isolates from patients with suspected typhoid fever attending hospitals in Bangkok, Nonthaburi, Loi, and Srakaew, in Thailand between 1973 and 1992 were available for genome sequencing in this study (Fig 1 and S1 Table). At the time of original isolation, bacterial cultures were transferred on nutrient agar slants to the department of Enteric Diseases, Armed Forces Research Institute of Medical Sciences (AFRIMS), Bangkok, Thailand for identification and antimicrobial susceptibility testing. At AFRIMS, bacterial isolates were subcultured on Heptac Enteric agar (HE) and identification was performed by biochemical testing on Kligler iron agar slants, tryptone broth for indole, lysine decarboxylase medium, ornithine decarboxylase medium, urease test, mannitol and motility media (Becker Dickenson, Thailand). Serological agglutination was performed using Salmonella O antiserum and Salmonella Vi antiserum (Dako, USA). Bacterial strains were stored frozen at -70°C in 10% skimmed milk or lyophilised in 10% skimmed milk; lyophilized ampoules were stored at 2-8°C. Prior to DNA extraction for sequencing, lyophilized bacteria were rehydrated with trypticanse soy broth, inoculated on McConkey agar and incubated at 37°C for 18-24 hours. If bacteria were stored frozen in skimmed milk, organisms were inoculated directly onto McConkey agar after thawing and then incubated at 37°C for 18-24 hours.

Antimicrobial susceptibility testing against ampicillin, chloramphenicol, cephalothin, gentamicin, kanamycin, neomycin, sulfisoxazole, trimethoprim/sulfamethoxazole, and tetracycline was performed by disk diffusion according to Clinical and Laboratory Standards Institute (CLSI) [25-28].

Genome sequencing and SNP analysis

Genomic DNA from the 44 S. Typhi from Thailand was extracted using the Wizard Genomic DNA Extraction Kit (Promega, Wisconsin, USA). Two μg of genomic DNA was subjected to indexed WGS on an Illumina Hiseq 2000 platform at the Wellcome Trust Sanger Institute, to generate 100 bp paired-end reads. For analysis of SNPs, paired end Illumina reads were
mapped to the reference sequence of S. Typhi CT18 (accession no. AL513382) [20] using the RedDog (v0.4) mapping pipeline, available at https://github.com/katholig/RedDog. RedDog uses Bowtie (v2.2.3) [30] to map reads to the reference sequence, then high-quality SNPs called with quality scores above 30 are extracted from the alignments using SAMtools (v0.1.19) [31]. SNPs were filtered to exclude those with less than 5 reads mapped or with greater than 2.5 times the average read depth (representing putative repeated sequences), or with ambiguous base calls. For each SNP that passed these criteria in any one isolate, consensus base calls for the SNP locus were extracted from all genomes (ambiguous base calls and those with phred quality scores less than 20 were treated as unknowns and represented with a gap character). SNPs with confident homozygous allele calls (i.e. phred score >20) in >95% of the S. Typhi genomes (representing a ‘soft’ core genome of common S. Typhi sequences) were concatenated to produce an alignment of alleles at 45,893 variant sites. The resultant allele calls for 68 of these SNPs were used to assign isolates to previously defined lineages according to an extended S. Typhi genotyping framework [32].

doi:10.1371/journal.pntd.0005274.g001
SNPs called in phage regions, repetitive sequences (354 kb; ~7.4% of bases in the CT18 reference chromosome, as defined previously [33]) or recombiant regions (~186 kb; ~4% of the CT18 reference chromosome, identified using Gubbins (v1.4.4) [34]) were excluded, resulting in a final set of 1,850 SNPs identified in an alignment length of 4,275,037 bp for the 44 isolates. SNP alleles from Paratyphi A strain 12601 [35] were also included as an outgroup to root the tree. For global context, raw read data [4] were also subjected to genotyping analysis and those isolates sharing the genotypes that were observed in the Thai collection (n = 340, details in S2 Table) were subjected to the same SNP analyses, resulting in a final set of 9,700 SNPs for a total of 386 isolates.

**Phylogenetic and SNP analysis**

Maximum likelihood (ML) phylogenetic trees (Figs 1 and 2) were constructed using the 1,850 and 9,700 bp SNP alignments, respectively, using RAxML (v 8.2.3) [36] with a generalized time-reversible model and a gamma distribution to model site specific recombination (GTR+Γ substitution model; GTRGAMMA in RAxML), with Felsenstein correction for ascertainment bias. Support for ML phylogenies was assessed via 100 bootstrap pseudoanalyses of the alignments. For the larger tree containing global isolates, clades containing only isolates from only a single country were collapsed manually in R using the drop.tip() function in the ape package [37]. Subtrees were extracted for each subtclade, which are therefore each rooted by the other subclades. Pairwise SNP distances between isolates were calculated from the SNP alignments using the dist.gene() function in the ape package for R [37].

**Accessory genome analysis**

Acquired antimicrobial resistance (AMR) genes were detected, and their precise alleles determined, by mapping to the ARG-ANNOT database [38] of known AMR genes using SRST2 v0.11.5 [39]. Plasmid replicon sequences were identified using SRST2 to screen reads for replicons in the PlasmidFinder database [40, 41]. Raw read data was assembled de novo with SPAdes (v 3.5.0) [42] and circular contigs were identified visually and extracted using the assembly graph viewer Bandage (v0.7.0) [43]. These putative plasmid sequences were annotated using Prokka (v1.10) [44] followed by manual curation. Where IncH1 plasmid replicons were identified using SRST2, and their presence confirmed by visual inspection of the assembly graphs, IncH1 plasmid MLST (pMLST) sequence types were determined using SRST2 [13, 39, 45, 65]. Where resistance genes were detected from short read data, Bandage was used to inspect their location in the corresponding de novo assembly graph in order to determine whether they were encoded in the bacterial chromosome or on a plasmid. Assembled contigs were concatenated and putative prophage genomes were identified with the PHAge Search Tool (PHAST) [47], and their novelty determined by BLASTN analysis against the GenBank database. Pairwise alignments between novel and known prophage sequences were visualised using the genoPlotR package for R [48].

**Nucleotide sequence and sequence read data accession numbers**

Raw sequence data have been submitted to the European Nucleotide Archive (ENA) under project PRJEB5381; individual sample accession numbers are listed in S1 and S2 Tables. Assembled phage and protein sequences were deposited in GenBank, accession numbers are listed in Table 1.
Fig. 2. Zoomed in phylogenies showing relationships of Thai S. Typhi to global isolates. Maximum likelihood trees including S. Typhi isolates from the Thai and global collections are shown, for each genotype that was observed amongst the Thai isolates. (A) Genotype 2.0.0 tree. (B) Genotype 2.1.7 tree. (C) Genotype 2.2.0 tree. (D) Genotype 2.3.4 tree. (E) Genotype 2.4.0 tree. (F) Genotype 3.0.0 tree. (G) Genotype 3.1.2 tree. (H) Genotype 3.2.1 tree. (I) Genotype 3.4.0 tree. (J) Genotype 4.1.0 tree. Colored branches and nodes indicate country of origin, according to the inset legend. Year of isolation is shown to the right of pink and red Thai isolates obtained before and after the introduction of the immunization program;
grey and black, non-Thai isolates obtained before and after the introduction of the immunization program. Thai isolates are also labeled to indicate their city of origin: L, Loei; B, Bangkok; S, Sriracha; N, Nonthaburi. SNP distances between isolates as well as AUR plasmids are labelled, with any resistance genes indicated in italics. Branch lengths are indicative of the number of SNPs.

Results

The population structure of S. Typhi in Thailand

All 44 S. Typhi isolates collected between 1973 and 1992 were subjected to WGS and SNP analysis. Genome-wide SNPs were used to construct a ML phylogeny and isolates were assigned to previously defined genotypes [32] using a subset of SNPs (see Methods). These analyses subdivided the population into ten distinct genotypes, each corresponding to a specific lineage in the ML phylogeny (Fig 1). Genotype 3.2.1 (which includes the reference genome CT18, isolated from Vietnam in 1993 [29]) was the most common (n = 14, 32%), followed by genotype 2.1.7 (n = 10, 23%). Genotypes 2.0 (n = 1, 2%) and 4.1 (n = 3, 7%) were observed only in 1973 (pre-vaccine period). Genotypes 2.1.7 (n = 10, 23%), 2.3.4 (n = 1, 2%), 3.4.0 (n = 2, 5%), 3.0.0 (n = 3, 7%), 3.1.2 (n = 2, 5%), were observed only after 1981 (post-vaccine period). Each of these post-immunization genotypes was from a single location and time period (Fig 1), consistent with short-term localized transmission. The only exceptions were the two S. Typhi 3.1.2 isolates, that were from Sriracha in 1989 and Bangkok in 1992 and separated by just 4 SNPs. Genotypes 2.1.2 and 2.4.0 were observed amongst both pre- and post-vaccine isolates.

Thai S. Typhi in the context of a global genomic framework

Based on the Thai S. Typhi genotyping results we hypothesised that the post-immunization typhoid infections in Thailand resulted from occasional re-introduction of S. Typhi from outside the country, as opposed to long-term persistence of S. Typhi lineages within Thailand. To explore this possibility, and to provide a global context for our analysis, we examined 1,832 S. Typhi genomes from a recently published global collection that included isolates from 65

Table 1. Summary of mobile genetic elements observed in S. Typhi isolates from Thailand.

<table>
<thead>
<tr>
<th>Isolate</th>
<th>Genotype</th>
<th>Name</th>
<th>Replicons detected and/or attachment sites</th>
<th>Size (no. putative genes)</th>
<th>Accession number</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>004</td>
<td>2.0.0</td>
<td>pTy004_01</td>
<td>FI6 (gOCH42)</td>
<td>106,898 bp (133)</td>
<td>KX833209</td>
<td>Cryptoic, Phage defence (IhA protein)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>X1</td>
<td>38,260 bp (49)</td>
<td>KX833212</td>
<td>Phage defence (Abortive Infection)</td>
</tr>
<tr>
<td>031</td>
<td>3.0.0</td>
<td>pTy031_01</td>
<td>NA</td>
<td>40,835 bp (53)</td>
<td>KX833210</td>
<td>Phage defence (Restriction Modification)</td>
</tr>
<tr>
<td>042</td>
<td>3.0.0</td>
<td>pTy042_01</td>
<td>NA</td>
<td>40,835 bp (53)</td>
<td>KX833210</td>
<td>Phage defence (Restriction Modification)</td>
</tr>
<tr>
<td>051</td>
<td>3.2.1</td>
<td>pTy051_01</td>
<td>H1</td>
<td>&gt;215 kbp</td>
<td>N/A.</td>
<td>AMR (sul1, tetA1, intB), aadA1</td>
</tr>
<tr>
<td>052</td>
<td>3.1.2</td>
<td>Prop/age STYPH</td>
<td>atnLAAGCCTGGTCAG atnRAAAGCCTGGTCAG</td>
<td>28,949 bp (39)</td>
<td>KX833211</td>
<td>Cryptoic</td>
</tr>
<tr>
<td>054</td>
<td>4.1.0</td>
<td>Prop/age STYP2</td>
<td>atnLATTGTAATGGAAGCTGATGTCGACGCTATTCGAGCAG</td>
<td>34,780 bp (50)</td>
<td>KX833213</td>
<td>Cryptoic</td>
</tr>
</tbody>
</table>
countries [3]. Genome-wide SNP-based ML trees for each of these genotypes, showing the relationships between Thai and global isolates, are shown in Fig 3. In general, post-vaccine Thai isolates were closely related to recent isolates sourced from neighbouring countries including Vietnam, Laos and Cambodia (Fig 2), consistent with regional endemic circulation. In contrast, most pre-vaccine isolates had no close neighbours in the global collection, particularly 2.0.0 strains (Fig 2A), suggesting they may have been Thailand-specific lineages that have died out following the vaccine program. The S. Typhi genomes in the global collection were mainly isolated 2–3 decades after the Thai isolates as we did not have access to contemporaneous isolates from these countries that could identify specific transfer events. However, all but three of the post-vaccine Thai isolates shared shorter SNP distances with isolates from neighbouring countries than they did with pre-vaccination Thai isolates (see Fig 3), consistent with these cases being caused by occasional re-introduction of genotypes circulating in the region. Notably, Thai S. Typhi 3.2.1 that were isolated in 1986–7 clustered separately from the 1973 pre-vaccine isolates (>60 SNPs apart), but closely with isolates from Vietnam and Cambodia (differing by as few as 7 SNPs, Fig 2H). Post-vaccine Thai S. Typhi 2.4 formed two distinct groups that were not consistent with direct descendence from earlier isolates (Fig 2B). These data are therefore consistent with transfer of S. Typhi into Thailand from neighbouring countries during the post-immunization program era, although the long-term circulation of ancestral populations in Thailand remains an unlikely alternative explanation.

Acquired antimicrobial resistance

We identified acquired AMR genes in the genomes of four S. Typhi genotype 3.2.1 that were isolated in Srakaew in 1986 (Fig 1, Table 1). These isolates shared the same four AMR genes:
snl1 (sulphonamides), catA1 (chloramphenicol), tetB (tetracyclines), and aadA1 (aminoglycosides) which were carried on near-identical plasmids of IncHI1 plasmid sequence type 2 (PST2). Although the presence of insertion sequences (IS) in these plasmids prevented the complete sequences from being assembled, the regions of these plasmids encoding the AMR genes were identical in all assemblies. This commonality suggests they are a single plasmid (referred to as pTy03(1) in Fig 1 and Table 1) that was likely acquired in a common ancestor of this clade. The chromosomal and IncHI1 plasmid sequences for these four isolates were very closely related to those of a 1993 Vietnamese isolate (Viet1 distra_1993) in the global S. Typhi collection [4, 45], consistent with regional transfer.

Other plasmids and mobile genetic elements

We identified three non-AMR related plasmids amongst the Thai isolates (Fig 1, Table 1). pTy004 (genotype 2,2), pTy004 (genotype 2,2) carried two novel plasmids that assembled into circular sequences, pTy004_01 and pTy004_02. The largest, pTy004_01, was a novel variant of the cryptic plasmid pHCM2 [29, 49] (Fig 1). pTy004 was isolated in Bangkok in 1973, making pTy004_01 the earliest example of a pHCM2-like plasmid reported to date. pTy004_01 was distant from other pHCM2-like plasmids in the global S. Typhi genome collection, sharing 92% coverage and 99% nucleotide identity with the reference sequence pHCM2 of S. Typhi CT18 (genotype 3.2,1) which was isolated approximately 20 years later in Vietnam [29]. The pTy004_01 sequence (Fig 1) appears to be ~2 kbp larger than pHCM2, and encodes an additional tRNA-Lys as well as an insertion of a hypothetical protein (orf17) into a putative DNA polymerase gene (HCM2.0015c in pHCM2, divided into orf16 and orf18 in pTy004_01). Plasmid pTy004_02 was ~80 kbp in size and similar to E. coli plasmid pFQ2 (65% coverage, 98% nucleotide identity), encoding genes for conjugation, chromosomal partitioning, addiction systems, and an abortive infection protein (orf44). Three isolates (Ty031, Ty042, and Ty049) all of genotype 3,0,0 and obtained from Srlaew in 1986, carried a ~40 kbp cryptic plasmid that we named pTy03(0,1). This plasmid was similar to that carried by Enterobacter hormaechei strain CAV1176 (83% coverage, 96% identity) and encoded genes for chromosomal partitioning, addiction systems, and a putative restriction modification system (orf33–orf34).

PHAST analysis revealed the presence of novel intact prophages in three Thai S. Typhi isolates (Fig 1, Table 1). Two S. Typhi 3.1.2 isolated from Srlaew in 1989 and Bangkok in 1992.
shared a novel phage STYP1 that was similar to fAA91-ss infective for Shigella sonnei (Fig 5A). However, the S. Typhi phage lacked the cytolysin distending toxin cdt genes and the IS21 element found in phage fAA91-ss [56]. This prophage sequence had a mosaic architecture, incorporating a number of putative insertions of phage tail fiber genes that were not present in the fAA91-ss reference genome (Fig 5A). Additionally, a single isolate of genotype 4.1 obtained from Bangkok in 1973 contained a novel S1V-like phage, here named STYP2, that lacked the serotype conversion gene Gtr cluster and IS1 element of phage S1V [51]. Again, the novel Thai phage variant also encoded novel tail fiber genes not in the S1V reference genome, as well as a Dam methylase gene (orf37) (Fig 5B).

Discussion

These data provide a historical insight into the population structure of S. Typhi in Thailand in 1973 (pre-immunization program, n = 11) and 1981–1992 (post-immunization program, n = 33). It has been reported that the national S. Typhi immunization program in Thailand, which commenced in 1977, was highly effective in reducing the burden of typhoid fever [14]. Our data are consistent with the hypothesis that the vaccine program successfully depleted the endemic S. Typhi population to the extent that most subsequent typhoid cases resulted from sporadic introduction of non-indigenous S. Typhi, rather than long-term persistence of the
S. Typhi in Thailand before and after Immunization

pre-vaccine era population. It is apparent that these introductions were sometimes accompanied by limited local transmissions, resulting in small, localized outbreaks, but we found no evidence to suggest that these result in the establishment of stable local source populations. Notably, the post-immunization S. Typhi isolates from Loi (in the north of Thailand near the border with Laos, from which it is separated by the Mekong river) were most closely related to Laos isolates, whilst those from the capital Bangkok and nearby Nonthaburi and Srakaew districts were closely related to other isolates from across Southeast Asia (Fig 2), suggesting there may have been multiple routes of import into Thailand.

Our study is limited by the sample of isolates available for analysis, which was small and reflects opportunistic sampling of sporadic local cases in the four sites and historical storage. A larger collection of historical isolates from Thailand and neighboring countries in the 1970s and 1980s would help to further elucidate the epidemiological patterns of S. Typhi before and after the vaccination program. However, from our data, it is notable that the Thai isolates clustered according to site, consistent with limited local transmission rather than dissemination of lineages between locations. The only exception to this was two genotype 3.1.2 isolates, which were collected from Srakaew in 1989 and Bangkok in 1992 and differed by only 4 SNPs. This is consistent with either transfer between these cities in Thailand following an initial introduction into the country, or two independent transfers into Thailand from a common source. The phylogenetic structure is most suggestive of the latter, but denser samples from Thailand and/or potential source populations would be required to resolve this with confidence.

While our sample is small, this study is nevertheless the largest to date exploring genetic diversity among S. Typhi from Thailand. An earlier global haplotyping study that included seven Thai isolates [32] identified five distinct haplotypes in Thailand (H1, 1969; H2, 1990; H3, 1989; H4, 1997; H5, 2002; H6, 1984; H7, 1998; H8, 2002), three of which are related to genotypes that we identified amongst Thai strains in this study (H7, H8, H9, H10, H11, H12). Genotype 4.3.1 (H5) was not found amongst our historical Thai isolates. This is consistent with previously published spatiotemporal analyses of the global isolate collection, which showed this rapidly expanding clone only began spreading throughout Asia after 1990 [4]. To our knowledge the only evidence to date of the presence of 4.3.1 (H5) in Thailand comes from the global study [4], in which three isolates were identified from 2010–2011, most likely introduced from India. Therefore, our genomic snapshot of the Thai S. Typhi population is consistent with previous insights and is likely reasonably representative for the study period. In the years following the vaccination program the prevalence of Typhoid fever in Thailand has continued to decline [53, 54]. The vaccination program has been credited with reducing disease incidence in Thailand and was followed by increased economic development in the region as well as improvements to both water and sanitation systems that have likely improved the control of such outbreaks [53, 54]. Consequently, Typhoid fever is no longer considered a serious public health threat in Thailand [53].

The presence of novel plasmids and prophages in the Thai isolates is also noteworthy. While small plasmids of unknown function have been observed in S. Typhi previously [55], they are infrequent compared to the IncH1 MDR plasmid and the cryptic plasmid pHCM2 [33]. Presumably, such plasmids are ephemeral, possibly because their maintenance imposes a fitness burden on the host cells so a strong selective advantage is required for retention [56, 57]. It is also possible that the lack of previous reports regarding the diversity of small plasmids in S. Typhi reflects a technological complexity, however, this is bypassed with high-throughput WGS and we detected negligible small plasmid content in the global collection of 1,832 genomes using the same screening approach [4, 32, 58]. Notably, few of the Thai plasmids share nucleotide sequence homology with those previously described in S. Typhi, but were closely related to those found in other Enterobacteriaceae. The novel pHCM2-like plasmid
(pTy004_01) and two additional plasmids (pTy004_02 and pTy031_01) harbored genes associated with phage resistance, which could provide protection against phage predation [59–62]. We also observed two novel prophages integrated into Thai genomes, which both showed variation in their phage tail structural regions compared to close neighbors found in Shigella/E. coli. These regions are typically responsible for binding of phage to host receptors [63–65], thus the variation in these regions may be associated with recent adaptations to the S. Typhi host. While genomic data from more recent S. Typhi collections shows limited evidence for genetic exchange with other organisms [4], the detection amongst older Thai isolates of both phage and plasmids that have been previously associated with E. coli/Shigella suggests that genetic exchange may have been more common in the past or in certain localized populations.

Overall, these data provide valuable historical insights into the S. Typhi populations circulating in Thailand during the 1970s and 1980s, and early examples of the two most common S. Typhi plasmids, as well as other mobile elements identified within the S. Typhi population. Importantly, while genomic epidemiology has been applied to study typhoid transmission, antimicrobial resistance evolution and antibiotic treatment failure in various settings [66–68], this study provides an important proof-of-principle demonstration that this approach can also provide useful insights into the impact of typhoid vaccines on circulating bacterial populations. This should motivate the adoption of WGS methods to monitor S. Typhi populations during future immunization programs and other large-scale interventions, which could potentially identify differential impacts on distinct genotypes.

Supporting Information
S1 Table. Isolate and sequencing details.
(OCX)
S2 Table. Global isolate and sequencing details.
(OCX)

Author Contributions
Conceptualization: SB KEH LB CJM DPT MAR.
Formal analysis: ZAD KEH.
Funding acquisition: GET KEH.
Investigation: DPT LB CJM AS MAR PVV THT.
Resources: LB CJM SB KEH.
Supervision: DPT.
Visualization: ZAD KEH.
Writing - original draft: ZAD.
Writing - review & editing: KEH SB.

References