

https://dspace.org/

Development and evaluation of new molecular epidemiological methods for analysis of salmonella TYPHI

Tau, Nomsa Pauline

2017
http://hdl.handle.net/10539/23279

Downloaded from a DSpace repository. https://dspace.org/

DEVELOPMENT AND EVALUATION OF NEW MOLECULAR EPIDEMIOLOGICAL METHODS FOR ANALYSIS OF *SALMONELLA* TYPHI

Nomsa Pauline Tau

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of requirements of the degree of Master of Science.

Johannesburg, 2017

DECLARATION

I Nomsa Tau, declare that this Dissertation is my own, unaided work. It is being submitted for the fulfillment of the degree of Master of Science at the School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.



Nomsa Tau

_____14th_____day of _____June_____2017_ in __Johannesburg_____

DEDICATION

To my son

Simphiwe Ndumiso Sibanyoni

LIST OF PRESENTATIONS

1. Nomsa Tau, Development and evaluation of a multiple-locus variable-number tandem-repeats analysis assay for subtyping *Salmonella* Typhi strains from Sub-Saharan Africa. Oral presentation, NICD Scientific Forum, National Institute for Communicable Diseases, Johannesburg, 26 February 2014

2. Nomsa Tau, Anthony Smith and Karen Keddy, Development and evaluation of a multiple-locus variable-number tandem-repeats analysis assay for subtyping *Salmonella* Typhi strains from Sub-Saharan Africa. Poster presentation, International Congress on Infectious Diseases, Cape Town, 02-05 April 2014

3. Nomsa Tau, Anthony Smith and Karen Keddy, Development and evaluation of a multiple-locus variable-number tandem-repeats analysis assay for subtyping *Salmonella* Typhi strains from Sub-Saharan Africa. Poster Presentation, Pathology Research and Development Congress (PathRed 2015), Johannesburg, 16 April 2015

LIST OF PUBLICATIONS

1. **Nomsa P. Tau, Anthony M. Smith, John R. Wain, Andrew Tarupiwa, Kalpy J. Coulibaly, Karen H. Keddy for GERMS-SA**. (In Press). Development and Evaluation of a Multiple-Locus Variable-Number Tandem-Repeats Analysis Assay for Subtyping *Salmonella* Typhi Strains from sub-Saharan Africa. *Journal of Medical Microbiology*

ABSTRACT

The typhoid fever causing *Salmonella* Typhi remains an important public health problem in Africa. More importantly, the emergence of the highly antimicrobial resistant H58 *Salmonella* Typhi haplotype is of greater concern. Rapid and highly discriminatory molecular methods are essential for prompt and effective epidemiological investigation of typhoid fever outbreaks. Traditional methods, such as pulsed-field gel electrophoresis (PFGE) are time-consuming and offer subjective discrimination of highly homologous isolates. On the contrary, molecular subtyping based on multiple-locus variable-number tandem-repeats (VNTR) analysis (MLVA) is a rapid, PCR-based method which has been successfully used for subtyping homogenous isolates of the *Salmonella* genus. This study describes the development and application of a MLVA assay for molecular characterization of *Salmonella* Typhi isolates from sub-Saharan Africa (SSA). This involved evaluation of thirteen VNTR loci using a validation panel consisting of 50 diverse *Salmonella* Typhi isolates. A MLVA assay consisting of five highly variable VNTR loci was adopted. The developed MLVA assay was used, along with PFGE, to characterize 316 *Salmonella* Typhi isolates from SSA. A total of 226 MLVA types were identified as compared to 143 PFGE fingerprint types. MLVA typing results indicated intracontinental spread of *Salmonella* Typhi. For the rapid identification of H58 *Salmonella* Typhi, a conventional PCR targeting a mutation that is exclusive to the H58 haplotype was employed on 105 isolates from South Africa as well as 121 isolates from other SSA countries. Approximately 54% (105/214) of the *Salmonella* Typhi isolates from South Africa and 62% (75/121) of the isolates from other SSA countries were identified as H58 *Salmonella* Typhi. The MLVA tool was able to discriminate among H58 *Salmonella* Typhi isolates. MLVA is viable alternative to PFGE for subtyping *Salmonella* Typhi and can be used as first-line assay for routine screening of *Salmonella* Typhi isolates in SSA, providing excellent discrimination of isolates.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Dr Anthony Smith, for his support, invaluable advice and guidance throughout my MSc. I would also like to thank my co-supervisor, Dr Karen Keddy, for constructive feedback and support towards my project. Thank you for granting me the opportunity to do my Masters.

To the great team that is the staff at the Centre for Enteric Diseases (CED), thank you for making the lab an enjoyable environment. Thank you to the scientist at CED for promoting a supportive, encouraging and fun office environment. A special thanks to Sanelisiwe Duze and Shannon Smouse for proofreading my dissertation.

Thank you to the Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa (GERMS-SA) as well as the Typhoid Fever Surveillance in Africa Program (TSAP) for the isolates used in this study. I would also like to thank the Global Disease Detection research grant for funding my project.

I would like to express my gratitude to the staff at Public Health England, Prof John Wain, Dr Katie Hopkins and Dr Satheesh Nair for training on Multiple-locus variable-number tandem-repeats analysis (MLVA). I learnt a lot during my stay in London.

Thank you to my friends, Ntombifuthi, Mimmy and Thamsanqa for the best wishes, encouragement and believing in me. I am very grateful to my best friend and husband, Bongani Sibanyoni, for his unending support, encouragement, understanding and believing in me throughout. I'm also grateful to my family for all their continuous support and encouragement throughout the years. Finally, I would like to thank my son, Simphiwe Sibanyoni for making my life an adventure.

"but with God all things are possible" Matthew 19:26

TABLE OF CONTENTS

Declaration ..... **i**

Dedication ..... **ii**

List of presentations ..... **iii**

List of publications ..... **iii**

Abstract ..... **iv**

Acknowledgements ..... **v**

Table of contents ..... **vi**

Nomenclature ..... **x**

List of figures ..... **xii**

List of tables ..... **xiv**

Chapter 1 ..... **1**

1.1 Background 1

1.2 The *Salmonella* bacterium 3

1.2.1 Typhoid fever 5

1.2.2 Pathogenesis of typhoid fever 6

1.2.3 *Salmonella* Typhi carriage 7

1.2.4 Typhoid fever prevention and vaccines 8

1.3 Epidemiology of typhoid in Africa 9

1.3.1 Mediterranean North African countries including Morocco, Algeria, Tunisia, Libya and Egypt 10

1.3.2 Nigeria 11

1.3.3 Ghana 11

1.3.4 Democratic Republic of Congo 12

1.3.5 Ethiopia 12

1.3.6 Malawi and Mozambique 12

1.3.7 Cameroon 13

1.3.8 Uganda 13

1.3.9 Kenya 13

1.3.10 Tanzania 14

1.3.11 South Africa 14

1.4 Diagnosis of typhoid fever 15

1.5 Phenotypic identification of *Salmonella* Typhi 17

1.5.1 Serotyping 17

1.5.2 Phage typing 18
1.5.3 Antimicrobial treatment and acquired resistance 19
1.6 Molecular characterization 23
1.6.1 The genome of *Salmonella* Typhi 23
1.6.2 Genotyping *Salmonella* Typhi 24
1.6.3 Pulsed-field gel-electrophoresis 24
1.6.4 Multi-locus sequence-typing 26
1.6.5 Multiple-locus variable-number tandem-repeats analysis 28
1.6.6 Whole-genome sequencing 30
1.7 The importance of studying the molecular epidemiology of *Salmonella* Typhi in the sub-Saharan African region 35
1.8 Aim 36
1.8.1 Specific objectives 36
Chapter 2 37
2.1 Surveillance programs 37
2.1.1 National Surveillance 37
2.1.2 Typhoid Fever Surveillance in Africa Program (TSAP) 37
2.2 Phenotypic characterization of *Salmonella* Typhi isolates 38
2.3 Crude genomic DNA extraction 38
2.4 Real-time PCR serotyping of *Salmonella* Typhi isolates 38
2.5 Pulsed-field gel electrophoresis (PFGE) 39
2.5.1 The PFGE reference standard – *Salmonella* Braenderup (strain H9812) 39
2.5.2 Preparation of bacterial plugs 40
2.5.3 Restriction enzyme digestion of genomic DNA and electrophoresis of DNA fragments. 40
2.5.4 Staining of agarose gels 40
2.5.5 BioNumerics analysis of PFGE bands 41
Chapter 3 42
3.1 Introduction 42
3.1.1 The study aim 43
3.2 Materials and Methods 44
3.2.1 Selection of bacterial isolates for the setting up of MLVA assay 44
3.2.2 Genotyping *Salmonella* Typhi from sub-Saharan Africa using MLVA 44
3.3 Results 48
3.3.1 Screening for length polymorphism of VNTR loci using simplex PCR 48
3.3.2 Measure of diversity at each VNTR locus 52
3.3.3 MLVA validation by nucleotide sequencing 54

3.3.4 Optimisation of the multiplex MLVA assay consisting of five VNTR loci 56
3.3.5 Subtyping of *Salmonella* Typhi isolates using the developed MLVA multiplex PCR assay 59
3.3.6 Discriminatory power of MLVA assay..... 61
3.4 Discussion 63
3.5 Conclusion..... 67
Chapter 4 68
4.1 Introduction 68
4.1.1 Aim..... 69
4.2 Materials and Methods..... 71
4.2.1 Crude DNA extraction..... 71
4.2.2 Serotype confirmation of *Salmonella* Typhi isolates..... 71
4.2.3 Conventional PCR identification of H58 *Salmonella* Typhi..... 71
4.2.4 Gel electrophoresis 72
4.3 Results..... 73
4.3.1 Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa (GERMS-SA study)..... 73
4.3.2 Multi-Country Typhoid Fever Surveillance in sub-Saharan Africa Program (TSAP Study) 78
4.4 Discussion 82
4.5 Conclusion..... 85
Chapter 5 86
5.1 Introduction 86
5.1.1 The study objectives 86
5.2 Materials and Methods..... 88
5.2.1 *Salmonella* Typhi isolates 88
5.2.2 Pulsed-field gel-electrophoresis (PFGE) 88
5.2.3 Multiple-locus variable-number tandem-repeats analysis (MLVA)..... 88
5.2.4 Data analysis 88
5.3 Results..... 89
5.3.1 Diversity of *Salmonella* Typhi isolates analysed by MLVA 89
5.3.2 MLVA typing of isolates from South Africa 89
5.3.3 Annual distribution of MLVA types amongst *Salmonella* Typhi in South Africa..... 94
5.3.4 MLVA typing of isolates from other sub-Saharan African Countries..... 98
5.3.5 Intracontinental transmission of *Salmonella* Typhi isolates..... 103
5.3.6 Comparison of PFGE analysis of *Salmonella* Typhi isolates from SSA to MLVA typing 105

5.3.7 Comparison of the discriminatory power of MLVA assay with PFGE analysis 112
5.4 Discussion 114
5.5 Conclusion 118
Chapter 6 119
6.1 Development of a five loci MLVA assay 119
6.2 The use of conventional PCR to study the prevalence of H58 *Salmonella* Typhi 120
6.3 MLVA offers higher discriminatory power than PFGE analysis for molecular subtyping *Salmonella* Typhi isolates from SSA 121
6.4 Concluding remarks 122
List of references 124
Appendix A Preparation of reagents for PFGE 136
Appendix B PCR amplification of VNTR loci 138
Appendix C Sequencing of VNTR loci using ABI Prism®BigDye Terminator cycle sequencing kit 140
Appendix D VNTR sequence alignments 141
Appendix E *Salmonella* serotyping flow chart 159
Appendix F Agarose gel electrophoresis 160
Appendix G MLVA profiles from MLVA typing of *Salmonella* Typhi from SSA 161
Appendix H Ethics clearance certificate 167
Appendix I Plagiarism declaration with Turnitin report 168

NOMENCLATURE

bp	base pair
CED	Centre for Enteric Diseases
CI	Confidence interval
cm	centimetre
°C	Degrees Celsius
dH_2O	Distilled water
dNTP	deoxynucleotide triphosphate
DI	Diversity index
DRC	Democratic Republic of Congo (DRC)
g	gram
GERMS-SA	Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa
IncHI1	Plasmid incompatibility type
L	Litre
MDR	multi-drug resistant
MIC	Minimum Inhibitory concentration
mg	milligram
$MgCl_2$	Magnesium chloride
ml	millilitre
MLST	Multi-locus sequence typing
MLVA	Multiple-locus variable-number tandem-repeats analysis
mM	millimolar
MST	minimum spanning tree
NHLS	National Health Laboratory Services
NICD	National Institute for Communicable Diseases
NTC	no template control

PCR	Polymerase Chain Reaction
PFGE	pulsed-field gel electrophoresis
PHE	Public Health England
rpm	revolutions per minute
SDS	sodium dodecyl sulphate
SSA	sub-Saharan Africa
SNP	single-nucleotide polymorphisms
ST	Sequence type
STyMT	*Salmonella* Typhi MLVA type
STyPT	*Salmonella* Typhi Pulsotype
TAE	Tris-Acetate-EDTA
TBE	Tris-borate EDTA
TE	Tris-EDTA
TSAP	Typhoid Fever Surveillance Program in Africa
µl	microliter
µM	micromolar
UPGMA	unweighted pair group method with arithmetic averages
USA	United States of America
V	Volts
v	volume
VNTR	variable-number tandem-repeats
wgMLST	Whole-genome MLST
WGS	Whole-genome sequencing
wg-SNP	Whole-genome SNP

LIST OF FIGURES

Chapter 1

Figure 1.1 Classification of *Salmonella* species....4
Figure 1.2 Global burden of typhoid fever and its geographical distribution, 2010 ..... **Error! Bookmark not defined.**
Figure 1.3 *Salmonella* Typhi infection of the epithelial cells in the intestinal tract 7
Figure 1.4 Population structure of the *Salmonella* Typhi H58 lineage and distribution across countries in the Sub-Saharan African continent 23
Figure 1.5 Outline of the whole genome sequencing workflow (Illumina platform). 32

Chapter 3

Figure 3.1 Dendrogram analysis of PFGE patterns of 50 *Salmonella* Typhi strains from sub-Saharan Africa 49
Figure 3.2 Geographical areas in sub-Saharan Africa as source of *Salmonella* Typhi strains included in the development of the MLVA assay 50
Figure 3.3 Bar diagram showing selection of fluorescent dyes based on amplicon size ranges for the five VNTR loci....58
Figure 3.4 Electropherogram showing all five VNTR loci incorporated in a multiplex PCR for MLVA assay of *Salmonella* Typhi isolates 58
Figure 3.5 MLVA cluster analysis of *Salmonella* Typhi isolates from sub-Saharan Africa 60

Chapter 4

Figure 4.1 Inter and intracontinental transfer of H58 *Salmonella* Typhi 70
Figure 4.2. Temporal distribution of 214 *Salmonella* Typhi isolates from the GERMS-SA study for the period 2012-2014. 73
Figure 4.3 Number of H58 *Salmonella* Typhi isolates identified in South Africa from 2012 - 2014. 74
Figure 4.4 Geographical distribution of H58 and non-H58 *Salmonella* Typhi within provinces in South Africa for the years 2012 to 2014....76
Figure 4.5 Bar graph representing *Salmonella* Typhi haplotypes identified in eight provinces in South Africa for the years 2012 to 2014....77
Figure 4.6 The Bar graph depicts haplotype specific typhoid fever infections by age group in South Africa, for the years 2012 to 2014....78
Figure 4.7 The bar graph showing the number of *Salmonella* Typhi isolates detected in SSA countries through the TSAP study for the years 2012 to 2014....79
Figure 4.8 Pie charts showing the prevalence H58 *Salmonella* Typhi isolates in SSA countries. 80
Figure 4.9 Map of Africa depicting the presence of H58 *Salmonella* Typhi in countries involved in typhoid fever surveillance through TSAP and GERMS-SA studies. 81

Chapter 5

Figure 5.1 Thirteen TSAP sentinel sites.....87
Figure 5.2 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from SSA collected over a three year period (n=316) i.....91
Figure 5.3 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from SSA collected over a three year period (n=316).....92
Figure 5.4 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected over a three year period (n=195).....93
Figure 5.5 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2012 (n=57). 95
Figure 5.6 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2013 (n=58). 96
Figure 5.7 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2014 (n=73). 97
Figure 5.8 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Burkina Faso collected over a three year period (n=12).....98
Figure 5.9 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Ghana collected over a three year period (n=34).....100
Figure 5.10 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Kenya collected over a three year period (n=56).....102
Figure 5.11 Outline map of Africa showing intracontinental distribution and transmission of *Salmonella* Typhi isolates based on MLVA assay. 104
Figure 5.12 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters I and II. 107
Figure 5.13 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters II and III. 108
Figure 5.14 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters III and IV.....109
Figure 5.15 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE cluster V. 110
Figure 5.16 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA.111

LIST OF TABLES

Chapter 2

Table 2.1 Primer and probe sequences for *Salmonella* species and *Salmonella* Typhi specific genes 39

Chapter 3

Table 3.1 VNTR loci showing no size variation amongst a panel of 50 *Salmonella* Typhi strains 51
Table 3.2 Simpson's diversity indices and primers selected for the amplification of 13 *Salmonella* Typhi VNTR loci 53
Table 3.3 Features of seven highly polymorphic *Salmonella* Typhi VNTR loci.....57

Chapter 4

Table 4.1 PCR primers selected for specific amplification of H58 *Salmonella* Typhi 72

Chapter 5

Table 5.1 Simpson's Index of Diversity 113
Table 5.2 Wallace coefficient and analytical 95% CI.....113

Chapter 1

Literature review

1.1 Background

Bacterial food-borne pathogens are a major cause of morbidity and mortality worldwide [1-3]. The widespread bacteria of the genus *Salmonella* are an important and leading cause of foodborne diseases and are mostly identified as aetiological agents of foodborne disease outbreaks [1, 3-5]. *Salmonellae* are transmitted from person to person and through consumption of contaminated food; hence making *Salmonella enterica* infections a huge problem in developing countries where there is poor sanitation and standard hygiene practices [6]. Epidemiological investigations play a huge role in public health improvement [2]. During food-borne disease outbreaks, epidemiological investigations are carried out in order to determine the primary sources of bacterial contamination. These can be used to link ill patients following consumption of contaminated food to the sources of bacterial contamination. Strong evidence linking exposure to incidence of infection within a population can be obtained through epidemiological investigations thereby improving public health management [2].

Traditional surveillance methodologies for *Salmonella enterica* species are based on phenotypic characterization of the bacterial pathogen [4]. The phenotypic techniques such as serotyping and phage typing are very useful; nonetheless they have a limited ability to further characterize predominant phenotypes within a species. Presently, newer methods based on characterization of bacterial genome are being employed [4, 7].

Molecular techniques have enhanced traditional surveillance and have become an essential tool for identifying, tracing and preventing dissemination of pathogenic bacteria [2]. These techniques have presented an exciting opportunity to learn about the genetic makeup of pathogenic bacteria consequently enhancing our knowledge and understanding of these organisms [2, 8]. It is through molecular epidemiological investigations that the true extent of

genetic diversity of pathogenic bacteria can be established [9]. Molecular subtyping techniques, in particular the application of molecular markers, are used to scrutinize variations in nucleotide sequences amongst bacterial isolates, and have become an integral part of epidemiological investigations of outbreak related isolates [8, 9]. These techniques are based on the premise that epidemiologically related isolates are derived from the clonal expansion of a single precursor, and that share characteristics that differ from epidemiologically unrelated isolates [4]. The application of molecular epidemiological techniques for subtyping food-borne bacterial pathogens is important to improve public health [2].

Molecular epidemiological techniques are fundamental tools in surveillance and outbreak investigations of human *Salmonella enterica* infections [10]. The use of these techniques for food-borne disease outbreaks caused by *Salmonellae* has resulted in a plethora of subtyping techniques with varying ability to differentiate homogenous serovars belonging to *Salmonella enterica* species [7, 8]. Currently, macro-restriction analysis of chromosomal DNA by pulsed-field gel electrophoresis (PFGE) is a common technique that has been used to discriminate *Salmonellae* at a DNA-level. PFGE has been used extensively for fingerprinting isolates in outbreak situations and is relatively inexpensive to use. Although PFGE has been used successfully in tracking the source of bacterial infection, this technique has limited ability to differentiate homogenous species and does not display equal sensitivity with different bacterial species [8].

Amongst the *Salmonella enterica* species, *Salmonella enterica* subspecies *enterica* serovar Typhi (*Salmonella* Typhi) is one of the serovars known to be highly homogenous [11]. *Salmonella* Typhi is the causative agent of typhoid fever, a grave systemic infection that remains an important cause of morbidity and mortality in developing countries [12, 13]. Molecular epidemiological investigation of *Salmonella* Typhi isolates is critical in understanding dissemination of the bacteria more especially in the African context where it is endemic [13, 14].

1.2 The *Salmonella* bacterium

Salmonellae belong to the family *Enterobacteriaceae* [15]. Isolates of this genus are Gram-negative, non-sporulating, facultative-anaerobic bacilli and are motile by peritrichious flagellae. The *Salmonella* bacterium emerged from *Escherichia coli* (*E. coli*) approximately 100 to 150 million years ago and has adapted to colonize various niches as it can be found as both a commensal and a pathogen in human and animal, and can also survive free in the environment [1, 15].

Biochemically, *Salmonella* isolates are characterized by glucose gas and hydrogen sulphide production, citrate utilization and fermentation of arabinose, rhamnose as well as trehalose. *Salmonella* isolates test negative for urease, indole and test positive for lysine- and ornithine-decarboxylase [16].

The *Salmonella* genus is divided into two species namely *Salmonella bongori* and *Salmonella enterica* (Figure 1.1) [17]. The species *Salmonella enterica* can be further subdivided into 6 subspecies including *Salmonella enterica* subspecies *enterica*, *Salmonella enterica* subspecies *salamae*, *Salmonella enterica* subspecies *arizonae*, *Salmonella enterica* subspecies *diarizonae*, *Salmonella enterica* subspecies *houtenae* and *Salmonella enterica* subspecies *indica*. The subspecies are further classified into serovars based on bacterial cell surface antigens [18].

According to the Kauffmann-White-Le Minor classification scheme, the serovars can be determined based on the immunologic reactivity of the two surface structures including the O-antigen (lipopolysaccharide) and the H-antigens (flagella proteins) [17, 18]. The Vi-antigen (subtype of the capsular K antigen) is only found in some pathogenic *Salmonella*. This antigenic diversity has led to the assignment of *Salmonella* isolates to approximately 2500 serovars. *Salmonella enterica* subspecies *enterica* is the most important subspecies, causing approximately 99% of all human and animal *Salmonella* infections. Clinically, this subspecies can be split into typhoidal *Salmonella* and non-typhoidal *Salmonella*. While the non-typhoidal *Salmonellae* are known for extra-intestinal disease and gastroenteritis, the typhoidal *Salmonellae* are characterized by invasive disease, and consist of *Salmonella enterica*

subspecies *enterica* serovar Typhi (*Salmonella* Typhi) and *Salmonella* enterica subspecies enterica serovar Paratyphi A, B and C (*Salmonella* Paratyphi A, B and C) [17, 18].

Salmonella Typhi isolates share common antigenic determinants including being serologically positive for lipopolysaccharide antigens O9 and O12, flagella protein antigen Hd and the polysaccharide capsular antigen Vi [18]. In contrast to the other ~2500 serovars of *Salmonella enterica* subspecies *enterica* (with the exception of *Salmonella* Paratyphi A, B and C), *Salmonella* Typhi is exclusively adapted to the infection of human host [18, 19]. This specificity for human suggests that *Salmonella* Typhi might be younger than the split that occurred several million years ago between human and higher primates, and is possibly as young as the expansion of the anatomically modern humans from Africa which occurred ~50,000 to 100,000 years ago [19].

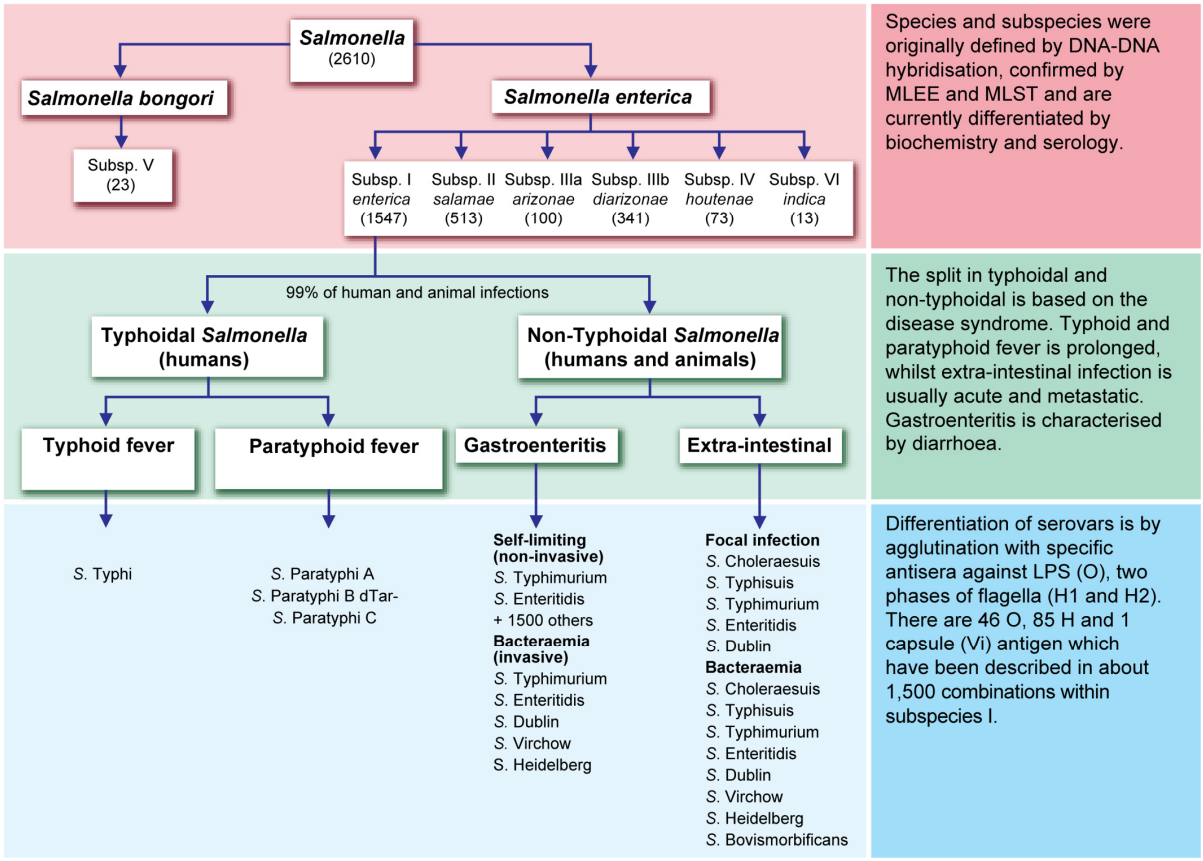
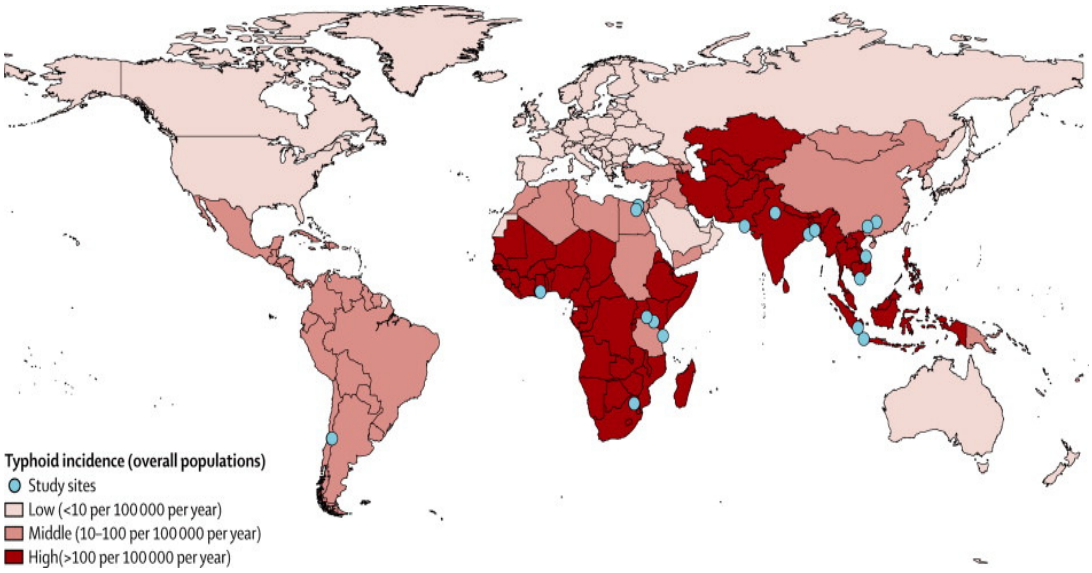


Figure 1.1 Classification of *Salmonella* species [17]

1.2.1 Typhoid fever

Salmonella Typhi is the aetiological agent for the life threatening systemic disease known as typhoid fever [20]. This invasive disease is characterized by high fever, headache, gastrointestinal symptoms (including abdominal pain, nausea, constipation and diarrhoea) and systemic infection which can be detected by isolation of the *Salmonella* Typhi bacterium from patient's blood or bone marrow [12, 21]

Recent global estimates indicate that typhoid fever causes 26.9 million illnesses annually [13, 20]. Although not common in industrialized countries, typhoid fever remains a major public health problem in developing countries. Typhoid fever is endemic in Latin America, Southeast Asia and Africa (Figure 1.2) where the disease is exacerbated by poor living conditions [20, 22-27]. The incidence of typhoid fever in the sub-Saharan African region has not been well characterized [20, 28]. In 2010, Buckle and colleagues conducted a study aimed at calculating the global burden of typhoid fever disease [13]. In their study, they concluded that there are approximately 724.6 typhoid fever cases per 100 000 population in Africa [13]. Even so, the true burden of typhoid fever in Africa is still unclear due to under-reporting as only severely ill patients seek hospital treatment and, lack of blood-culture facilities essential for diagnosis [23, 26, 29]. These estimates have relied on limited available data from countries in the sub-Saharan African region [23, 26, 29]



1.2.2 Pathogenesis of typhoid fever

Poor environmental sanitation and lack of provision of safe drinking water remain the main cause of typhoid fever endemicity in developing countries [20]. Contrary to other *Salmonella* serovars, humans are the only reservoirs for *Salmonella* Typhi [30]. The explanation behind this host specificity is still not clear. However, it has been reported that *Salmonella* Typhi could have undergone some mutation which could have led to specificity to the human host [30]. Typhoid fever is considered a disease of children, adolescents and elderly people [30, 31]. Although there are consistent reports of infections in very young children, these infections are generally uncommon [30]. This could be explained by the fact that young children display atypical response to *Salmonella* Typhi infection resulting in misdiagnosis [30].

Salmonella Typhi infectious dose varies between 1000 and 1 million organisms [32]. Once ingested, the bacteria survives the gastric-acid barrier in the stomach passing through the epithelial cells to proliferate in the Peyer's patch [33]. The *Salmonella* Typhi bacteria migrate into the mesenteric lymph nodes where they multiply [33, 34]. They are then released in the blood stream (primary bacteraemia) where they translocate to the reticuloendothelial cells of the liver and spleen. During the re-entry of *Salmonella* Typhi (secondary bacteraemia), the bacteria are then removed from the blood via the liver to infect the gallbladder [34]. The gallbladder has been reported to be the preferred site of *Salmonella* Typhi residence in persistently infected humans. Gallbladder infection by *Salmonella* Typhi results in excretion of the bacteria in urine and faeces (Figure 1.3) [34, 35]. The infection of the gallbladder might lead to reinfection of the small intestine with the second exposure of Peyer's patches to the *Salmonella* Typhi bacteria. This may result in inflammation, ulceration and necrosis ultimately complications include endocarditis, meningitis, renal failure and pneumonia [32, 34].

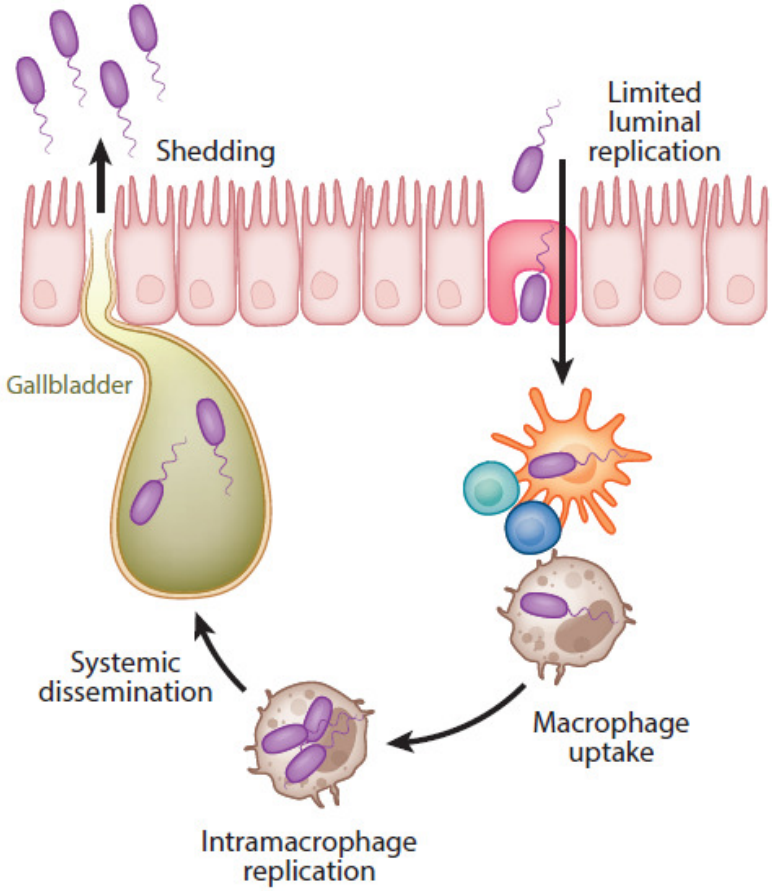


Figure 1.2 *Salmonella* Typhi infection of the epithelial cells in the intestinal tract [30]

1.2.3 *Salmonella* Typhi carriage

Salmonella Typhi reaching the gallbladder can establish an acute, active infection which is accompanied by inflammation or can persist in this organ long after the symptoms subside [32, 36]. Bile is an important digestive secretion in the liver that serves as potent emulsifying and antimicrobial agent in the gastric tract, however, *Salmonella* Typhi has adapted to withstand these harsh conditions [36]. These bacteria are concentrated in the bile-concentrated gallbladder. Reports indicate that gallbladder damage and the presence of gallstones contribute to carriage of *Salmonella* Typhi, even though the precise role of gallstone in carriage has not been described [36]. There are life-threatening risk factors associated with carriage of *Salmonella* Typhi in the gallbladder including cancer of the gallbladder, pancreas, lung and female colorectum [36-38].

Chronic typhoid fever carriage may develop as a result of acute typhoid fever infection or after subclinical infection with the bacterium [39]. It has been reported that subclinical cases are five times higher than acute cases, with about 10% of healthy individuals identified as carriers [39]. The infection of the gallbladder by *Salmonella* Typhi can result in long-term shedding of the bacterium through urine or faeces [34, 40]. This plays a major role in the transmission of the diseases. It has been reported that approximately 10% of untreated patients and 3% of treated patients become temporary carriers who shed the bacteria for months or years [34, 39, 40].

Since *Salmonella* Typhi is a human restricted pathogen, carriers are mostly responsible for endemic nature of typhoid fever and for community outbreaks [34]. Such was the case of a cook from New York City, Mary Mallon, also known as "Typhoid Mary", who in the early 20th century, served as a source for typhoid fever as she shed high levels of *Salmonella* Typhi [41]. While Typhoid Mary was asymptomatic and leading a normal life, reports indicate that she infected approximately 54 people with typhoid fever. During those days, she had to be quarantined to stop her from spreading typhoid fever further [41].

1.2.4 Typhoid fever prevention and vaccines

The provision of safe drinking water, hygienic food preparation and proper management of sewage systems could be the most effective way to prevent typhoid fever spread [28, 32]. This has been shown in a cross-sectional, laboratory-based surveillance study in the United States of America (USA) from the years 1999 -2006, where a dramatic decline in incidence rates and associated mortality of typhoid fever was observed following widespread implementation of municipal water and sewage treatment systems. Currently the majority of cases in the USA are associated with travel to typhoid fever endemic countries [28]. Unfortunately, the provision of safe drinking water and foodstuffs in the developing countries, particularly in rural areas, might not be achieved in the near future [42].

The use of vaccines against typhoid fever can provide short to medium term protection against this disease [43]. The first typhoid fever vaccine was developed in the late 1800s [44]. This vaccine was developed by Almroth Wright in Britain using a heat- and phenol-treated *Salmonella* Typhi preparation. The vaccine was used by the British Army during the Anglo-

Boer War in southern Africa in 1899 for protection against typhoid fever infection. Major Frederick Russell of the US Army Medical School later modified the vaccine by using inactivated whole-cell *Salmonella* Typhi. The widespread use of this vaccine continued amongst the soldiers and the Navy from the year 1909 [44]. It was later established that the efficacy of the whole-cell vaccine over a period of 3 years was 73%. Although whole-cell vaccines were very effective, the vaccine demonstrated a high rate of adverse effects [44, 45]. It is for this reason that the use of whole-cell typhoid fever vaccines was renounced [44].

Newer typhoid fever vaccines have been developed and these provide an alternative to the highly efficacious whole-cell vaccine [32, 46]. There are two commercially available typhoid fever vaccines including Ty21a (oral) and Vi polysaccharide (parental) vaccines [32, 46, 47]. The Ty21a vaccine is a live-attenuated vaccine given orally in adults and children six years or older. The vaccine has a 50% protection efficacy over 3 years and there are no significant adverse effects reported. The Vi vaccine is a purified Vi antigen and is administered as intramuscular injection in adults and children over two years of age [32]. Although this vaccine provides 60% efficacy over a period of two years, a booster dose is still required every after two years to maintain the protection against typhoid fever [46-48].

Currently, the commercially available vaccines are recommended for travellers to endemic regions, laboratory workers and household contacts of typhoid carriers, but are not used routinely [32]. Mass immunization coupled with adequate provision of safe drinking water has been recommended in typhoid fever endemic areas [32].

1.3 Epidemiology of typhoid in Africa

In Africa, typhoid fever is mainly a disease of young children, adolescents and the elderly [22]. Even though the true burden of this disease has not been established in SSA, there have been several reports of typhoid fever from a number of African countries [13, 22].

1.3.1 Mediterranean North African countries including Morocco, Algeria, Tunisia, Libya and Egypt

The incidence of typhoid fever in the Mediterranean North African countries (MNAC) is estimated at 10-100 cases per 100 000 populations for the years 2000 to 2005. Outbreaks in this area have been attributed to the consumption of sewage-contaminated and untreated water [49].

1.3.1.1 MNAC: Libya

For the years 1975 to 1980, *Salmonella* Typhi was identified amongst the most frequently reported *Salmonella* serotypes [50]. The incidence of typhoid fever in Libya has been estimated at 7-21 cases per 100 000 population between the years 2004 to 2006 [49].

1.3.1.2 MNAC: Tunisia

Studies have shown that *Salmonella* Typhi was identified amongst the eight most frequently isolated *Salmonella* serotypes in Tunisia [49, 51, 52]. Typhoid fever outbreaks in this country were recorded in 1999 in hospitalized patients in Sousse City, and in 2004 and 2005 in southeast of Tunisia [52, 53]. The incidence of typhoid fever in Tunisia has been estimated at 1-6 cases per 100 00 population per year [49].

1.3.1.3 MNAC: Morocco

The incidence of typhoid fever in Morocco is estimated at 8-17 cases per 100 000 population [49, 54]. Reports indicated that for the years 1999 to 2000, *Salmonella* Typhi dropped from being the most commonly isolated *Salmonella* serotype to being the third most common serotype in Morocco. In the year 2001, no typhoid fever cases were reported in Morocco, however, during the year 2002, *Salmonella* Typhi was the fourth most commonly isolated *Salmonella* serotype [49].

1.3.1.4 MNAC: Algeria

Salmonella Typhi has been identified amongst the most common causes of diarrhoea, predominately affecting male children in Algeria [49]. This pathogen was identified as the *Salmonella* serotype responsible for approximately 98% of the 3340 clinical *Salmonella* isolates recovered between the years 1986 to 1990 [55]. For the years 1985 to 2005, the

incidence of typhoid fever was estimated at 3-22 cases per 100 000 populations. Typhoid fever outbreaks in Algeria have resulted from the consumption of sewage contaminated village water reservoirs; and as a result, more than 100 water wells have been filled and 16 suspected drinking water supply points stopped in order to control typhoid fever spread in the country [49, 56, 57].

1.3.1.5 MNAC: Egypt

Typhoid fever is endemic in Egypt and is most commonly a disease of children and adolescents aged 3 to 17 years old [58]. The incidence of typhoid fever in Egypt is estimated at 59 cases per 100,000 populations per year [59]. Of most concern is the widespread prevalence of MDR *Salmonella* Typhi [59]. Wasfy and colleagues (2002) studied the trends of multi-drug resistance among *Salmonella* Typhi isolates from the years 1987 to 2000. Interestingly, they reported a significant shift in the prevalence of MDR *Salmonella* Typhi from nonexistence to greater than 60% and then decreasing to only 5% in a period of 15 years [58].

1.3.2 Nigeria

Typhoid fever is a major public health problem in Nigeria and has been rated eighth amongst widespread diseases affecting the population [60, 61]. Typhoid fever is endemic this country and affects people of all age groups [60, 62]. Reports indicate an alarmingly increasing rate of resistance amongst *Salmonella* Typhi isolates [63-67]. Adabara and colleagues (2012) studied the prevalence and antimicrobial susceptibility patterns of *Salmonella* Typhi amongst patients attending a Military hospital in Minna, Nigeria. *Salmonella* Typhi isolates were found to be resistant to ceftriaxone, cefuroxime, amoxicillin, ampicillin, ciprofloxacin and augmentin, all of which are antimicrobials of choice routinely used for the treatment of typhoid fever. Interestingly, these isolates were susceptible to chloramphenicol [63].

1.3.3 Ghana

Typhoid fever has been reported to be amongst the 20 leading causes of outpatient illness in Ghana, accounting for 0.92% of hospital admissions [68]. Marks and colleagues (2010) conducted a study which included approximately 1500 children less than 15 years of age who were admitted to the paediatric ward of Agogo Presbyterian Hospital from September 2007 to

November 2008. The highest incidence of typhoid fever (290/100 000 population per year) was identified in children 2 to 5 years of age. The incidence of typhoid fever was estimated at 190/100 000 population per year in children less than 5 years of age and 200/100 000 population per year in children 5 to 8 years of age. In children older than 8 years of age the incidence decreased with increasing years. Resistance to ciprofloxacin and ceftriaxone was identified in less than 10 % of *Salmonella* Typhi isolates [68].

1.3.4 Democratic Republic of Congo

Salmonella Typhi has been reported to be endemic in the Democratic republic of Congo (DRC) [69]. Muyembe-Tamfum and colleagues (2009) compiled the first report of multidrug-resistant *Salmonella* Typhi in Kinsasha between the years 2004 and 2005 [70]. In their study, *Salmonella* Typhi was isolated from blood of patients with peritonitis. The isolates were resistant to first-line antimicrobials but were susceptible to third-generation cephalosporins, quinolones and amoxicillin-clavulanic acid [70]. Lunguya and colleagues (2012) characterized 201 *Salmonella* Typhi isolates from the DRC [69]. Infected patient age ranged from less than 4 years of age to greater than 50 years of age with the median age of 5 years of age. Not only were the isolates resistant to first line antimicrobials but they also had decreased susceptibility to ciprofloxacin. PFGE analysis indicated 33 pulsotypes with 72% of the isolates showing a single profile [69].

1.3.5 Ethiopia

Typhoid fever infections have been reported in Ethiopia [22, 71, 72]. Bogale Worku (2000) reported typhoid fever in an Ethiopian children's hospital between the years 1984 to 1995 [72]. The study showed the prevalence of typhoid fever infections amongst children aged less than 2 years to adolescents aged 13 years with high case fatality rate of 15.7% due to chloramphenicol treatment of multi-drug resistant typhoid fever [72].

1.3.6 Malawi and Mozambique

The incidence of typhoid fever in Malawi and Mozambique is estimated at rates of 10-100 cases per 100 000 population per year [73]. In Malawi, typhoid fever remains a significant public health problem and is mostly a disease of school aged children and younger adults [23,

74, 75]. Although previously reported as uncommon in Mozambique, *Salmonella* Typhi caused an outbreak along Mozambiquean and Malawian borders in March to November 2009 [73, 76]. Lutterloh and colleagues (2012) studied 42 *Salmonella* Typhi isolates from the outbreak and reported resistance to first-line antimicrobials with four isolates having additional resistance to nalidixic acid and decreased susceptibility to ciprofloxacin. It was concluded that the outbreak was caused by a single clone as 83% (35/42) of the isolates were indistinguishable by PFGE analysis [73].

1.3.7 Cameroon

Even though there's only limited data regarding the incidence of typhoid fever in Cameroon, a report by Nkemngu and colleagues (2005) confirmed the presence of *Salmonella* Typhi in this country [77]. In this report, Nkemngu and colleagues (2005) indicated the emergence of multidrug resistant and nalidixic acid resistant *Salmonella* Typhi. This isolate was obtained from a 29 year old woman who presented to the St. John's hospital with symptoms of fever, emesis, myalgia and hepatosplenomegaly. *Salmonella* Typhi was isolated from blood culture and the isolate was found to be resistant to first-line antimicrobials and to nalidixic acid [77].

1.3.8 Uganda

A large laboratory-confirmed outbreak of typhoid fever was reported in Uganda from December 2007 to July 2009 [78]. From a total of 27 laboratory confirmed cases of typhoid fever, 76% were resistant to ampicillin, streptomycin, sulfisoxazole, tetracycline and cotrimoxazole and were susceptible to chloramphenicol. Resistance to nalidixic acid and ciprofloxacin were not detected. PFGE analysis revealed several different pulsotypes suggesting that multiple sources of infection were present. PFGE analysis also showed different pulsotypes for isolates collected from the same patient indicating co-infection with different *Salmonella* Typhi isolates [78].

1.3.9 Kenya

Typhoid fever in Kenya is mostly a disease of children aged 2 to 9 years of age [29, 79, 80]. Typhoid fever outbreaks have been well documented in this country [29, 79-83]. Kariuki and colleagues (2004) characterized 102 *Salmonella* Typhi isolates collected from three parts of

Kenya, including Nairobi Province, Embu district and Thika district hospital, during an typhoid fever outbreak which occurred during January 2000 to December 2002 [82]. Analysis of these isolates indicated an increase in the number of MDR *Salmonella* Typhi isolates. These isolates were seen to be gradually replacing the sensitive phenotype. PFGE analysis of these isolates indicated 2 circulating strains (including MDR and fully sensitive strains) [82]. Not only is MDR a public health problem in Kenya, but the increase is fluoroquinolone resistant *Salmonella* Typhi isolates has become a major public health concern in the treatment of typhoid fever and other bacterial infections in Kenya [29, 82]. Kariuki and colleagues (2010), through DNA sequence-based analysis, determined that the increase in the incidence of fluoroquinolone resistant *Salmonella* Typhi was attributed to *Salmonella Typhi* haplotype H58 [29]. This haplotype has been detected as the phylogenetic lineage of *Salmonella* Typhi responsible for the increase in antimicrobial resistant *Salmonella* Typhi isolates in Kenya [29].

1.3.10 Tanzania

Salmonella Typhi was found to be the leading cause of fever in patients in Pemba Island, Zanzibar and Tanzania mainland between March 2009 and December 2010 [84]. The report by Thriemer and colleague (2012) showed that typhoid fever affected patients with ages ranging from 10 months to 58 years with a mean age of 17.7 years and most of these patients were female (61%) [84].

1.3.11 South Africa

There have been several reports on typhoid fever outbreaks in South Africa. In the year 1992, Coovadia and colleagues (1992) reported the first outbreak of multidrug resistant *Salmonella* Typhi in South Africa [85]. A cluster of six typhoid fever cases were described from three adjacent districts in South Africa and all isolates showed resistance to first line antimicrobial treatment [85]. Antimicrobial susceptibility testing revealed high levels of resistance to ampicillin, chloramphenicol, tetracycline, trimethoprim, sulphamethoxazole and trimethoprim-sulphamethoxazole. These isolates were β-lactamase producers and high mortality rates (50%) were reported [85]. In the year 2003, national surveillance for *Salmonella* species was introduced. Data was collected through the national surveillance representing data from approximately 205 laboratories across the country [23]. Typhoid fever affects mainly school-aged children and younger adults [23]. Reports of MDR *Salmonella* Typhi indicate an increase

in the number of quinolone-resistant *Salmonella* Typhi [86, 87]. Several other typhoid fever outbreaks have been reported in this country [88, 89]. During the year 2005, an outbreak of typhoid fever occurred in the Delmas area in Mpumalanga [88]. During this outbreak, approximately 600 clinically diagnosed typhoid fever cases were reported. This outbreak was second to the one that occurred in the same municipality during the year 1993 [88, 90]. Molecular epidemiological techniques showed that the isolates from the 1993 as well as the 2005 outbreaks were related [88]. In the year 2010, an outbreak of typhoid fever was reported in the Pretoria district in Gauteng [89]. Eight laboratory confirmed cases of typhoid fever were reported during this outbreak and patients involved were students and friends of students who all ate at a common restaurant in the Pretoria area [89].

1.4 Diagnosis of typhoid fever

Typhoid fever diagnosis can be difficult as the presenting symptoms are similar to those of Malaria and dengue fever [91]. Isolation of *Salmonella* Typhi bacterium from blood still remains the gold standard for laboratory diagnosis [92]. Even so, the lack of blood culture facilities in most primary health care facilities in the developing countries makes typhoid fever diagnosis difficult [93].

The Widal test has been used in many developing countries as a tool for typhoid fever diagnosis [94-96]. This test was developed by F. Widal in 1896, and is based on agglutination reaction between *Salmonella* Typhi somatic lipopolysaccharide O antigen and the flagella H antigen. The Widal test is relatively cheap, easy to perform, requires unsophisticated facilities and minimal training; however, the value of the test has been questioned due to lack of sensitivity and specificity as the O antigen and the H antigens targeted in the test are shared by many *Enterobacteriaceae* species [91, 93, 94].

Newer diagnostic tests have been developed for typhoid fever diagnosis [91]. These include the TUBEX® and Typhidot® which are based on direct detection of IgM antibodies against a host of specific *Salmonella* Typhi antigens. In a study by Keddy and colleagues (2011) the Widal test was compared to blood culture test and newer typhoid rapid antibody tests including the semi-quantitative slide agglutination test, the TUBEX® and Typhidot® [97]. Although

acceptable as diagnostic tests, none of these rapid tests exceeded blood culture performance in typhoid fever diagnosis [97].

PCR has been successfully used as a tool for diagnosis of typhoid fever [91, 98, 99]. In a study by Hashimoto and colleagues (1995), a PCR assay targeting the sequence that encodes the Vi antigen (ViaB) was developed [98]. This assay was specific for *Salmonella* Typhi due to the limited distribution of the Vi antigen amongst organisms. In order to detect *Salmonella* Typhi from blood, Hashimoto and colleagues used a nested-PCR strategy combined with the PCR primers targeting the ViaB region, and this offered a useful tool for rapidly and specifically detecting *Salmonella* Typhi from clinical specimens. Additionally, a nested PCR that targets the *H1-d* gene has been used to amplify *Salmonella* Typhi specific genes from blood [98].

In another study, Song and colleagues (1993) developed a PCR targeting the flagellin gene of *Salmonella* Typhi [100]. Unlike the flagella antigen and d-antigen that are not specific structures of *Salmonella* Typhi, the flagellin gene of *Salmonella* Typhi has unique nucleotide sequences that are specific for the organism. Song and colleagues developed a nested-PCR that targets fragments of the flagellin gene that are specific to *Salmonella* Typhi. This PCR has become a promising tool for rapid diagnosis of typhoid fever [91, 99, 100].

Rapid detection of *Salmonella* Typhi isolates has been improved with the development of real-time PCR assays [101, 102]. Real-time PCR is a fluorogenic based PCR method which utilizes an internal fluorogenic probe that is specific to the target gene [101]. During PCR, the amplification of the target gene is recognized and computationally monitored by the increase in fluorescence resulting from the fluorescent probe moiety. The most commonly used fluorogenic PCR-based method is the TaqMan® assay. This is based on a linear fluorogenic probe which requires the 5’ to 3’ endonuclease activity of the DNA polymerase. For the TaqMan® real-time PCR, the linear fluorescent moiety is conjugated to one end of the target sequence and a quencher moiety is attached to the other end of the target sequence. During PCR amplification of the target sequence, the quencher moiety is cleaved off, resulting in the fluorescence emission by the fluorogenic moiety. Contrary to the traditional PCR method, which uses agarose gels to detect amplicons of the target gene, real-time chemistries have the

advantage of timeous detection of target amplification by measuring kinetics of the reaction in early PCR phases [101-103].

Ranjbar and colleagues (2014) developed a TaqMan® real-time PCR assay for the rapid detection of *Salmonella* Typhi [101]. The PCR targeted a putative fimbrial protein (*staG*) of the *Salmonella* Typhi. The TaqMan® real-time PCR assay was tested against various *Salmonella* enterica serovars as well as non-*Salmonella* microorganisms. This PCR was highly sensitive, rapid and specific detection of *Salmonella* Typhi in clinical samples [101].

Nga and colleagues (2010) developed a real-time PCR assay for the detection of *Salmonella* Typhi and *Salmonella* Paratyphi A [102]. In this study, the PCR targeted sequences unique to *Salmonella* Typhi and *Salmonella* Paratyphi A. These unique sequences were identified by aligning whole genome sequences of *Salmonella* Paratyphi A strain ATCC9150, *Salmonella* Typhi strains Ty2 and CT18 as well as shotgun-sequenced strains. Although the real-time PCR exhibited high specificity for *Salmonella* Typhi and *Salmonella* Paratyphi, a lack of sensitivity attributed to the low physiological level of these pathogens in blood was observed. However, the real-time PCR assay showed 100% sensitivity on culture positive bone marrow biopsies. This increased sensitivity was attributed to the high bacterial loads in bone marrow biopsies [102].

1.5 Phenotypic identification of *Salmonella* Typhi

Salmonella enterica isolates are classically identified and typed using phenotypic methods such as biochemical profiling, serotyping, phage typing as well as antimicrobial resistance profiling. Microbial typing is essential for establishing and understanding clonal relations between microbial isolates.

1.5.1 Serotyping

Serotyping is used as the basis for the classification of *Salmonella* enterica isolates [8, 104]. This methodology does not have the capacity to fingerprint the isolates in a sensitive manner; however, it still remains useful in surveillance programs. Serotyping deciphers the antigenic

makeup of the pathogen by identifying variants of the somatic (O) and the flagella (H) antigens through reaction with specific antisera [8]. The O-antigen is the saccharide component of the lipopolysaccharide layer that is exposed on the cell surface of bacteria [104, 105]. The reaction of the O-antigen with specific antisera forms the basis of *Salmonella* serotyping. Most *Salmonella* isolates possess two different copies (phase I and phase II) of genes encoding the H-antigen. Even so, only one H-antigen is expressed at a time. The H-antigens are determined by reaction towards specific antisera. *Salmonella* isolates are characterized and classified by their antigenic properties according to the serological Kauffmann-Le Minor scheme [104, 105]. Currently, 2,463 *Salmonella* serovars have been identified [18].

1.5.2 Phage typing

Phage-typing of *Salmonella* isolates is a useful tool for identification and subtyping of these pathogens [91]. A bacteriophage is defined as a virus which specifically attacks bacteria. The method is based on bacterial cell lysis by a specific bacteriophage. The lysis that occurs as a result of the viral infection is specific and important in the identification of the bacterium. A plaque assay, in which a clear zone results from the bacterial lysis, is used for enumeration of the lytic phages [91].

In 1938, Craigie and Yen developed a phage-typing scheme for *Salmonella* Typhi [106]. Following the description of the Vi antigen of the *Salmonella* Typhi, Craigie and Yen discovered bacteriophages that could only attack bacteria possessing the Vi-antigen. The Vi-phages specifically attack bacteria that possess the Vi-antigen, such as *Salmonella* Typhi, and these phages cannot be absorbed into bacterial cells lacking the Vi-antigen. Four virulent Vi-phages were described and designated I, II, III and IV. These phages were serologically different and had different physical properties. The Vi-phage II showed adaptations that would lyse strains similar to that which it had been last grown. Through successive adaptation of the Vi-antigen II, *Salmonella* Typhi strains could be further subdivided. Thirty three internationally recognized Vi types have been recorded. The Vi-phage typing scheme for *Salmonella* Typhi was published by Craigie and Yen in 1938 and has since become the method of choice for the epidemiological subtyping of *Salmonella* Typhi isolates [106, 107].

The advantage of phage typing resides in the simplicity of its implementation, which requires only basic laboratory equipment [8]. However, the ambiguous lysis reactions are common drawbacks, and careful coordination between reference laboratories is required in order to ensure reproducibility of the assay. The method is also limited by the number of available phages [8]

1.5.3 Antimicrobial treatment and acquired resistance

Typhoid fever can be fatal if untreated [12]. Considering the risk of relapse and chronic carriage, it is essential to treat all typhoid fever cases [108]. It is recommended that patients be treated with antimicrobials as soon as a diagnosis has been made rather than after the availability of antimicrobial susceptibility test results. Treatment with appropriate antimicrobials is essential to prevent typhoid fever case fatalities and spread [12, 108-110].

1.5.3.1 Chloramphenicol treatment

Following its availability in 1948, chloramphenicol became widely used as the treatment of choice for typhoid fever infections [111-115]. Chloramphenicol treatment of typhoid fever during this time resulted in a dramatic decrease in illness and mortality associated with typhoid fever [115]. The antimicrobial effect of chloramphenicol comes from its ability to cause prompt and dramatic inhibition of protein synthesis in drug sensitive isolates [115]. It is a metabolic antagonist as it prevents protein chain elongation by inhibiting the peptidyl transferase activity of the bacterial ribosome. Chloramphenicol does not block the progression of the growing peptide but interferes directly with substrate binding [116].

Most *Salmonella* Typhi isolates remained sensitive to chloramphenicol treatment throughout the 1950s to the 1960s. Even though there were sporadic reports of typhoid fever resistance to chloramphenicol [117-119], it was not up until 1972 that this resistance became a public health problem, as epidemics caused by chloramphenicol resistant *Salmonella* Typhi were reported in Mexico and India [110, 118, 120, 121]. Subsequently, epidemics caused by chloramphenicol resistant *Salmonella* Typhi were reported in Vietnam, Korea and Peru [108, 121]. Most of the resistant isolates not only showed resistance to chloramphenicol, but also showed resistance to tetracycline, aminoglycosides and sulphonamides [121]. Chloramphenicol resistance in these

isolates was encoded by a plasmid that belonged to the incompatibility group H (IncH). Streptomycin, sulphonamides and tetracycline-resistance were also carried on this plasmid [108, 122].

1.5.3.2 Ampicillin treatment

Following the introduction of ampicillin as treatment of choice for chloramphenicol resistant *Salmonella* Typhi, epidemics caused by *Salmonella* Typhi isolates resistant to ampicillin and chloramphenicol, with additional resistance to streptomycin, sulphonamides and tetracyclines were reported in developing countries including Pakistan and India [29, 108, 123]. Such multi-drug resistant (MDR) isolates were subsequently isolated in Vietnam and in some countries in Africa including South Africa, Egypt and Kenya [29, 85, 124, 125].

1.5.3.3 Co-trimoxazole treatment

In 1981, following the introduction of co-trimoxazole as treatment for MDR *Salmonella* Typhi, plasmid mediated resistance to chloramphenicol and trimethoprim was identified in sporadic *Salmonella* Typhi isolates [122].

1.5.3.4 Multi-drug resistant *Salmonella* Typhi

The emergence of multi-drug resistant (MDR) *Salmonella* Typhi has become a global health problem with serious impact on treatment of typhoid fever [126]. MDR *Salmonella* Typhi harbour incHI1 incompatibility type plasmid. Wain and colleagues (2003) described two incHI1 incompatibility plasmids responsible for the MDR phenotype of *Salmonella* Typhi, including the pHCM1 that encodes transferable multiple antibiotic resistance and the R27 plasmid, a precursor of pHCM1 that encodes tetracycline resistance. The pMCH1 plasmids contain coding sequences (CDS) that are involved in resistance to antimicrobial agents. These include coding sequences which represent the transposable genetic element (Tn9) which determines chloramphenicol resistance, coding sequences which represent the Tn10 element associated with tetracycline resistance, coding sequences which are homogenous to the sulphonamide resistance gene, coding sequences which are identical to the trimethoprim resistance gene and coding sequences which are putative β-lactamases [126].

1.5.3.5 Ciprofloxacin treatment

Following reports of MDR *Salmonella* Typhi, ciprofloxacin, a class of fluoroquinolones, became the treatment of choice for typhoid fever [12, 127]. Fluoroquinolones target the bacterial topoisomerases, in particular DNA gyrase protein (GyrA) thereby inhibiting DNA replication [128].

Chromosomal mediated resistance is emerging as a result of selective pressure on the bacterial population due to uncontrolled use of fluoroquinolones [129]. In most cases, fluoroquinolone resistance is chromosomally encoded and is associated with single point mutations that alter the amino acid structure of DNA gyrase. This mutation occurs in the A subunit of DNA gyrase (*gyrA*) known as the quinolone-resistance determining-region (QRDR). In *Salmonella* Typhi isolates, a single point mutation conferring an amino acid substitution at codons Serine 83 (Ser83 →Phe) or aspartate 87 (Asp87 → Gly or Tyr) of the gene encoding GyrA protein (*gyrA*) can confer resistance to nalidixic acid [129, 130]. The evolution of fluoroquinolone resistance is often preceded by mutations which induce nalidixic acid-resistance; therefore, reduced susceptibilities to fluoroquinolones can be predicted by the detection of resistance to nalidixic acid [128, 131]. The accumulation of more mutations in DNA gyrase genes may result in an increase in the minimum inhibitory concentration (MIC) for fluoroquinolone [132].

Currently, the rate of nalidixic acid resistance in *Salmonella* Typhi is increasing and there are an increasing number of reports on the emergence of *Salmonella* Typhi with decreased susceptibility to ciprofloxacin [29, 127, 131-133]. Recent reports implicate *Salmonella* Typhi haplotype H58 (H58 *Salmonella* Typhi) as the single clone that has been mostly associated with nalidixic acid resistance [132]. The H58 *Salmonella* Typhi lineage is a highly clonal haplotype of *Salmonella* Typhi that possesses the IncHI1 MDR plasmid as well as point mutations conferring resistance to quinolones [132]. This H58 *Salmonella* Typhi lineage, which first emerged in Southeast Asia, has now become the causative agent of most typhoid fever outbreaks [132, 134]. In addition, the H58 *Salmonella* Typhi has disseminated into parts of Africa and is spreading rapidly [29, 132, 133]. A study done in Kenya has shown that this

haplotype has replaced the antimicrobial-susceptible strains of *Salmonella* Typhi and has been associated with an increase in the incidence of MDR *Salmonella* Typhi [29]. From the phylogenetic analysis of a global collection of *Salmonella* Typhi isolates from 63 countries based on whole-genome sequences, Wong and colleagues (2015) showed cases of closely related *Salmonella* Typhi isolates from different countries clustering together, indicating the likelihood of spread of *Salmonella* Typhi isolates from one African country to another [132]. This is depicted in figure 1.4 which shows H58 *Salmonella* Typhi isolates from South Africa and Malawi clustering closely together. In their study, Wong and colleagues (2015) indicated a major on-going clonal replacement of non-H58 *Salmonella* Typhi haplotypes with this H58 *Salmonella* Typhi clade since 1990. High predominance of this lineage has also been seen in Eastern and Southern Africa [132].

Currently, treatment options for typhoid fever with reduced susceptibility to fluoroquinolones, are ceftriaxone (extended-spectrum cephalosporin) and azithromycin (macrolide) [135].

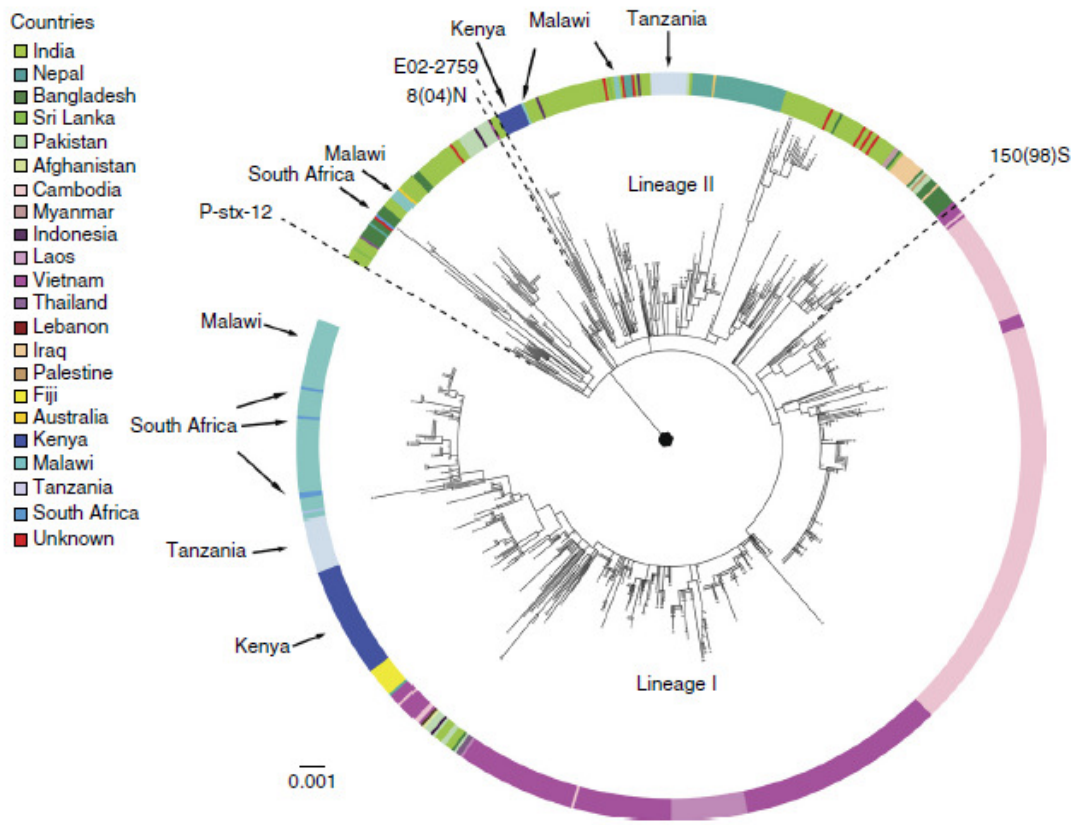


Figure 1.3 Population structure of the *Salmonella* Typhi H58 lineage and distribution across countries in the Sub-Saharan African continent [132]

1.6 Molecular characterization

1.6.1 The genome of *Salmonella* Typhi

The genomic features of individual species, subspecies, serovars and different isolates can be scrutinised using sequencing [136]. Through genomic sequencing, the genetic blue print of bacteria can be examined and compared to genomes of bacteria of similar phenotypic traits [136]. Chromosomal scrutiny of *Salmonella* Typhi strains CT18 and TY2 have indicated an incredible degree of conservation amongst *Salmonella* Typhi isolates [136]. It is believed that *Salmonella* Typhi belongs to a single clone that evolved from the same progenitor. The low rate of variation of single-nucleotide polymorphisms in *Salmonella* Typhi advocates the notion that the bacterium may be as young as 30 000 years and thereby denoting that the *Salmonella* Typhi diverged significantly later than the estimated divergence of *Salmonella* enterica and *E. coli*. The differentiation of *Salmonella* Typhi isolates is a very cumbersome process because of the high homogenous characteristic of the organism. Several genotypic methods have been

implemented in attempt to differentiate the homogenous *Salmonella* Typhi isolates. There is still a need for newer genetic tools that will improve the abilities to examine the genomic structure of *Salmonella* Typhi isolates [136].

1.6.2 Genotyping *Salmonella* Typhi

Molecular characterization, in particular of *Salmonella* Typhi isolates, plays a very critical role in public health management [137]. In order to understand the factors that contribute to the epidemiology of *Salmonella* Typhi in endemic areas, it is important to propose new and efficient methods which can be used for rapid characterization of *Salmonella* Typhi isolates. Molecular subtyping methods, in particular the application of molecular markers, have become an integral part of epidemiological investigations of *Salmonella* Typhi infections. These methods are most valuable in surveillance for improving treatment, control and prevention measures and during outbreak investigations [11, 138]. Molecular subtyping techniques offer rapid, robust, portable, sensitive and objective results and can be used to differentiate epidemiologically unrelated *Salmonella* Typhi isolates [137].

Several molecular subtyping methods have been explored in order to differentiate homogenous *Salmonella* Typhi isolates. PFGE, Multi-locus sequence-typing (MLST), whole-genome sequencing (WGS) and multiple-locus variable-number tandem-repeats analysis (MLVA) are a few techniques commonly used to characterize *Salmonella* Typhi at molecular level [137].

1.6.3 Pulsed-field gel-electrophoresis

Macro-restriction of chromosomal DNA using PFGE is commonly used for the differentiation of isolates at DNA-level [8, 137]. This technique has been employed by PulseNet international as a primary methodology for tracking foodborne infections world-wide. PFGE has been used successfully to perform comparative chromosomal DNA analysis amongst *Salmonella* Typhi isolates and has been used successfully in typhoid fever outbreak investigations [87, 139-142]. The method relies on the digestion of chromosomal or plasmid DNA by rare cutting restriction endonucleases. The resulting large DNA fragments (greater than 50 000 base pairs) are separated by use of periodic changes in the direction of the electric field during electrophoresis.

PFGE has added great value in epidemiological surveillance and outbreak investigations of pathogenic isolates and has been used frequently for bacterial isolate characterization [10, 142]. The validity of PFGE has been well established for the most common bacterial pathogens including *Salmonella*, *Shigella*, *E. coli*, *Campylobacter*, *Clostridium*, *Cronobacter* and *Listeria*. The standardization of the PFGE protocols for these pathogens by PulseNet International has enhanced the ability to compare fingerprint data between laboratories [10, 137]. PulseNet International has also established the international fingerprinting database from which emerging outbreaks can be detected and from where the spread of pathogenic bacteria can be monitored [7]. The greatest strength of PFGE is that it is has good discriminatory power [137, 143]. The PFGE approach scrutinizes the whole bacterial genome, for this reason, insertion and deletion mutations resulting from environmental pressure and from recombination can alter the genome by introducing new restriction sites which will, as a result, alter the PFGE pattern of the specific isolate [137, 143]. The acquisition of mobile genetic elements, such as resistance plasmid, may also introduce new restriction sites thereby altering the PFGE pattern of the isolate [143].

The successful use of PFGE in outbreak investigations of typhoid fever has been documented [10, 89, 139-142]. In 2010, Smith and colleagues used PFGE to study and track a typhoid fever outbreak from South Africa to Australia [89]. A cluster of six typhoid fever cases were reported from patients who had previously visited a common restaurant. PFGE analysis of these isolates revealed an indistinguishable PFGE fingerprint type (pulsotypes) indicating that the outbreak was caused by a single *Salmonella* Typhi isolate. Epidemiological investigations revealed an additional typhoid fever case from a Canadian barman who had worked at the same restaurant while on a working holiday in South Africa. The barman, who was later identified as the index case, departed from South Africa and journeyed to Australia where he was diagnosed and treated with typhoid fever on arrival. The PFGE pattern of this *Salmonella* Typhi isolate was obtained through the PulseNet International networks and was later found to be indistinguishable to the isolates from the Pretoria outbreak [89].

Although PFGE has been successfully used in outbreak detection of *Salmonella* isolates, the subtyping capabilities of this technique are limited in *Salmonella* Typhi isolates [144, 145]. Studies have shown that *Salmonella* Typhi frequently undergoes changes in the chromosomal

gene order [146]. Chromosomal rearrangement involves the movement or reshuffling of DNA throughout the genome, from one location to another. The chromosome in *Salmonella* Typhi show significant rearrangement in gene order in the wild-type isolates [146]. Therefore, a single *Salmonella* Typhi isolate may have multiple PFGE patterns [145]. This phenomenon was demonstrated in a study by Echieta and Usera (1998) were 85 *Salmonella* Typhi isolates belonging to eight different typhoid fever outbreaks that occurred in Spain between 1989 and 1994 were analysed. In five of the eight outbreaks, the strains were indistinguishable by PFGE. However, the remaining three outbreaks presented two PFGE patterns. The detection of two PFGE types in the very same outbreak isolate was due to chromosomal rearrangements [145].

Other major drawbacks with PFGE are that the technique is technically demanding, time-consuming and labour-intensive [137]. Additionally, the technique does not display equal sensitivity with different bacterial species. Analysis of the PFGE patters can suffer from some subjectivity, making it prone to errors [137].

1.6.4 Multi-locus sequence-typing

Multi-locus sequence-typing (MLST) is a well-established technique aimed at exploiting the unambiguous nature of nucleotide sequence data for bacterial characterization [10, 135, 147]. MLST has become a common tool used for epidemiological studies and for the analysis of the molecular evolution of pathogens [17, 135, 137]. The technique is based on the analysis of sequence data from a selection of house-keeping genes [143, 147]. The alleles from the house-keeping gene are assigned allele numbers based on a complete match to alleles in the global database. The combination of these allele numbers from the selected house-keeping genes makeup a sequence type and a clonal complex [143, 147].

In the year 2002, Kidgell and colleagues published an MLST scheme for studying the clonality of *Salmonella* Typhi isolates [19]. This MLST scheme involved PCR amplification of fragments (432 to 501 bp) from a selection of 7 housekeeping genes including *aroC* (chorismate synthase), *dnaN* (DNA polymerase III beta subunit), *hemD* (uroporphyrinogen III cosynthase), *hisD* (histidinol dehydrogenase), *pure* (phosphoribosylaminoimidazole carboxylase), *sucA* (alpha ketoglutarate dehydrogenase) and *thrA* (aspartokinase+homoserine

dehydrogenase) [8, 10, 19]. These 7 housekeeping genes were selected from the *Salmonella* Typhi CT18 genome based on their scattered position, that they are flanked by genes with known functions and that the selected housekeeping genes as well as their flanking regions are selectively neutral. Following amplification of the 7 housekeeping genes, the PCR products were purified and resulting fragments subjected to nucleotide sequencing [10, 19]. The trimmed sequences from the 7 housekeeping genes were submitted to a publically accessible database (http://mlst.warwick.ac.uk/mlst/dbs/Senterica) which currently contains more than 2300 isolates that have been characterized [17]. For each locus, alleles were assigned arbitrary numbers and based on the combination of allele numbers from all 7 housekeeping genes (allelic profile) the sequence type was determined [137, 148].

The success of MLST in discriminating bacterial isolates has been shown to be dependent on the number of the housekeeping genes selected [8, 143]. In a study conducted by Fakhr and colleagues (2005), MLST analysis of 85 *Salmonella enterica* subspecies *enterica* serovar Typhimurium (*Salmonella* Typhimurium) isolates using only 4 housekeeping genes including *manB* (phosphomannomutase), *glnA* (glutamasesynthetase), *pduF* (1-2-propanediol utilization factor) and *spaM* (virulence gene), showed no diversity amongst the isolates [10, 143]. On the contrary, an MLST scheme consisting of 7 housekeeping genes was successful in discriminating *Salmonella* Typhimurium isolates and identifying the highly invasive MLST sequence type (ST) 313 clone, which has emerged as a huge public health problem in sub-Saharan Africa [148].

Even so, MLST has been found to be poorly suited for the discrimination of *Salmonella* Typhi isolates [19, 149]. In a study conducted by Dahiya and colleagues (2013), MLST analysis of 30 *Salmonella* Typhi isolates using seven housekeeping genes distinguished the isolates into 2 MLST sequence types, namely ST1 and ST2 [149]. Kidgell and colleagues (2002), used seven housekeeping genes to distinguish 26 *Salmonella* Typhi isolates into 3 MLST sequence types including ST1, ST3 and ST8[19]. Since *Salmonella* Typhi is homogenous, MLST analysis of this pathogen is uninformative [135].

PFGE has been found to be more discriminatory than MLST [10, 143]. MLST analysis is limited only to parts of a series of genes while PFGE screens the entire genome. With MLST analysis, little or no variation in nucleotide sequences outside of the housekeeping genic region can render no discrimination between isolates. Yet, insertions, deletions and plasmid presence can alter PFGE profile resulting in diversity in PFGE patterns [10, 143].

Since MLST is sequenced-based, it is considered to be objective and reproducible, and allows for inter-laboratory comparison [10, 148]. The ability for MLST to detect slowly accumulating mutations within a bacterial species renders this technique a valuable tool for international and national surveillance [149]. However, due to its labour-intensive, time-consuming and low discriminatory power for *Salmonella* Typhi, MLST is not suitable for epidemiological typhoid fever surveillance and outbreak investigations [10, 150].

1.6.5 Multiple-locus variable-number tandem-repeats analysis

In recent years, variable-number tandem-repeats (VNTR) have been used with increasing frequency for the molecular subtyping of pathogenic bacteria [151-154]. These are short nucleotide sequences organized as tandem repeats at specific loci. These repeated sequences are widespread throughout the genome of bacteria and can vary in repeat number from strain to strain. VNTR are a good source of polymorphism believed to have resulted mostly from slipped-strand mispairing during replication. It has been reported that insertion and deletion mutation are more prevalent in DNA templates that contain repeated sequences. Replication slippage can be induced at a repeat sequence when the template strand and its copy shift their relative position leading to part of the template being copied twice or missed out. The resulting template is a new polynucleotide strand with larger or smaller number of repeat units. Thus, replication slippage generates new length variants, adding to the collection of alleles already present in the bacterial population hence making VNTR highly diverse. Variation within VNTR has also been attributed to DNA recombination between homogenous repeat sequences [151-154].

Multiple-locus variable-number tandem-repeats analysis (MLVA) is an assay in which several VNTR loci can be combined to generate strain specific profiles used to discriminate genetically

homogenous strains [155]. The MLVA typing of VNTR is rapid, reliable and can provide greater discriminatory capacity as VNTR are highly diverse [155]. The MLVA procedure is a PCR-based method involving amplification of several VNTR loci, followed by fragment size determination using capillary electrophoresis with an internal size standard and then determination of the actual number of repeats at each locus [156].

MLVA has been used successfully and has been very effective in subtyping genetically homogenous strains [157-160]. Recently, there has been great success in the development of highly-discriminatory MLVA assays for subtyping *Salmonella* enterica serovars including *Salmonella* Typhimurium and *Salmonella enterica* subspecies *enterica* serovar Enteritidis (*Salmonella* Enteritidis) [151-154]. In a recent report by Liu and colleagues (2016), MLVA was used to study the molecular epidemiology of *Salmonella* Enteritidis infection in the Guangdong province, China [161]. A seven-locus MLVA assay was used to characterize 147 sporadic *Salmonella* Enteritidis isolates into 33 MLVA types showing higher discrimination as compared to PFGE from which 29 pulsotypes were identified. Liu and colleagues (2016) also used the MLVA assay to detect outbreaks and find their sources successfully. Seven *Salmonella* Enteritidis isolates collected during the year 2014 were found to possess a common MLVA type. Epidemiological investigations determined this to be an outbreak of *Salmonella* Enteritidis infection with six epidemiologically related isolates. Liu and colleagues (2016) determined the source of infection to be a bread countertop prepared with contaminated cakes [161].

Wuyts and colleagues (2013) evaluated the added value of MLVA typing for surveillance and outbreak detection of *Salmonella* Typhimurium isolates [162]. Isolates in this study were collected from Belgium where *Salmonella* Typhimurium has been identified as the most frequently isolated *Salmonella* serovar from humans with an average of 1,985 isolates per year. Of the 5,698 isolates collected between the years 2010 and 2012, approximately 1,420 isolates were randomly selected for analysis. A five-locus MLVA assay characterized the isolates into 414 MLVA types. The MLVA assay identified 30 MLVA types that were common throughout the three-year period, and also identified unique MLVA types in each year of study. MLVA

showed high discriminatory power and showed the ability to improve public health surveillance [162].

1.6.6 Whole-genome sequencing

Whole-genome sequencing (WGS) technique has revolutionized molecular typing of pathogenic bacteria by providing rapid and accurate identification of variation in the bacterial genome [137, 163-165]. This has become the most powerful and most highly recommended tool for epidemiological investigation [163, 166]. The most remarkable aspect of WGS is its ability to combine typing results which could have been obtained from several different typing techniques in order to give accurate discrimination of bacterial pathogens [163].

Contrary to the Sanger sequencing technique where genome sequencing projects took years before completion, the current WGS techniques provide rapid and high throughput technology [167]. WGS technique involves extraction of bacterial genomic DNA, which can be done using commercially available kits. It uses a shotgun sequencing approach for the sequencing of large genomic DNA fragments [168]. Shotgun sequencing involves enzymatic or mechanical breakdown of DNA into multiple segments that are sequenced using the chain termination method to produce short reads (500–700bp). High-throughput data in form of short reads are assembled based on overlapping regions to form contigs. Contigs are joined together to form a complete sequence [165, 168, 169].

Through sequencing of one bacterial genome, a large amount of data can be obtained including antimicrobial resistance profile, MLST sequence types, virulence genes, serotype prediction as well as single-nucleotide polymorphisms (SNP) [165, 168]. WGS provides a complete unambiguous typing of different bacterial pathogens due to its potential to resolve single-base differences between genomes [165, 168].

There are high-throughput sequencing platforms developed for WGS analysis with the capacity for large scale DNA sequencing [170]. Through the use of these platforms, sequence data can be determined from amplified DNA fragments without the need for cloning of DNA fragments.

There are several numbers of WGS platforms that are commercially available, including the Illumina Solexa technology, ABI SOLiD and the Ion torrent technology. These platforms differ on base read-length generated as well as error rates. The Illumina/Solexa genome analyzer (Figure 1.5) is one of the most commonly used platforms [167, 170].

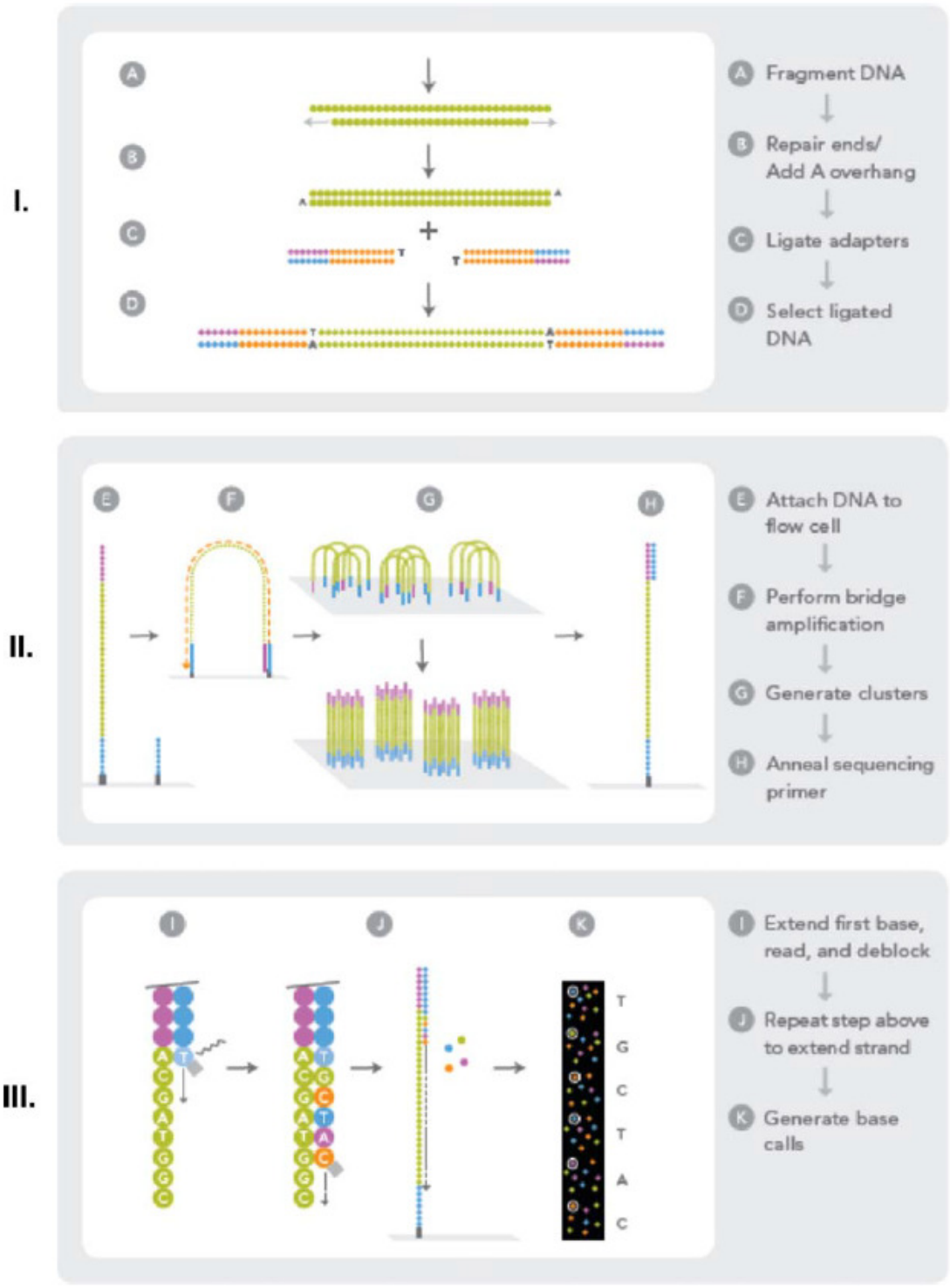


Figure 1.4 Outline of the whole genome sequencing workflow (Illumina platform). (I) Following genomic DNA fragmentation, (II) the library is applied on to a solid surface of a flow cell. The attached fragments form bridge molecules which are amplified by means of isothermal amplification resulting in a cluster of identical fragments. The DNA fragments are amplified following denaturation and annealing of sequencing primers. (III) The amplified DNA fragments are subjected to sequencing-by-synthesis using 3'labelled nucleotides[167].

1.6.6.1 Analysis of whole genome sequence data

1.6.6.1.1 Whole genome MLST

Phylogenetic analysis of whole-genome sequencing data includes MLST and single nucleotide polymorphism (SNP) analysis [171, 172]. Contrary to the traditional MLST which begins with a PCR amplification step using primers that are specific for the targeted house-keeping genes and followed by Sanger sequencing of each gene, whole-genome MLST (wgMLST) provides a more discriminatory and rapid alternative [171]. WgMLST is regarded as the gene-by-gene approach and involves *in silico* comparison of whole genome allelic variants between isolates. The added advantage of using wgMLST is that the number of targeted loci is not limited [173]. Analysis of the large amount of data generated through *in silico* MLST can be performed using user-friendly software and tools available on the Center for Genomics Epidemiology (http://www.genomicepidemiology.org) server as well as the CLC Genomics workbench tool [174].

1.6.6.1.2 Whole-genome SNP

SNPs represent the most abundant source of genetic variation within bacterial genomes and are considered good markers to study diversity [174]. SNPs occur as a result of substitution mutations as they emerge in strains within the same serotype [174, 175]. Contrary to the taxonomically informative or canonical SNP-based approaches, whole-genome SNP (wg-SNP) analysis provides a robust and unbiased method to resolve closely related species [175]. SNPs from whole-bacterial genomes can be identified using online pipelines, such as the genobox available on the Centre for Genomic Epidemiology server. Once the SNPs have been identified, genetic diversity of strains can be analysed through the construction of phylogenetic SNP trees [175]. WgSNP analysis has been found to be a superior method for clustering outbreak related isolates of *Salmonella* as better resolution between outbreak and sporadic isolates has been observed [172, 176, 177].

Taylor and colleagues (2015), showed the ability of WGS to distinguish between epidemiologically linked and unrelated isolates [172]. In their study, Taylor and colleagues (2015) analysed 28 *Salmonella* Enteritidis isolates from seven distinct food-borne outbreaks as

well as 27 sporadic isolates. SNP-based phylogenetic tree analysis showed all isolates within the same outbreak to be closely related with zero to three SNP differences identified between the isolates. Clusters of outbreak isolates differed from the nearest non-outbreak isolates by an average of 42 SNPs and they differed from sporadic isolates by an average of 60 SNPs. Furthermore, wgSNP analysis was able to distinguish between isolates from different outbreaks thus showing that the outbreaks did not originate from a common source [172].

WGS has proven to be a superior molecular method used to distinguish between epidemiologically linked and unrelated isolates. However, the challenge that still remains with WGS is the requirement of skilled bioinformatics experts to aid in extraction and analysis of data that is important for clinical microbiology, infection control and public health [178]. In order for WGS to be used for routine diagnostics and surveillance, sequence data needs to be transformed to clinically relevant information which can be easily understood by public health professionals and clinicians with limited bioinformatics skills [137, 163, 178, 179]. Although online analysis tools are available for use by non-bioinformatics experts, these tools only provide data similar to PFGE, MLST, serotyping, antimicrobial resistance profiling, virulence genes and SNP determination. The ability to formulate novel research questions is tremendously diminished without the skills provided by bioinformatics experts. With the skills of bioinformatics experts, data produced by WGS can be invaluable for developments of new subtyping techniques and conventional typing methods can be optimized [137, 163, 178, 179].

1.7 The importance of studying the molecular epidemiology of *Salmonella* Typhi in the sub-Saharan African region

Typhoid fever has proven to be a public health problem in Africa [29, 49, 78, 84, 85]. More concerning is the increase in the number of isolates showing resistance to current treatment regimes. Currently, fluoroquinolones are recommended for treatment of typhoid fever, but the emergence of the super-strain H58 *Salmonella* Typhi and the spread thereof has made treatment of typhoid fever a challenge [87, 180, 181]. The H58 *Salmonella* Typhi has become a huge public health problem in Africa, yet very little is known about the emergence, evolution and transmission of the H58 lineage across Africa. Of great concern is the on-going H58 *Salmonella* Typhi lineage epidemic that has been reported across Africa [132]. Several studies recommend the use of Azithromycin which has proven to be effective against fluoroquinolone resistant *Salmonella* Typhi [182, 183]. Even so, reports of Azithromycin resistant *Salmonella* Typhi are emerging [184, 185]. Increasing resistance of *Salmonella* Typhi isolates to newly used antimicrobials presents a public health challenge. Fundamental stringent control measures within SSA countries are required [14]. Undoubtedly, the use of molecular techniques can provide the ultimate method to discriminate these isolates. Furthermore, through the use of standardized molecular methods, the genetic makeup of the organism could be compared within the countries in the continent providing information to model transmission dynamics and inform vaccine trial to control dissemination from one country to another [14, 132]. MLVA has proven to be a highly discriminatory yet affordable molecular method which has potential to be used in African countries for subtyping *Salmonella* Typhi isolates [155].

1.8 Aim

The aim of the research presented in this dissertation was to design a highly-discriminatory MLVA typing assay for the analysis of *Salmonella* Typhi strains. The assay was used to study the population structure of these *Salmonella* Typhi isolates (including *Salmonella* Typhi H58) and compare the discriminatory capacity of MLVA with that of PFGE.

1.8.1 Specific objectives

a) To use VNTR markers to explore their potential in the characterization of *Salmonella* Typhi isolates
b) To establish a highly-discriminatory MLVA assay consisting of five VNTR markers for the typing for *Salmonella* Typhi isolates
c) To compare MLVA genotyping data to the commonly used PFGE genotyping data
d) PCR screen *Salmonella* Typhi isolates for *Salmonella* Typhi H58-specific mutations
e) To investigate the whether H58 *Salmonella* Typhi is associated with particular a MLVA profile

Chapter 2

General materials and methods

2.1 Surveillance programs

2.1.1 National Surveillance

The Centre for Enteric Diseases (CED) of the National Institute for Communicable Diseases (NICD), a division of the National Health Laboratory Services (NHLS), serves as a reference centre for enteric pathogens (including *Salmonella*) in South Africa. There are approximately 200 laboratories (NHLS and Private) across the country that participates in the national laboratory surveillance through the Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa (GERMS-SA). For each year of surveillance demographic details are captured for all isolates, including information about the patient, isolate collection date and site, and the information is stored in a centralized GERMS-SA Microsoft Access database [186]. All laboratories submit *Salmonella* isolates to the CED on Dorset Egg transport media. The case definition for typhoid fever is the isolation of *Salmonella* Typhi from any normally-sterile sites (e.g. blood, cerebrospinal fluid (CSF) and tissue) as well as gastrointestinal sites (e.g. stools and rectal swabs) by positive culture.

2.1.2 Typhoid Fever Surveillance in Africa Program (TSAP)

Typhoid Fever Surveillance Program (TSAP) in Africa is a multi-country surveillance network aimed at establishing the true burden of typhoid fever in the African continent [187]. The TSAP study was conducted in public healthcare facilities in Burkina Faso, Ethiopia, Ghana, Guinea Bissau, Kenya, Madagascar, Senegal, South Africa, Sudan and Tanzania [187, 188]. The case definition for typhoid fever includes positive blood culture from patients presenting to sentinel sites around Africa with a history of fever for 72 hours or a recorded axillary temperature of >37.5°C or rectal 38°C [189]. Participating sites submit a subculture *Salmonella* Typhi to the CED onto Cary & Blair transport media.

2.2 Phenotypic characterization of *Salmonella* Typhi isolates

Upon receipt at the CED, all *Salmonella* isolates from transport media (Dorset Egg and Cary & Blair) were subcultured onto non-selective 5% blood agar plates (Diagnostic Media Products (DMP), Sandringham, South Africa) using sterile loop. The blood agar plates were incubated at 37°C for 18-24 hours. The automated VITEK-2 compact system (BioMérieux, Marcy L'Étoile, France) was used to confirm biochemical identification.

Specific anti-sera (Statens Serum Institut, Copenhagen, Denmark; Remel Europe Ltd, Dartford, Kent, UK; and BioMérieux, Marcy-I'Ètoile, France) were used to serotype *Salmonella* Typhi isolates according to the Kauffman-Le Minor scheme [104]. Pure heavy growths of *Salmonella* Typhi were frozen in 10% skim milk (DMP) at -70°C for further analysis.

2.3 Crude genomic DNA extraction

Bacterial isolates were cultured on 5% blood agar (Diagnostic Media Products) and incubated overnight at 37°C. Half a loopful of bacterial culture was suspended in autoclaved TE Buffer (10 mM Tris: 1 mM EDTA, pH 8.0) and incubated for 25 minutes at 95°C. The boiled suspensions were centrifuged in a table-top centrifuge at 12000 rpm for 3 minutes. Approximately 50 µl of the supernatant was transferred into a clean 1.5 ml Eppendorf tube and were used as the template DNA for PCR assays. The supernatants were stored in the freezer at -20°C.

2.4 Real-time PCR serotyping of *Salmonella* Typhi isolates

The real-time PCR assay targeted DNA sequences that were unique to *Salmonella* species as well as sequences unique to *Salmonella* Typhi. The ttrRSBCA gene, which is located near the *Salmonella* pathogenicity island 2, required for tetrathionate respiration in *Salmonella*, was selected for specific detection of *Salmonella* isolates [103]. For *Salmonella* Typhi, the STY0201 gene, encoding a putative fimbral-like adhesion protein located at position 210,264 in the *Salmonella* Typhi CT18 chromosomal genome (Accession number NC_003198) was selected [102]. Primers and probes targeting these genes were used as previously described (Table 2.1).

Table 2.1 Primer and probe sequences for *Salmonella* species and *Salmonella* Typhi specific genes

Isolate	Gene	PCR primer/probe sequences	Amplicon size	Reference
Salmonella species	ttrRSBCA	Forward (ttr-6): CTCACCAGGAGATTACAACATGG Reverse (ttr-4): AGCTCAGACCAAAAGTGACCATC Probe (ttr-5): 6FAM - CACCGACGGCGAGACCGACTTT-BHQ1	95 bp	[103]
Salmonella Typhi	STY0201	Forward: CGCGAAGTCAGAGTCGACATAG Reverse: AAGACCTCAACGCCGATCAC Probe: Cy5 - CATTTGTTCTGGAGCAGGCTGACGG-BBQ	131 bp	[102]

All PCRs were carried out on an Applied Biosystems 7500 real-time PCR system (Applied Biosystems, Weiterstadt, Germany). Each 50µl PCR reaction contained 25 µl TaqMan Gene Expression Master Mix, 12 µl internal positive control mix and 1µl DNA template. The final primer concentrations were 1µM and 2µM for the *Salmonella* species and *Salmonella* Typhi assays, respectively.

2.5 Pulsed-field gel electrophoresis (PFGE)

PFGE analysis is still commonly used for molecular characterization of most bacterial strains, including molecular subtyping of *Salmonella*. Analysis of *Salmonella* Typhi isolates by PFGE is carried out as part of routine surveillance in the CED and is performed using the standardized PulseNet protocol for *Salmonella*, *E. coli* and *Shigella sonnei* [190].

2.5.1 The PFGE reference standard – *Salmonella* Braenderup (strain H9812)

Normalization and accurate comparison of PFGE gel images required a well-characterized size standard [191]. A *Salmonella* enterica subspecies enterica serotype Braenderup (*Salmonella* Braenderup) strain H9812 was included as a size standard on at least three lanes in all PFGE gels. This strain was digested with *Xba*I restriction enzyme and produced an even distribution of band sizes.

2.5.2 Preparation of bacterial plugs

Overnight bacterial culture was resuspended into 2ml of cell suspension buffer (Appendix A) to a turbidity of approximately 0.7 (MicroScan Turbidity Meter, Dade Behring, West Sacramento, CA). 400µl of the bacterial suspension was mixed with 20µl of Proteinase K (20mg/ml stock) and heated to 37°C. A volume of 280 µl of 1% SeaKem® Gold agarose (Lonza, Rockland, USA) was then added to the bacterial suspension and the mixture was dispensed into plugs moulds and allowed to solidify for 15 min at room temperature.

For bacterial cell lyses, plugs were incubated at 55°C for 2 hours in 5 ml cell lysis buffer (Appendix A). Following cell lysis, plugs were washed twice with preheated dH2O for 15 min at 55°C and four times with Tris-EDTA (TE) buffer (Appendix A) for 15 min at 55°C.

2.5.3 Restriction enzyme digestion of genomic DNA and electrophoresis of DNA fragments

For the restriction enzyme digestion of genomic DNA, 2-2.5 mm slices of agarose plugs were incubated for 15 min at 37°C in 100µl of *Xba*I restriction buffer [Roche Diagnostics GmbH, Mannheim, Germany] (Appendix A). The plugs were then incubated at 37°C for 3 hours in 150µl of restriction enzyme master mix containing 50 units of *Xba*I restriction enzyme (Appendix A). The plugs were placed on a comb in a casting stand and allowed to air-dry for 3-5 min. 1% SeaKem Gold agarose (Appendix A) was poured into the casting stand and the gel was allowed to polymerize for 20 min at room temperature.

PFGE was performed in a CHEF-DR III electrophoresis chamber (Bio-Rad Laboratories, Hercules, CA) filled with 3L of 0.5x Tris-borate EDTA [TBE] (Appendix A) buffer cooled to 14°C. The PFGE was run at 6 V/cm at an angle of 120° with an initial switch time of 2.2 sec, a final switch time of 63.8 sec and a run time of 19 hours.

2.5.4 Staining of agarose gels

The gel was stained in 0.5x TBE buffer containing 1 mg/ml ethidium bromide [EtBr staining solution] (Appendix A) for 20 minutes and destained three times with dH2O at 20 minute intervals. The gel image of PFGE patterns was captured using the Quantity 1-D analysis software (Bio-Rad Laboratories, Hercules, CA) and saved in TIFF format to enable electronic transfer into the BioNumerics software for further analysis.

2.5.5 BioNumerics analysis of PFGE bands

The generated TIFF image was imported into the BioNumerics (version 6.5) Software (Applied Maths, Sint-Martens-Latem, Belgium) for analysis. The analysis of the image involved marking of PFGE bands of the isolates and these were normalized against the *Salmonella* Braenderup strain H9812 PFGE bands. Dendrograms of the patterns were created using the unweighted pair group method with arithmetic averages (UPGMA), with analysis of banding patterns incorporating the Dice-coefficient at an optimization setting of 1.5% and a position tolerance setting of 1.5% for the band migration distance. In this study, a cluster was defined as a group of isolates with indistinguishable PFGE patterns.

Chapter 3

Development and validation of multiple-locus variable-number tandem-repeats analysis for molecular sub-typing of *Salmonella* Typhi isolates

3.1 Introduction

MLVA is a DNA-based molecular subtyping method widely used to study the molecular epidemiology of pathogenic bacteria [192]. This molecular subtyping method has been successfully used for phylogenetic profiling as well as epidemiological investigations of *Salmonella enterica* species. MLVA uses PCR assisted amplification of VNTR loci and fragment size determination of the PCR amplicons by capillary electrophoresis with an internal size standard. This method has been used to study outbreaks as well as epidemiologically unrelated isolates, and has been very effective in typing homologous clones [138, 155].

Numerous approaches have been made in an effort to improve molecular subtyping of *Salmonella* Typhi using MLVA [11, 193-195]. The availability of two fully sequenced *Salmonella* Typhi genomes (CT18 and Ty2) motivated the determination of VNTRs in the *Salmonella* Typhi genome, from which MLVA assays have been designed. In 2003, Lui and colleagues identified five VNTR markers (designated TR1 to TR5) for molecular subtyping of *Salmonella* Typhi isolates from Asian countries. Only three markers (TR1, TR2 and TR3) showed variations amongst 59 *Salmonella* Typhi isolates [11]. In 2004, Ramisse and colleagues identified five other VNTR markers (designated Sal02, Sal06, Sal10, Sal15 and Sal20) which were used together with two previously reported VNTR markers (TR1 and STTR5) to distinguish 27 French *Salmonella* Typhi isolates into twenty-five MLVA profiles [194]. Octavia and Lan (2009), identified two more VNTR markers (TR4500 and TR4699) and together with five previously reported markers (Sal02, Sal16, Sal20, TR1 and TR2) distinguished 73 global *Salmonella* Typhi isolates into 70 MLVA profiles [193]. The latest report published by Tien and colleagues (2011) proposed the use of two novel VNTR markers (Sty2 and Sty3) in addition to nine previously reported VNTR markers (TR4500, TR4600, Sal02, Sal06, Sal16, Sal20, TR1, TR2 and TR5) to establish

accurate genetic relationships amongst closely related *Salmonella* Typhi isolates from Taiwan [195]. Although these MLVA assays were able to distinguish between *Salmonella* Typhi isolates, there is still no standardized set of VNTR loci for the molecular subtyping of homologous *Salmonella* Typhi strains.

3.1.1 The study aim

The aim of this current study was to design a highly-discriminatory MLVA typing assay for the analysis of *Salmonella* Typhi isolates from SSA. The assay will be used to study the population structure of these *Salmonella* Typhi isolates and compare the discriminatory capacity of MLVA with PFGE.

3.1.1.1 Objectives

a. To use VNTR markers to explore their potential in studying the molecular epidemiology of *Salmonella* Typhi isolates
b. To establish a highly-discriminatory MLVA assay consisting of five VNTR markers for the typing for *Salmonella* Typhi isolates

3.2 Materials and Methods

3.2.1 Selection of bacterial isolates for the setting up of MLVA assay

A total of 50 isolates were selected from a potential 1080 *Salmonella* Typhi isolates from the culture collection at the CED (Chapter 2, section 2.1) for the evaluation of the MLVA assay. The isolates in the validation panel were selected to given a good representation of the diverse *Salmonella* Typhi PFGE pulsotypes within the CED database, different specimen collection dates and geographic areas within the SSA region including various provinces in South Africa (Eastern Cape, Free State, Gauteng, Kwa-Zulu Natal, Limpopo, Mpumalanga, Northern Cape, North West and Western Cape provinces) as well as sporadic isolates from Zimbabwe and Ivory Coast. Isolates associated with previous South African *Salmonella* Typhi outbreaks were also included in the panel.

3.2.2 Genotyping *Salmonella* Typhi from sub-Saharan Africa using MLVA

3.2.2.1 Genomic DNA extractions

Crude DNA extraction was performed using the boiling method as described in section 2.3 of chapter 2.

3.2.2.2 Selection of VNTR loci and PCR primers for MLVA assay

For this study, VNTR loci from previously published literature were evaluated as markers to explore their potential in determining the strain relatedness of the *Salmonella* Typhi isolates. A total of 13 VNTR loci (Table 3.2) were identified and selected for analysis. The lengths of the repeat sequences at each VNTR locus ranged from three base pairs (bp) to 26bp. MLVA was performed incorporating previously described VNTR primers [11, 193-195].

3.2.2.3 Screening for length polymorphism of VNTR using simplex PCR

The evaluation of length polymorphism at each VNTR locus began with simplex PCR amplification of the VNTR loci on the 50 *Salmonella* Typhi isolates. PCRs were performed using the Qiagen multiplex PCR kit (Qiagen, Hilden, Germany). The evaluation of each VNTR was carried out using the conventional PCR method and the forward primers were labelled with fluorophores. Each of the forward primers were labelled with one of four fluorophores, including FAM (blue), VIC (green), NED (yellow) or PET (red). Each 25µl reaction contained 12.5µl of the Qiagen master mix, 2.5µl Qiagen Q-solution, l µl of each of the forward and reverse primers, 1µl bacterial lysate suspension and autoclaved distilled water (dH_2O) to adjust the final volume to 25µl (Appendix B). PCR amplification of the VNTR loci was performed in an Applied Biosystem 2720 cycler (Applied Biosystems, Weiterstadt, Germany) and the PCR conditions included an initial denaturation at 95ºC for 15 min followed by 35 cycles of a three step cycle protocol: 94ºC for 60 sec, 55°C for 90 sec and 72ºC for 90 sec, and a final extension at 72ºC for 10 min. PCR amplicons were diluted 1:35 in sterile distilled water. Two microliter aliquots of the dilutions were mixed with 11µl Hi-Di formamide solution (Applied Biosystems) and 0.2µl GeneScan 600 LIZ® size standard v2.0 (Applied Biosystems). These samples were evaluated by capillary electrophoresis on the Applied Biosystems 3500 genetic analyser (Applied Biosystems, Weiterstadt, Germany) and fragments sizes were analysed using the Gene-Mapper Software (Applied Biosystems). DNA fragments were automatically allocated to length bins and alleles were assigned based on the bin fragment sizes. The VNTR fragment sizes (in base pairs) were entered into the BioNumerics 6.5 software (Applied Maths, Sint-Martens-Latem, Belgium) as character values and a dendrogram was constructed using a categorical coefficient with a 1.5 tolerance and an unweighted pair group method (UPGMA). Length polymorphism at each VNTR locus was defined as insertion or deletion of a repeat sequence. For example, VNTR locus TR2 has an 8bp repeat sequence, therefore, a deletion or insertion of 8bp in this VNTR locus would be considered as variation.

3.2.2.4 VNTR measure of diversity

The discriminatory power at each VNTR locus was determined by Simpson's DI. The Simpson's DI does not only depend on the number of alleles present at each locus but also

takes into consideration the equitability with which the alleles are distributed at each locus. VNTR loci with a Simpson's DI closer to 1 are better markers to differentiate the strains for epidemiological purposes [196, 197]. In order to determine the measure of diversity and the degree of polymorphism at each VNTR locus, the Simpson's index of diversity (Simpson's DI) and 95% confidence intervals (CIs) were calculated using an online tool available at the Public Health England (PHE) website (http://www.hpa-bioinformatics.org.uk/cgi-bin/DICI/DICI.pl). The Simpson's DI for the MLVA assay as well as PFGE was calculated and the Wallace coefficient was determined in order to assess the congruence between the MLVA assay and PFGE analysis, via an online tool (https://www.comparingpartitions.info/).

3.2.2.5 Nucleotide sequencing of VNTR loci

Nucleotide sequencing was performed in order to determine the presence of repeat sequences as well as to determine the size of the flanking regions in each of the seven most variable VNTR loci. For each VNTR locus, PCR amplicons of ten *Salmonella* Typhi isolates representing various fragment sizes were selected. These PCR amplicons served as template DNA in a PCR cycle sequencing reaction using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) (Appendix C) and an Applied Biosystems 3500 genetic analyser. DNASTAR Lasergene (version 8.0) software (DNASTAR Inc., Madison, WI, USA) was used to analyse DNA sequences and multiple-sequence alignments were constructed using the BioEdit software package, version 7.2.5 (Ibis Biosciences, Carlsbad, Calif, USA).

3.2.2.6 Development of a 5-loci MLVA assay

The five VNTR loci that were confirmed to harbour repeat sequences, had conserved flanking regions and had the highest degree of polymorphism were selected for the development of the MLVA multiplex PCR assay. Each 25 µl reaction contained 1 µl of the bacterial lysate suspension, 12.5 µl of the Qiagen master mix, 2.5 µl Qiagen Q-solution and primer concentrations as shown on (Table 3.2). The forward primers for TR1, TR2, Sal02, Sal20 and TR4699 were labelled with PET, 6-FAM, 6-FAM, VIC and NED

fluorophores, respectively. PCR amplicons were diluted 1:35 in sterile distilled water and reactions were subsequently carried out as described above (section 3.2.2.3).

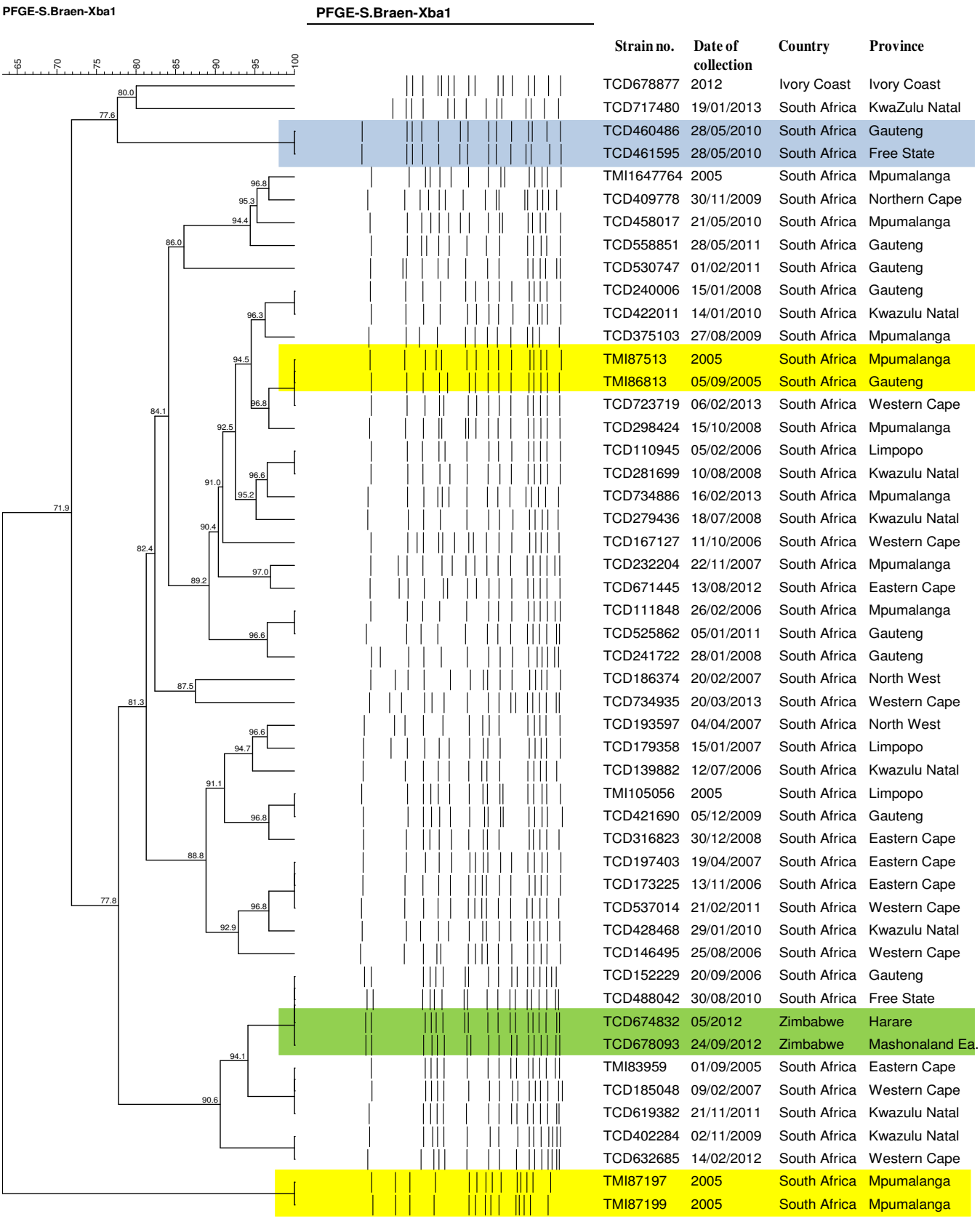
3.3 Results

The MLVA method for subtyping *Salmonella* Typhi isolates was developed for capillary electrophoresis and tested within this study.

3.3.1 Screening for length polymorphism of VNTR loci using simplex PCR

A validation panel consisting of 50 *Salmonella* Typhi isolates was used to evaluate 13 potential VNTR loci. At total of 34 PFGE patterns (Figure 3.1) were represented in the validation panel. The validation panel consisted of two representative isolates from Zimbabwe, one isolate from Ivory Coast and 47 isolates from South Africa. Amongst isolates from South Africa, nine were collected from Gauteng, ten from Mpumalanga, eight from Kwa-Zulu Natal, seven from Western Cape, five from Eastern Cape, three from Limpopo, two from Free State, another two from North West and one isolate from Northern Cape provinces (Figure 3.2). Included in this panel were two representative isolates from Mpumalanga collected during the 2005 typhoid fever outbreak as well as two representative isolates from a typhoid fever outbreak that occurred in Gauteng, 2010.

All of the 13 previously published VNTR loci selected were able to produce a PCR product at an annealing temperature of 55°C. Five VNTR loci were monomorphic and showed no variation amongst the isolates (Table 3.1). The remaining eight VNTR loci were polymorphic and showed variations between the isolates and were further evaluated for inclusion in the MLVA assay.



Strain no.	Date of collection	Country	Province
TCD678877	2012	Ivory Coast	Ivory Coast
TCD717480	19/01/2013	South Africa	KwaZulu Natal
TCD460486	28/05/2010	South Africa	Gauteng
TCD461595	28/05/2010	South Africa	Free State
TMI1647764	2005	South Africa	Mpumalanga
TCD409778	30/11/2009	South Africa	Northern Cape
TCD458017	21/05/2010	South Africa	Mpumalanga
TCD558851	28/05/2011	South Africa	Gauteng
TCD530747	01/02/2011	South Africa	Gauteng
TCD240006	15/01/2008	South Africa	Gauteng
TCD422011	14/01/2010	South Africa	Kwazulu Natal
TCD375103	27/08/2009	South Africa	Mpumalanga
TMI87513	2005	South Africa	Mpumalanga
TMI86813	05/09/2005	South Africa	Gauteng
TCD723719	06/02/2013	South Africa	Western Cape
TCD298424	15/10/2008	South Africa	Mpumalanga
TCD110945	05/02/2006	South Africa	Limpopo
TCD281699	10/08/2008	South Africa	Kwazulu Natal
TCD734886	16/02/2013	South Africa	Mpumalanga
TCD279436	18/07/2008	South Africa	Kwazulu Natal
TCD167127	11/10/2006	South Africa	Western Cape
TCD232204	22/11/2007	South Africa	Mpumalanga
TCD671445	13/08/2012	South Africa	Eastern Cape
TCD111848	26/02/2006	South Africa	Mpumalanga
TCD525862	05/01/2011	South Africa	Gauteng
TCD241722	28/01/2008	South Africa	Gauteng
TCD186374	20/02/2007	South Africa	North West
TCD734935	20/03/2013	South Africa	Western Cape
TCD193597	04/04/2007	South Africa	North West
TCD179358	15/01/2007	South Africa	Limpopo
TCD139882	12/07/2006	South Africa	Kwazulu Natal
TMI105056	2005	South Africa	Limpopo
TCD421690	05/12/2009	South Africa	Gauteng
TCD316823	30/12/2008	South Africa	Eastern Cape
TCD197403	19/04/2007	South Africa	Eastern Cape
TCD173225	13/11/2006	South Africa	Eastern Cape
TCD537014	21/02/2011	South Africa	Western Cape
TCD428468	29/01/2010	South Africa	Kwazulu Natal
TCD146495	25/08/2006	South Africa	Western Cape
TCD152229	20/09/2006	South Africa	Gauteng
TCD488042	30/08/2010	South Africa	Free State
TCD674832	05/2012	Zimbabwe	Harare
TCD678093	24/09/2012	Zimbabwe	Mashonaland Ea.
TMI83959	01/09/2005	South Africa	Eastern Cape
TCD185048	09/02/2007	South Africa	Western Cape
TCD619382	21/11/2011	South Africa	Kwazulu Natal
TCD402284	02/11/2009	South Africa	Kwazulu Natal
TCD632685	14/02/2012	South Africa	Western Cape
TMI87197	2005	South Africa	Mpumalanga
TMI87199	2005	South Africa	Mpumalanga

Figure 3.1 Dendrogram analysis of PFGE patterns of 50 *Salmonella* Typhi strains from sub-Saharan Africa. Highlighted in Yellow are strains isolated from Mpumalanga during the typhoid fever outbreak in 2005. Highlighted in blue are strains isolated during the typhoid fever outbreak in Pretoria, Gauteng 2010. Highlighted in green are strains isolated from Zimbabwe

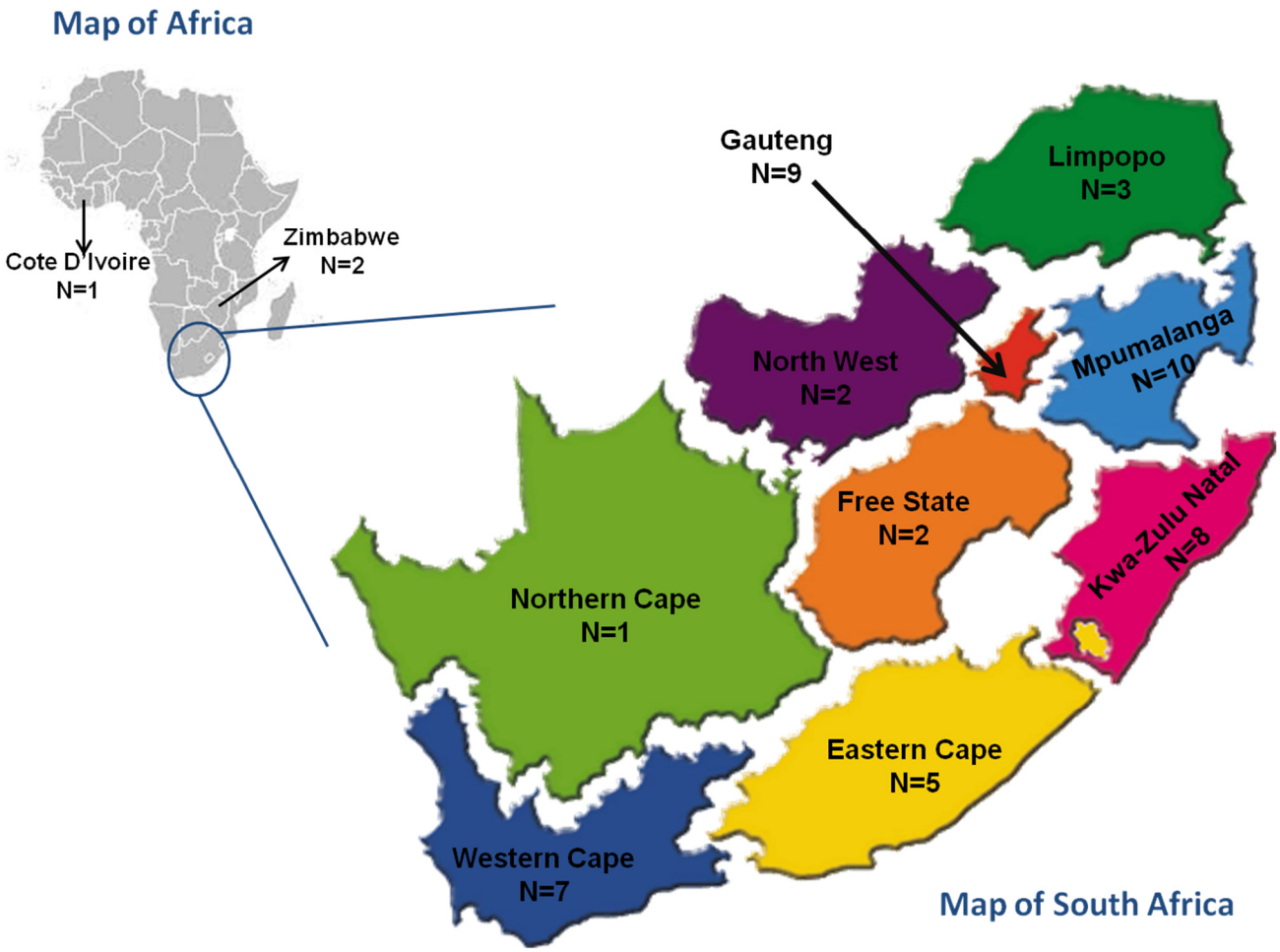


Figure 3.2 Geographical areas in sub-Saharan Africa as source of *Salmonella* Typhi strains included in the development of the MLVA assay

Table 3.1 VNTR loci showing no size variation amongst a panel of 50 *Salmonella* Typhi strains

Isolate #	Strain number	VNTR locus TR4	VNTR locus TR5	VNTR locus Sal06	VNTR locus Sal10	VNTR locus Sal15
1	TMI83959	418	173	171	190	187
2	TMI86813	418	174	171	189	187
3	TMI87197	418	174	171	190	187
4	TMI87199	418	174	171	190	187
5	TMI87513	418	174	171	190	187
6	TCD105056	418	174	171	190	187
7	TCD110945	418	174	171	190	187
8	TCD111848	418	174	171	190	187
9	TCD139882	418	174	171	189	187
10	TCD146495	418	174	171	190	187
11	TCD152229	419	175	171	190	187
12	TCD167127	419	174	170	189	187
13	TCD173225	419	175	170	190	187
14	TCD179358	419	175	170	190	187
15	TCD186374	419	174	172	191	187
16	TCD185048	419	174	170	190	187
17	TCD193597	419	175	172	191	187
18	TCD197403	419	174	178	191	187
19	TCD232204	419	175	171	190	187
20	TCD241722	419	174	171	191	187
21	TCD240006	419	174	170	191	187
22	TCD281699	419	174	171	191	187
23	TCD279436	419	174	171	191	187
24	TCD298424	420	174	171	190	187
25	TCD316823	419	175	178	192	187
26	TCD375103	419	174	172	191	187
27	TCD402284	419	175	172	192	187
28	TCD409778	419	175	171	191	187
29	TCD421690	419	175	172	192	187
30	TCD422011	419	175	178	192	187
31	TCD428468	419	175	172	192	187
32	TCD458017	419	175	172	192	187
33	TCD460486	419	175	172	192	187
34	TCD461595	419	175	172	192	187
35	TCD488042	419	175	172	192	187
36	TCD525862	419	174	172	191	187
37	TCD530747	419	174	172	191	187
38	TCD537014	419	175	177	190	187
39	TCD558851	418	175	172	192	187
40	TCD619382	419	175	172	192	187
41	TCD632685	419	175	172	191	187
42	TCD671445	419	175	171	191	187
43	TCD717480	419	175	172	192	187
44	TCD723719	419	175	172	192	187
45	TCD734935	419	175	172	191	187
46	TCD734886	419	175	172	191	187
47	TCD687093	419	175	171	191	187
48	TCD674832	419	174	171	191	187
49	TCD678877	418	175	171	191	187
50	TMI1647764	418	171	171	190	187

3.3.2 Measure of diversity at each VNTR locus

In this study, the Simpson's DI for the 13 VNTR loci ranged from 0.00 to 0.940 and averaged at 0.529 (Table 3.2). The most variable VNTR loci were TR2, TR4699 and Sal02 with Simpson's DI of 0.940, 0.921 and 0.916 respectively. Five more VNTR loci that were variable included TR1, Sal16, Sal20, TR3 and TR4500 exhibiting Simpson's DI of 0.868, 0.839, 0.730, 0.684 and 0.607 respectively. Even though TR3 showed a higher Simpson's DI of 0.684, this locus was excluded from the study as it contained a long repeat sequence (26 bp repeat length); preference was given to short repeat sequences which show more variability [198]. The remaining five VNTR loci had low diversity indices with the Simpson's DI ranging from 0.00 to 0.339. Of these, three (VNTR locus TR4, Sal10 and Sal 15) were monomorphic and showed no variation amongst the isolates (Table 3.1). The two other VNTR loci (TR5, Sal06) were found unsuitable for the MLVA assay as they showed poor variation and had low diversity indices of 0.039 and 0.339.

Table 3.2 Simpson's diversity indices and primers selected for the amplification of 13 *Salmonella* Typhi VNTR loci

VNTR Locus	Diversity Index	Confidence Interval	Repeat sequence	VNTR Primer sequences	VNTR Primer references
TR2	0.940	0.916 - 0.964	CCAGTTCC	Forward: CCCTGTTTTTCGTGCTGATACG Reverse: CAGAGGATATCGCAACAATCGG	[11]
TR4699	0.921	0.892 - 0.950	TGTTGG	Forward: CGGGCAATTCGAGATAGGTA Reverse: AACCTCCCTGTATCTACCAA	[193]
Sal02	0.916	0.896 - 0.936	TACCAG	Forward: CGATAGACAGCACCAGCAGA Reverse: TCGCCAATACCATGAGTACG	[195]
TR1	0.868	0.836 - 0.900	AGAAGAA	Forward: GCCAACGATCGCTACTTTTT Reverse: CAAGAAGTGCGCATACTACACC	[193, 194]
Sal16	0.839	0.805 - 0.874	ACCATG	Forward: TGCAGTTAATTTCTGCGATCA Reverse: CCTTCCGGATGTATGTGACC	[195]
Sal20	0.730	0.657 - 0.804	CAG	Forward: CAGCCGACACAACTTAACGA Reverse: ACTGTACCGTGCGCGTTT	[193]
TR3	0.684	0.605 - 0.763	CGCGGGGATCGGTTTATCCCCGCTGG	Forward: CGAAGGCGGAAAAAACGTCCTG Reverse: TGCGATTGGTGTCGTTTCTACC	[11]
TR4500	0.607	0.530 - 0.684	GGACTC	Forward: CGTTGCTGCTCCGAAAT Reverse: GCGGTGAAGTGGAAAAAG	[193]
Sal06	0.339	0.185 - 0.494	CTCAAT	Forward: TTGGTCGCGGAACTATAACTG Reverse: CTTCGTCTGATTGCCACTCC	[194]
TR5	0.039	0.000 - 0.114	CGTCACG	Forward: TGAAAACCGGCTCGTAGCAGTG Reverse: CATACGGTTACTGCGGATTGG	[11]
Sal15	0.000	0.000 - 0.132	No data published	Forward: GTGACCGGTTGAGTTTGCAT Reverse: GGCAGGTTGTACCAGTTCGT	[194]
Sal10	0.000	0.000 - 0.132	ACGCCGCTGCCG	Forward: AAGCGACGTTCTTCTGCAAC Reverse: TGGAATATGATGGCATGACG	[194]
TR4	0.000	0.000 - 0.132	GAAATAAAAAATG	Forward: AAAAGCCCGTCTAGTCTTGCAG Reverse: ATCCTTCGGTATCGGGGTATCC	[11]

3.3.3 MLVA validation by nucleotide sequencing

The seven VNTR loci (TR1, TR2, Sal02, Sal16, Sal20, TR4500 and TR4699) that were determined to be highly variable using the Simpson's DI were selected for nucleotide sequencing in order to verify that variation observed in these loci occurred within the tandem repeat regions.

VNTR locus TR1

Variation at VNTR TR1 was determined to be in the tandem repeat regions. The flanking regions for this VNTR locus were consistent for all ten *Salmonella* Typhi isolates that were sequenced and determined to be 39bp (forward flanking) and 116bp (reverse flanking) in size (Appendix D1).

VNTR locus TR2

Nucleotide sequencing showed variation within the tandem repeat region of the VNTR locus TR2. The forward and reverse flanking regions for VNTR locus TR2 were determined to be 191bp and 105bp respectively (Appendix D2).

VNTR locus Sal02

Nucleotide sequencing of the ten *Salmonella* isolates revealed variation in the VNTR locus Sal02 was within the tandem repeat region as the forward and reverse flanking regions for this VNTR locus were consistent and determined to be 136bp and 59bp, respectively (Appendix D3).

VNTR locus TR4699

Nucleotide sequencing of the ten *Salmonella* Typhi isolates determined the forward and reverse flanking regions for VNTR locus TR4699 to be 38bp and 137bp in size. The flanking regions were consistent for all ten isolates and variation in this VNTR locus was observed within the tandem repeat region (Appendix D4).

VNTR locus Sal16

Through nucleotide sequencing, the reverse flanking region for VNTR locus Sal16 was determined to be 49bp in size and was consistent amongst the ten *Salmonella* Typhi isolates that were sequenced. The determination of the forward flanking regions was challenging as variable sizes were observed. Three different sizes for the forward flanking region were observed and determined to be 126bp, 138bp and 162bp. Nucleotide sequencing revealed an additional tandem repeat region upstream of the VNTR locus Sal16 repeat region consisting of a 12bp repeat sequence, ACCACCATTACG (Appendix D5).

VNTR Locus TR4500

Nucleotide sequencing of the VNTR locus TR4500 showed that variation observed in this locus was not within the tandem repeat region but was observed in the flanking regions. A seven base pair 'TTGCCAC' insertion sequence was identified in several *Salmonella* Typhi isolates. Although the size of the forward flanking region was consistent at 76bp, the size of the reverse flanking region varied between two sizes including, 195bp and 202bp (Appendix D6).

VNTR locus Sal20

Nucleotide sequencing of VNTR locus Sal20 showed that variation in this locus occurs within the tandem repeat region as the sizes of the forward and reverse flanking regions were consistent and determined to be 83bp and 80bp respectively (Appendix D7)

3.3.4 Optimisation of the multiplex MLVA assay consisting of five VNTR loci

The development of the final MLVA multiplex PCR assay involved pooling the PCR amplicons of five highly variable VNTR loci into a single PCR reaction. The use of capillary electrophoresis for analysis of PCR amplicon sizes required labelling of forward VNTR primers with fluorescent dyes in order to enable differentiation of each of the VNTR loci PCR amplicons. Four fluorescent dyes were available for selection and these included VIC, FAM, NED and PET. VNTR loci with overlapping PCR amplicon sizes were assigned different fluorescent dyes. The selection of the fluorescent dyes was based on assessment of VNTR PCR amplicon sizes of the 50 *Salmonella* Typhi isolates from the simplex PCR assays in order to determine the size range of each VNTR locus (Figure 3.3).

The development of the multiplex PCR assay began by incorporating equal amounts of VNTR primer concentrations in the PCR reaction. This resulted in various VNTR amplicons exhibiting stronger signals than others. According to the GeneScan guide, fluorescent dyes have variable signal strength, with VIC and FAM having stronger signals than NED and PET [199]. Consequently, the primer concentrations of the various VNTR loci were adjusted to show equal signal strength on the electropherogram (Figure 3.4).

Table 3.3 Features of seven highly polymorphic *Salmonella* Typhi VNTR loci

VNTR Locus	Gene	Product	Repeat sequence	Unit length	Forward offset (bp)	Reverse offset (bp)	MLVA multiplex PCR primer concentrations	Reference(s)
TR1	-	Intergenic region between yedD and yedE	AGAAGAA	7	39	116	1.2 µM	[11, 193]
TR2	-	Intergenic region between arcD and yffB	CCAGTTCC	8	191	105	1.2 µM	[11, 193]
TR4699	sefC	Outer membrane fimbral usher protein	TGTTGG	6	38	137	0.8 µM	[193]
Sal02	citT	Citrate carrier	TACCAG	6	136	59	1 µM	[193, 194]
Sal16	-	Intergenic region between STY3169 (pseudogene) and STY3172	ACCATG	6	90	91, 97, 103, 109, 115, 127	-	[193]
Sal20	ftsN	Cell division protein	CAG	3	83	80	0.5µM	[193, 194]
TR4500	STY4635	Hypothetical protein	GGACTC	6	76	195, 202	-	[193]

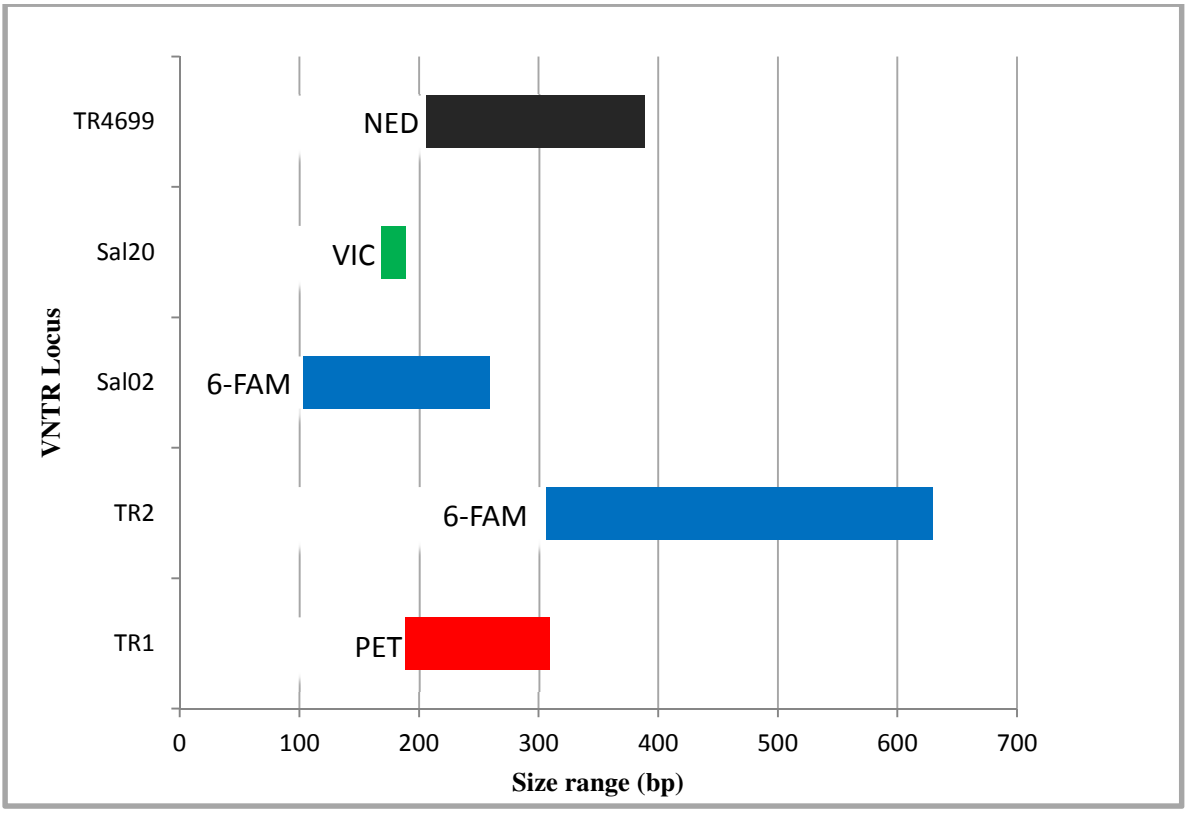


Figure 3.3 Bar diagram showing selection of fluorescent dyes based on amplicon size ranges for the five VNTR loci

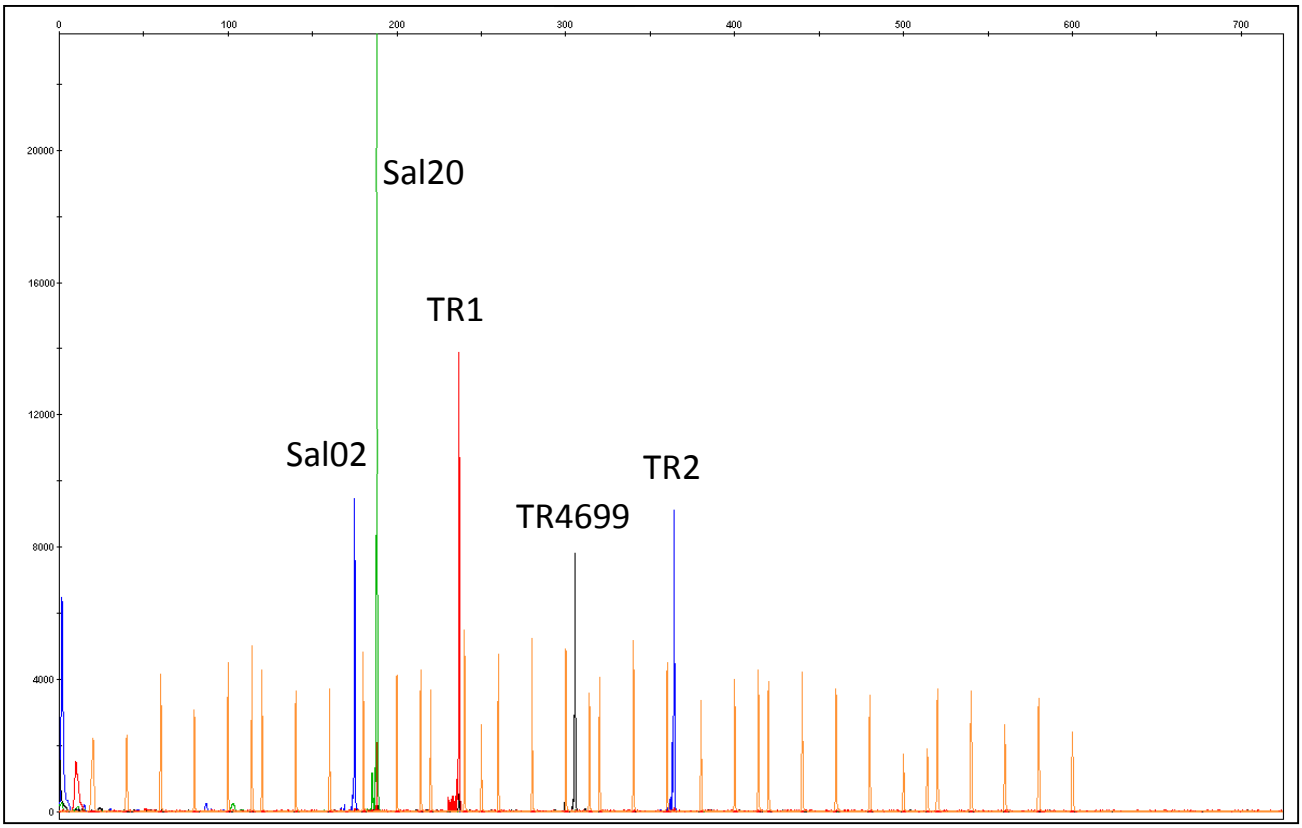
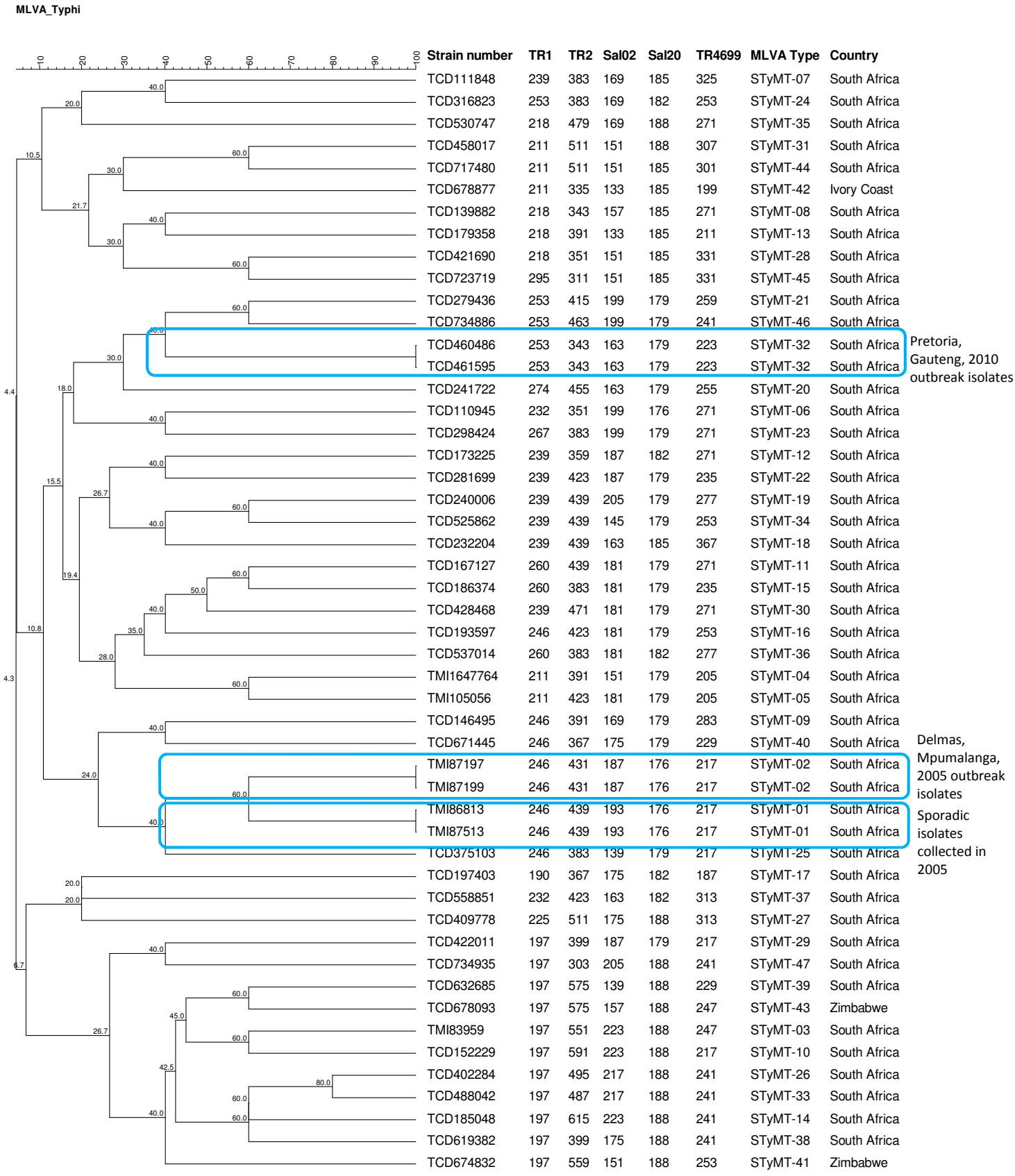


Figure 3.4 Electropherogram showing all five VNTR loci incorporated in a multiplex PCR for MLVA assay of *Salmonella* Typhi isolates

3.3.5 Subtyping of *Salmonella* Typhi isolates using the developed MLVA multiplex PCR assay

The MLVA multiplex PCR assay developed in this study was used to type 50 *Salmonella* Typhi validation isolates. Length bins were allocated DNA fragments and bin numbers were assigned based on the bin fragment sizes. The fragment sizes for each of the five VNTR loci were combined in a string consisting of integers to form a MLVA profile and each MLVA profile was assigned a MLVA type (STyMT). As an example, for isolate TCD111848, the fragment sizes for VNTR locus TR1, TR2, Sal02, Sal20 and TR4699 were determined to be 239bp, 383bp, 169bp, 185bp and 325bp respectively. Therefore, the MLVA profile for isolate TCD111848 is 239-383-168-1845-253 and this profile was designated MLVA type STyMT-206.

The MLVA assay was able to segregate the isolates into 47 MLVA types (Figure 3.5). Three clusters (consisting of isolates with indistinguishable MLVA profiles) were identified from the dendrogram. One of the clusters consisted of the two representative isolates from the Gauteng, 2010 outbreak [89]. Another cluster consisted of two representative isolates collected during typhoid fever outbreaks in Mpumalanga, 2005. The third cluster consisted of sporadic isolates which were collected in the year of the Mpumalanga 2005 outbreak. These isolates were unrelated to the two representative isolates from the Mpumalanga outbreak. MLVA profiles of the two representative isolates from Zimbabwe were unrelated to each other and unrelated to profiles from South African isolates. MLVA profiles for isolates from Ivory Coast were also determined and found to be unrelated to profiles from South Africa and Zimbabwe.



Strain number	TR1	TR2	Sal02	Sal20	TR4699	MLVA Type	Country	
TCD111848	239	383	169	185	325	STyMT-07	South Africa	
TCD316823	253	383	169	182	253	STyMT-24	South Africa	
TCD530747	218	479	169	188	271	STyMT-35	South Africa	
TCD458017	211	511	151	188	307	STyMT-31	South Africa	
TCD717480	211	511	151	185	301	STyMT-44	South Africa	
TCD678877	211	335	133	185	199	STyMT-42	Ivory Coast	
TCD139882	218	343	157	185	271	STyMT-08	South Africa	
TCD179358	218	391	133	185	211	STyMT-13	South Africa	
TCD421690	218	351	151	185	331	STyMT-28	South Africa	
TCD723719	295	311	151	185	331	STyMT-45	South Africa	
TCD279436	253	415	199	179	259	STyMT-21	South Africa	
TCD734886	253	463	199	179	241	STyMT-46	South Africa	
TCD460486	253	343	163	179	223	STyMT-32	South Africa	Pretoria, Gauteng, 2010 outbreak isolates
TCD461595	253	343	163	179	223	STyMT-32	South Africa	
TCD241722	274	455	163	179	255	STyMT-20	South Africa	
TCD110945	232	351	199	176	271	STyMT-06	South Africa	
TCD298424	267	383	199	179	271	STyMT-23	South Africa	
TCD173225	239	359	187	182	271	STyMT-12	South Africa	
TCD281699	239	423	187	179	235	STyMT-22	South Africa	
TCD240006	239	439	205	179	277	STyMT-19	South Africa	
TCD525862	239	439	145	179	253	STyMT-34	South Africa	
TCD232204	239	439	163	185	367	STyMT-18	South Africa	
TCD167127	260	439	181	179	271	STyMT-11	South Africa	
TCD186374	260	383	181	179	235	STyMT-15	South Africa	
TCD428468	239	471	181	179	271	STyMT-30	South Africa	
TCD193597	246	423	181	179	253	STyMT-16	South Africa	
TCD537014	260	383	181	182	277	STyMT-36	South Africa	
TMI1647764	211	391	151	179	205	STyMT-04	South Africa	
TMI105056	211	423	181	179	205	STyMT-05	South Africa	
TCD146495	246	391	169	179	283	STyMT-09	South Africa	
TCD671445	246	367	175	179	229	STyMT-40	South Africa	
TMI87197	246	431	187	176	217	STyMT-02	South Africa	Delmas, Mpumalanga, 2005 outbreak isolates
TMI87199	246	431	187	176	217	STyMT-02	South Africa	
TMI86813	246	439	193	176	217	STyMT-01	South Africa	Sporadic isolates collected in 2005
TMI87513	246	439	193	176	217	STyMT-01	South Africa	
TCD375103	246	383	139	179	217	STyMT-25	South Africa	
TCD197403	190	367	175	182	187	STyMT-17	South Africa	
TCD558851	232	423	163	182	313	STyMT-37	South Africa	
TCD409778	225	511	175	188	313	STyMT-27	South Africa	
TCD422011	197	399	187	179	217	STyMT-29	South Africa	
TCD734935	197	303	205	188	241	STyMT-47	South Africa	
TCD632685	197	575	139	188	229	STyMT-39	South Africa	
TCD678093	197	575	157	188	247	STyMT-43	Zimbabwe	
TMI83959	197	551	223	188	247	STyMT-03	South Africa	
TCD152229	197	591	223	188	217	STyMT-10	South Africa	
TCD402284	197	495	217	188	241	STyMT-26	South Africa	
TCD488042	197	487	217	188	241	STyMT-33	South Africa	
TCD185048	197	615	223	188	241	STyMT-14	South Africa	
TCD619382	197	399	175	188	241	STyMT-38	South Africa	
TCD674832	197	559	151	188	253	STyMT-41	Zimbabwe	

Figure 3.5 MLVA cluster analysis of *Salmonella* Typhi isolates from sub-Saharan Africa

3.3.6 Discriminatory power of MLVA assay

The established MLVA assay targeted five VNTR loci including TR1, TR2, Sal02, Sal20 and TR4699. The discriminatory power of the MLVA assay was calculated using the Simpson's DI applied to the *Salmonella* Typhi isolate panel. Of the 50 *Salmonella* Typhi isolates that were analysed, 47 MLVA types were identified. For the MLVA assay, Simpson's DI was calculated at 0.998 (95% confidence interval (CI) 0.995 – 1.000) [Table 3.3]. This was a high value compared to the Simpson's DI for PFGE analysis of the same strains which was calculated at 0.984 (95% CI 0.974 – 0.994). With non-overlapping CIs, the difference in Simpson's DI for MLVA and PFGE were statistically significant (P = 0.013).

The Wallace coefficient was calculated in order to assess the congruence between the MLVA assay and PFGE analysis. This coefficient indicates the probability that 2 strains that cluster together by one subtyping method could also be clustered together using another subtyping method [200]. The Wallace coefficient between MLVA assay and PFGE pulsotypes was 67% (Table 3.4). In contrast, the Wallace coefficient showed that PFGE could only poorly (9%) predict the results of the MLVA typing method.

Table 3.3 Simpson's Index of diversity for MLVA assay and PFGE

	Number of patterns	Diversity Index	Confidence Interval
PFGE Pattern	34	0.982	(0.970-0.994)
MLVA assay	47	0.998	(0.994-1.000)

Table 3.4 Wallace coefficient and analytical 95% CI of MLVA assay versus PFGE

	PFGE Pattern	MLVA assay
PFGE Pattern		***0.091*** (0.000-0.278)
MLVA assay	0.667 (0.333-1.000)	

3.4 Discussion

Epidemiological investigations are important for the control and prevention of typhoid fever outbreaks. Similarly, molecular subtyping of isolates serves as an important epidemiological investigation tool used to improve surveillance and epidemiological investigations. The MLVA typing scheme has been successfully used in outbreak investigation of many pathogenic bacteria. This is a fast and straightforward method with less complicated laboratory protocols compared to the widely used PFGE analysis method. MLVA uses PCR assisted amplification of VNTR loci and size determination of the PCR amplicons by capillary electrophoresis. Fragment analysis by capillary electrophoresis is performed for accurate determination of PCR amplicons sizes in base pair.

In this study, we present a MLVA assay based on five highly polymorphic VNTR loci for studying the genetic diversity of *Salmonella* Typhi isolates from SSA. VNTR locus TR1 located on the intergenic region between *yedD* and *yedE* carried a seven base pair repeat motif, AGAAGAA [193]. This VNTR locus was first characterized by Liu and colleagues [11]. In their study, Liu and colleagues found VNTR locus TR1 to be highly variable in typing *Salmonella* Typhi isolates from Asian countries. VNTR locus TR1 was also explored in other studies and was found to be highly variable, exhibiting Simpson's DI ranging between 0.87 to 0.90 [193-195]. In our study, VNTR locus TR1 exhibited a Simpson's DI of 0.87 (Table 3.2). Due to high variations observed, VNTR locus TR1 was selected for inclusion in our MLVA assay.

VNTR locus TR2 was also first characterized by Liu and colleagues (2013) [11]. This VNTR locus is located in the intergenic region between *arcD* and *yffB* genes and carries an eight base pair sequence motif, CCAGTTCC. Several studies including the study by Liu and colleagues (2013) described VNTR locus TR2 as a highly variable VNTR locus [11, 193-195] with an average Simpson's DI of 0.95. In our study, VNTR locus TR2 exhibited Simpson's DI of 0.94. VNTR locus TR2 was the most variable VNTR locus compared to other VNTR loci evaluated in our study and was therefore selected for inclusion in our MLVA assay.

VNTR locus Sal02 was first described in 2004 by Ramisse and colleagues [194]. This VNTR locus is located on the *citT* gene encoding the Citrate carrier [193] and carries a six bp repeat sequence, TACCAG. VNTR locus Sal02 was previously identified as highly polymorphic VNTR locus, with Simpson's DI ranging from 0.87 to 0.92 [11, 193-195]. Similarly, in our study, VNTR locus Sal02 showed a high Simpson's DI of 0.92 (Table 3.2). Therefore, this VNTR locus was also selected for inclusion in our MLVA assay.

VNTR locus TR4699 was first described by Octavia and Lan in 2009 [193]. This VNTR locus is located on a *sefC* gene encoding the outer membrane fimbral usher protein and carries a TGTTGG repeat [193]. Octavia and Lan (2009) found this VNTR locus to be highly variable with a Simpson's DI of 0.95. VNTR locus TR4699 was also explored in a study by Tien and colleagues (2012) and was found to be highly variable with a Simpson's DI of 0.92 [195]. In our study, VNTR locus TR4699 also showed a high Simpson's DI of 0.92 and was selected for inclusion in our MLVA assay.

VNTR locus Sal16 is located in the intergenic region between the pseudogene STY3169 and STY3172 [193]. This VNTR locus has a six bp repeat motif, ACCATG. Octavia and Lan (2009) found VNTR locus Sal16 to be highly variable with a Simpson's DI of 0.83. In a previous report, VNTR locus Sal16 was found to be inconsistent for *Salmonella enterica* subspecies *enterica* serovar Typhimurium strain LT2 [194]. Reports indicate that the inconsistencies in this VNTR locus are due to genome sequencing errors [193, 194]. In our study, VNTR locus Sal16 exhibited a high Simpson's DI of 0.84. Even though high Simpson's DI was observed for this VNTR locus, nucleotide sequencing revealed an additional tandem repeat region upstream of the VNTR locus Sal16 tandem repeat region. This region carried a 12bp repeat motif, ACCACCATCACG. Due to this region being a tandem repeat region, it is predisposed to high rates of polymorphism which will result in alteration of the product size of the VNTR locus Sal16. As a result, the correct number of repeat units for this VNTR locus cannot be determined. Designing new forward primers for VNTR locus Sal16 would not have eliminated the problem as the two tandem repeat regions are separated only by 12bp sequence, ACCACCATTACG. Consequently, VNTR locus Sal16 was not selected for inclusion in our MLVA assay.

VNTR locus TR4500 was first described in a study by Octavia and Lan [193] with an aim to develop a MLVA assay for the analysis of *Salmonella* Typhi strains from Australia. This VNTR locus is located on gene STY4635 of the *Salmonella* Typhi bacterium, and is a hypothetical protein with a consensus sequence consisting of six nucleotides, GGACTC [193]. Octavia and Lan (2009) found VNTR locus TR4500 to be highly variable with Simpson's DI of 0.691. In our study, VNTR locus TR4500 exhibited a Simpson's DI of 0.607, slightly lower than the Simpson's DI reported by Octavia and colleagues. A seven bp 'TTGCCAC' insertion sequence was identified in several *Salmonella* Typhi isolates. Due to this insertion, the correct number of repeat units for each isolate could not be determined as not all isolates harboured the seven bp sequence. The designing of new primers could not be considered as an option due to the fact that the insertion sequence is adjacent to the tandem repeat region. It is for this reason that the VNTR locus TR4500 was excluded from our MLVA assay.

VNTR locus Sal20 was first described by Ramisse and colleagues (2004) in France [194]. Their study was aimed at developing a MLVA assay for analysis of *Salmonella enterica* subspecies *enterica* isolates. They identified VNTR locus Sal20 as one of the most variable VNTR loci exhibiting a Simpson's DI of 0.81 [194]. VNTR locus Sal20 carries a three bp sequence, CAG, and is located on the *ftsN* gene coding for the cell division protein in *Salmonella* Typhi isolates [193]. This VNTR locus was found to be more variable in *Salmonella* Typhi than in other *Salmonella* species [194]. In our study, VNTR locus Sal20 showed a Simpson's DI of 0.730 and showed variation amongst the *Salmonella* Typhi isolates. This VNTR locus was included in our MLVA assay.

The diversity at each VNTR locus is a function of both the number of alleles as well as their distribution frequency within a population and, high diversity indices indicate more variable VNTR loci. The combination of the five highly variable VNTR loci in the MLVA assay resulted in a highly discriminatory molecular typing scheme. We assessed the discriminatory power of the MLVA assay developed for *Salmonella* Typhi using a panel of 50 isolates from SSA previously characterized by PFGE. The panel used in this study consisted of diverse collection of isolates that were epidemiologically unrelated as well as

closely related isolates. PFGE analysis of these isolates showed 34 unique pulsotypes. However, high discriminatory levels were achieved with the MLVA assay presented in this study. The MLVA assay consisting of the selected five VNTR loci (TR1, TR2, Sal02, Sa20 and TR4699) was able to differentiate the 50 *Salmonella* Typhi isolates into 47 MLVA types. The Simpson's DI for the MLVA assay was 0.998 indicating high discrimination abilities. The combination of VNTR loci with high diversity levels enabled differentiation of closely related and unrelated isolates as clusters of isolates from recent outbreaks in South Africa namely; Delmas, Mpumalanga outbreak in 2005 and the Pretoria, Gauteng outbreak in 2010. The MLVA assay was also able to distinguish between isolates from Ivory Coast, South Africa and Zimbabwe.

To evaluate congruence between MLVA and PFGE, the Wallace coefficient was calculated. The Wallace index between MLVA and PFGE was 67%, indicating good correlation between MLVA and PFGE. However, there was lower probability that isolates assigned to the same cluster by PFGE would be assigned to the same cluster by MLVA. MLVA assay showed higher discrimination of *Salmonella* Typhi isolates than PFGE. Although capillary electrophoresis is required in order to correctly determine PCR amplicon sizes, the speed and simplicity in processing as well as data interpretation makes MLVA a more suitable molecular subtyping technique.

The limitation in this study was the restricted number of isolates used to validate the MLVA assay. Testing a large number of isolates could help establish the discriminatory capacity of the MLVA assay in typing isolates from local and intercontinental outbreaks. Additionally, *Salmonella* Typhi isolates from other African countries were represented in low numbers. Numerous requests were put out to other African countries for *Salmonella* Typhi isolates; however, this was met with no response or false promises to send. The stability of the VNTR loci was not established. Further studies are required to determine whether VNTR loci remain stable in *Salmonella* Typhi isolates that have gone through a series of laboratory processing.

3.5 Conclusion

Our study describes a MLVA assay consisting of five VNTR markers for the rapid analysis of *Salmonella* Typhi isolates from SSA which can assist epidemiological investigations of strain relatedness and detection of typhoid fever outbreaks in SSA. The assay represents a high-throughput typing method that is a rapid and highly discriminatory molecular tool that may be a viable alternative to PFGE for subtyping *Salmonella* Typhi isolates from SSA. We have shown that the MLVA assay developed in this study is suitable to characterize *Salmonella* Typhi from the SSA as isolates from Ivory Coast, Zimbabwe and South Africa were discriminated using this MLVA assay. The assay was also able to differentiate between outbreak strains and epidemiologically unrelated strains as well as identify isolates that were closely related to the outbreak strains.

This MLVA assay was validated using isolates from SSA as we only had African strains to work with, and specifically wanted to ensure that it would be applicable to African laboratories. Even so, the implementation of this assay in laboratories across the world will enable inter-laboratory comparison of *Salmonella* Typhi strains.

Chapter 4

The prevalence of H58 *Salmonella* Typhi in sub-Saharan Africa

4.1 Introduction

In sub-Saharan Africa (SSA), the spread of *Salmonella* Typhi is exacerbated by the lack of safe drinking water and good hygiene practices [28, 201]. Even more concerning is the spread of the highly clonal H58 *Salmonella* Typhi (H58 haplotype) which has been documented in some parts of Africa [29, 132]. These report the spread and transmission of H58 *Salmonella* Typhi from Asia to sub-Saharan Africa. Kariuki and colleagues (2010) studied the association of typhoid fever with a dominant MDR *Salmonella* Typhi haplotype [29]. In their study, WGS and bioinformatics approaches were used to discriminate *Salmonella* Typhi isolates into distinct phylogenetic lineages. Approximately 1500 single nucleotide variations on the *Salmonella* Typhi genome were interrogated using a novel SNP typing method. The H58 *Salmonella* Typhi was found to be the predominant strain in Kenya accounting for approximately 76% of all *Salmonella* Typhi isolates. In their study, Kariuki and colleagues (2010) also demonstrated the H58 haplotype, previously shown to be widespread in Asia, was predominant in Kenya, and was replacing the antimicrobial-susceptible strains [29].

Wong and colleagues (2015) conducted a phylogeographical analysis of the H58 *Salmonella* Typhi [132]. In their study, whole-genome sequences of approximately 1,850 global *Salmonella* Typhi isolates were analysed. A simple SNP-based typing scheme stratifying the *Salmonella* Typhi population into haplotypes was used to map the isolates into phylogeny. Through this typing scheme, they determined that 47% of the global *Salmonella* Typhi population belonged to the H58 haplotype. They also showed that there was inter- and intracontinental transmission events of the H58 haplotype as a number of very closely related isolates from different countries were identified. Their data suggested that South East Asia was the early hub for the H58 haplotype from which it was propagated to many locations around the world including East Africa. Data also indicated that the H58 haplotype was predominant amongst the eastern and southern African typhoid fever isolates. The detection of the H58 haplotype in Kenya, Tanzania, Malawi and South Africa provided compelling

evidence for multiple introductions of H58 *Salmonella* Typhi from South East Asia into the African continent. In their study, Wong and colleagues (2015) revealed evidence of an unreported wave of transmission of the H58 haplotype from Kenya to Tanzania and on to Malawi and South Africa (Figure 4.1). This demonstrated an on-going epidemic of H58 typhoid fever across the SSA countries [132].

The surveillance of the H58 *Salmonella* Typhi in SSA is key in controlling the dissemination of this pathogen in the continent. The most reliable genotypic method that has been used to identify the H58 *Salmonella* Typhi isolates is WGS [29, 132]. However, for routine use in most laboratories in the SSA region, implementation of this methodology would be difficult due to lack of required equipment, lack of expertise and lack of funds. In a study by Holt and colleagues (2008), WGS analysis of approximately 20 *Salmonella* Typhi isolates identified 20 genomic deletions suitable for genotyping *Salmonella* Typhi isolates [128, 202]. Amongst the H58 haplotype, a 993bp deletion affecting the STY 1507 and STY 1508 genes was identified. A polymerase chain reaction (PCR) assay targeting this deletion was then described and has been found suitable for the identification of isolates belonging to H58 haplotype [202].

4.1.1 Aim

In our current study, we used the recently developed conventional PCR-based typing assay [202] for the rapid and easy detection of H58 haplotype amongst *Salmonella* Typhi isolates from SSA.

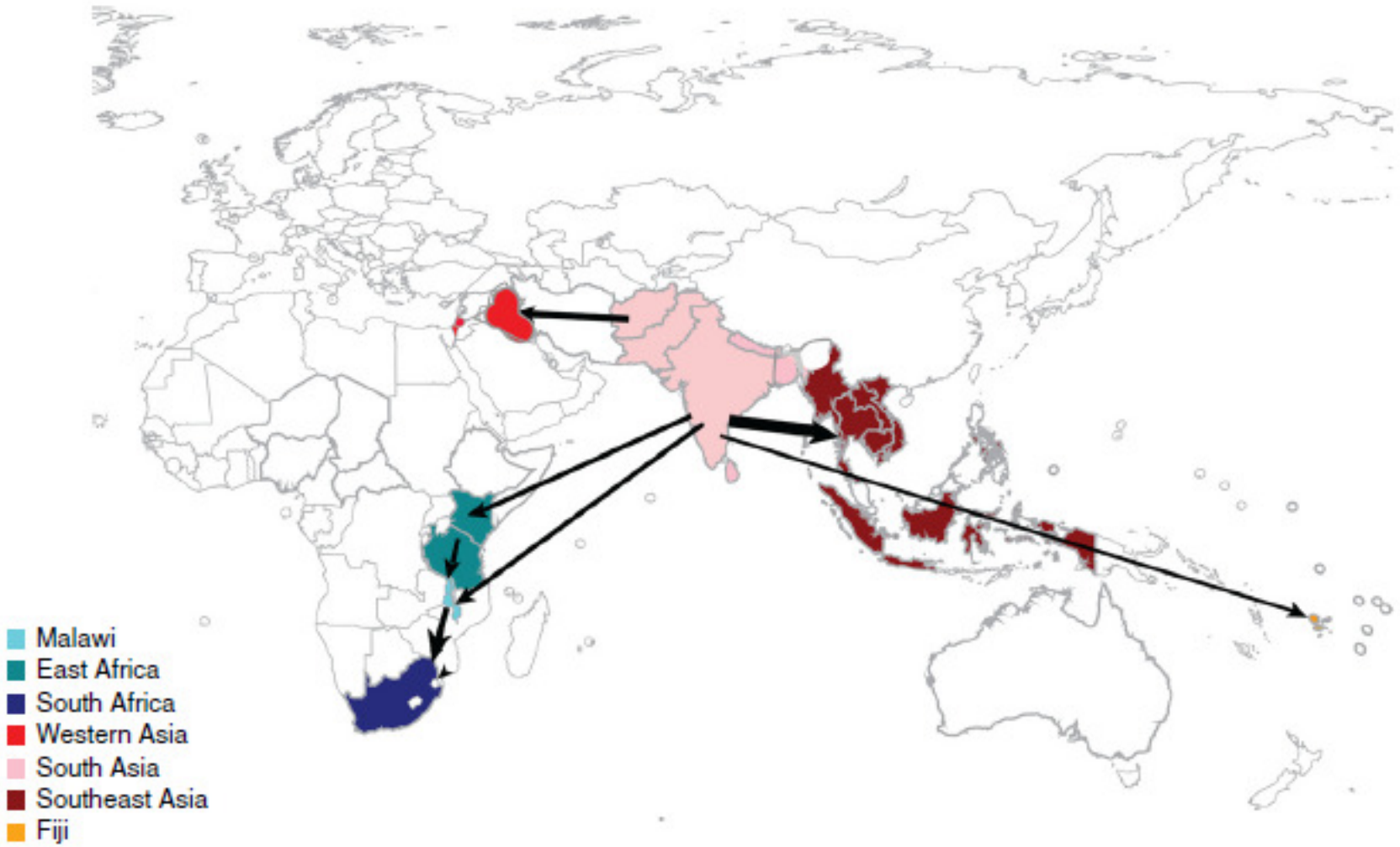


Figure 4.1 Inter and intracontinental transfer of H58 *Salmonella* Typhi [132]

4.2 Materials and Methods

4.2.1 Crude DNA extraction

Genomic DNA extractions were performed as described in Chapter 2 (section 2.3).

4.2.2 Serotype confirmation of *Salmonella* Typhi isolates

4.2.2.1 Serotyping of *Salmonella* Typhi isolates by slide agglutination

All *Salmonella* isolates collected through the GERMS-SA programme were subjected to traditional slide agglutination serotyping as part of routine surveillance (See Chapter 1 section 2.3). Specific antisera (Statens Serum Institut, Copenhagen, Denmark and BioMérieux, Marcy-I'Étoile, France) were used to serotype *Salmonella* Typhi isolates using the Kauffmann-Le Minor scheme (Appendix E).

4.2.2.2 Real time PCR serotyping of *Salmonella* Typhi isolates

Real-time PCR was employed in order to confirm the serotypes of *Salmonella* Typhi isolates collected through the TSAP study. The real-time PCR was performed as described in Chapter 2 (section 2.4).

4.2.3 Conventional PCR identification of H58 *Salmonella* Typhi

Salmonella Typhi isolates from the GERMS-SA and TSAP programmes were subjected to conventional PCR in order to determine the prevalence of H58 *Salmonella Typhi*. PCR identification of the H58 haplotype was carried out using a conventional PCR method adapted from that described previously [202]. The PCR primers described by Murgia and colleagues (2016) were designed to produce an 107bp amplicon to indicate presence of the deletion in H58 *Salmonella* Typhi isolates and an 1100bp amplicon to indicate absence of the deletion in non-H58 *Salmonella* Typhi isolates. Primer sequences are shown in Table 4.1.

PCR reactions were performed using Amplitaq® Gold DNA polymerase, GeneAmp® 10X PCR buffer and Magnesium chloride ($MgCL_2$) (Applied Biosystems, Foster City, California) incorporating 1.5 units of enzyme per reaction. Reactions were performed in a final volume of 25 µl, consisting of $MgCL_2$ and deoxynucleotide triphosphate (dNTP) (Celtic molecular

diagnostics) at concentrations of 2 mM and 200 µM respectively; with primer concentrations at 1.2 µM each. A no template control (NTC) was included in the reactions to rule out possible contamination. *Salmonella* Typhi NCTC 8385 was used as a positive control for non-H58 *Salmonella* Typhi and a *Salmonella* Typhi clinical strain number TCD0186374 was used as a positive control for H58 *Salmonella* Typhi. All reactions were carried out in 0.2 ml thermo-tubes (Thermo Scientific) and were performed in a Bio-Rad (iCycler) thermal cycler (Bio-Rad). The thermal cycling conditions were as follows: Initial denaturation at 94 °C for 2 minutes, followed by 30 cycles at 94 °C for 30 seconds, 60 °C for 30 seconds, 70 °C for 1 minute, and a final extension at 72 °C for 5 minutes.

Table 4.1 PCR primers selected for specific amplification of H58 *Salmonella* Typhi

Primer	**Primer sequence**	**Position on chromosome of *Salmonella* Typhi CT18**
H58-Forward	GCAGGCAAAATCGAAATCAG	1466515 – 1466534
H58-Reverse	CAAACCGTTGAATCGGAAGT	1467614 – 1457595

4.2.4 Gel electrophoresis

A 3 µl volume of loading dye (2.5% Bromophenol blue solution) (Sigma-Aldrich, St Louis, MO USA) was added to the 25 µl of PCR product (Appendix F). A 6 µl volume of the mixed solution was loaded on a 1.5% SeaKem LE agarose gel (Whitehead Scientific, Rockland, USA) containing 10 mg/ml ethidium bromide (Sigma-Aldrich) (Appendix A).

Gel electrophoresis was performed in 1X TAE buffer Tris (Merck): Acetate (Merck): EDTA (Merck) at 140 V (voltage) for 50 min (Appendix F). The gel image was visualized using the Quantity 1-D analysis software (Bio-Rad Laboratories, Hercules, CA).

4.3 Results

4.3.1 Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa (GERMS-SA study)

4.3.1.1 Annual typhoid fever isolates

A total of 214 *Salmonella* Typhi isolates were collected from patients hospitalized with typhoid fever across South Africa from January 2012 to December 2014. An increase was observed in the annual number of isolates collected with most of the isolates (45%) collected during the year 2014 (Figure 4.2).

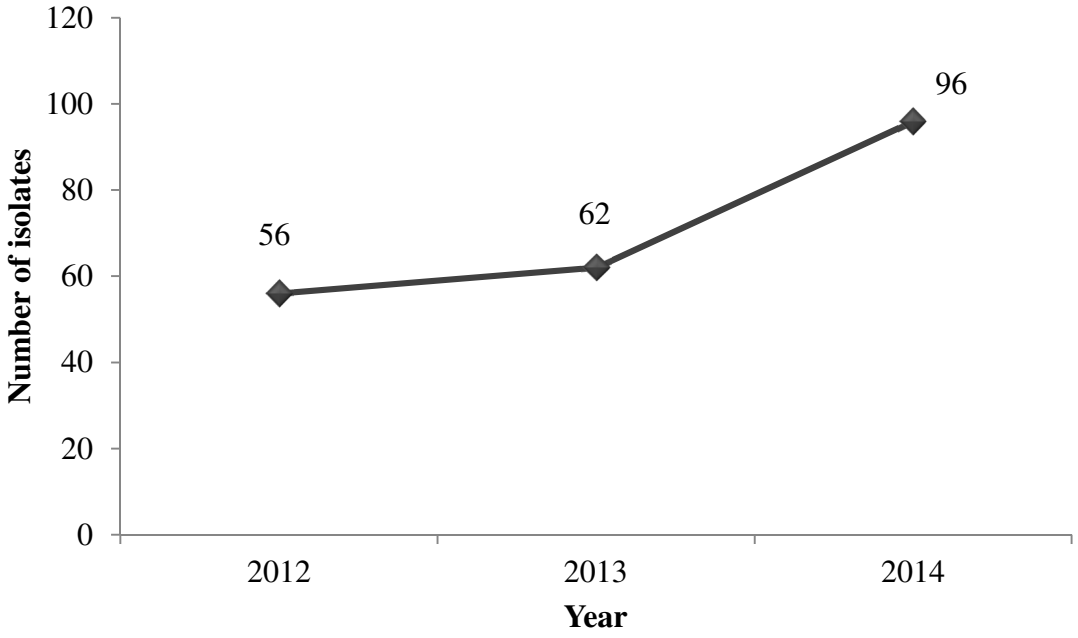


Figure 4.2. Temporal distribution of 214 *Salmonella* Typhi isolates from the GERMS-SA study for the period 2012-2014.

4.3.1.2 PCR screening of H58 *Salmonella* Typhi

Of the 214 isolates that were collected during the study period, 195 isolates were viable; these isolates were subjected to the H58 *Salmonella* Typhi conventional PCR. Of the 195 isolates that were processed, 107bp amplicons were identified in 54% (n=105) of the isolates. This confirmed the presence of a 993bp deletion which is only present in all H58 *Salmonella*

Typhi isolates. For the remaining isolates (n=90), 1100bp amplicons were identified; these isolates belonged to non-H58 *Salmonella* Typhi haplotype groups.

There was an increase in the number of H58 *Salmonella* Typhi isolates identified from the year 2012 to 2014. In the year 2012, 50% (n/N=28/56) of the isolates were identified as H58 *Salmonella* Typhi. This number slightly decreased to 43% (n/N=25/58) in 2013; however, during the year 2014, 64% (n/N=52/81) of the isolates were identified as H58 *Salmonella* Typhi (Figure 4.3). The identification of H58 *Salmonella* Typhi in more than half the number of the isolates (54%) collected between 2012 and 2014 is concerning. The association of H58 *Salmonella* Typhi with MDR as well as reduced susceptibility to fluoroquinolones means that current treatment regimes, such as ciprofloxacin, cannot be used for routine typhoid fever treatment. This then means that more expensive antimicrobial drugs would be required as alternatives.

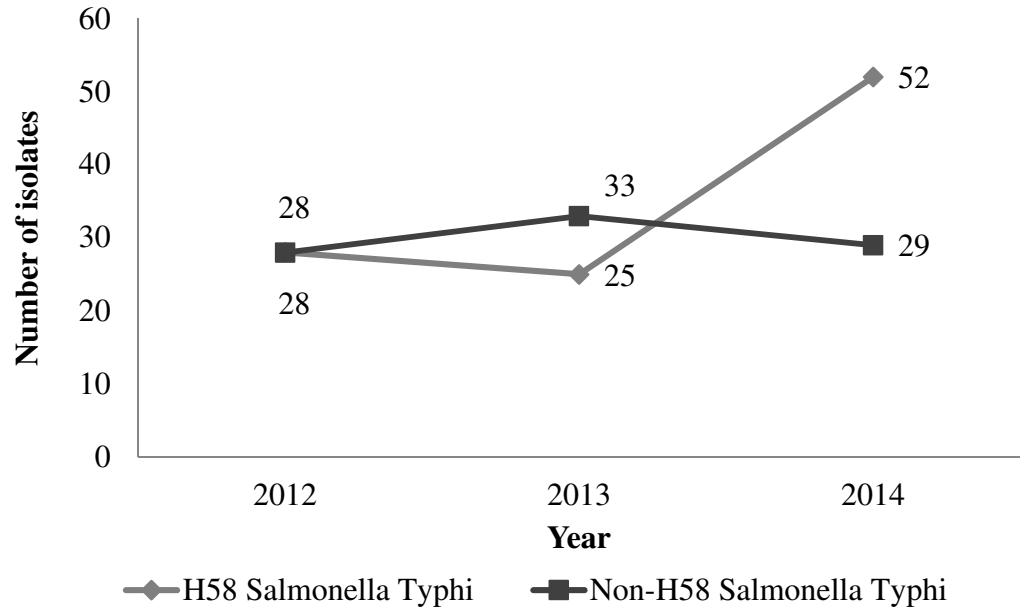


Figure 4.3 Number of H58 *Salmonella* Typhi isolates identified in South Africa from 2012 - 2014.

4.3.1.3 H58 *Salmonella* Typhi by geographical area

Typhoid fever cases were reported in all but one province (n=8) in South Africa (Figure 4.4). Most typhoid fever cases were reported from Gauteng, followed by Western Cape, Kwa-Zulu Natal and Mpumalanga (Figure 4.5).

The H58 *Salmonella* Typhi was identified in all eight provinces (Figure 4.4). This strain was predominant in Gauteng and Western Cape provinces accounting for 57% (n/N=43/75) and 61% (n/N=27/44), respectively, of the *Salmonella* Typhi isolates collected from these provinces. Even though the H58 *Salmonella* Typhi was identified in the Kwa-Zulu Natal province, findings indicated that non-H58 haplotypes were the predominant cause of typhoid fever, accounting for 62% of isolates collected in this province.

4.3.1.4 Prevalence of H58 *Salmonella* Typhi in age groups

The H58 *Salmonella* Typhi was identified in all age groups and caused a slightly higher number of illnesses than other *Salmonella* Typhi haplotypes except in age groups 6–10 and 11–15 and 55+ years (Figure 4.6). Even though the H58 *Salmonella* Typhi was not the principal cause of infection in the 6-10 years age group, approximately 36% of the isolates in this age group belonged to the highly resistant haplotype. Additionally, 83% (n=5) of typhoid fever illness in in children aged 0-5 years was caused by the resistant H58 haplotype.

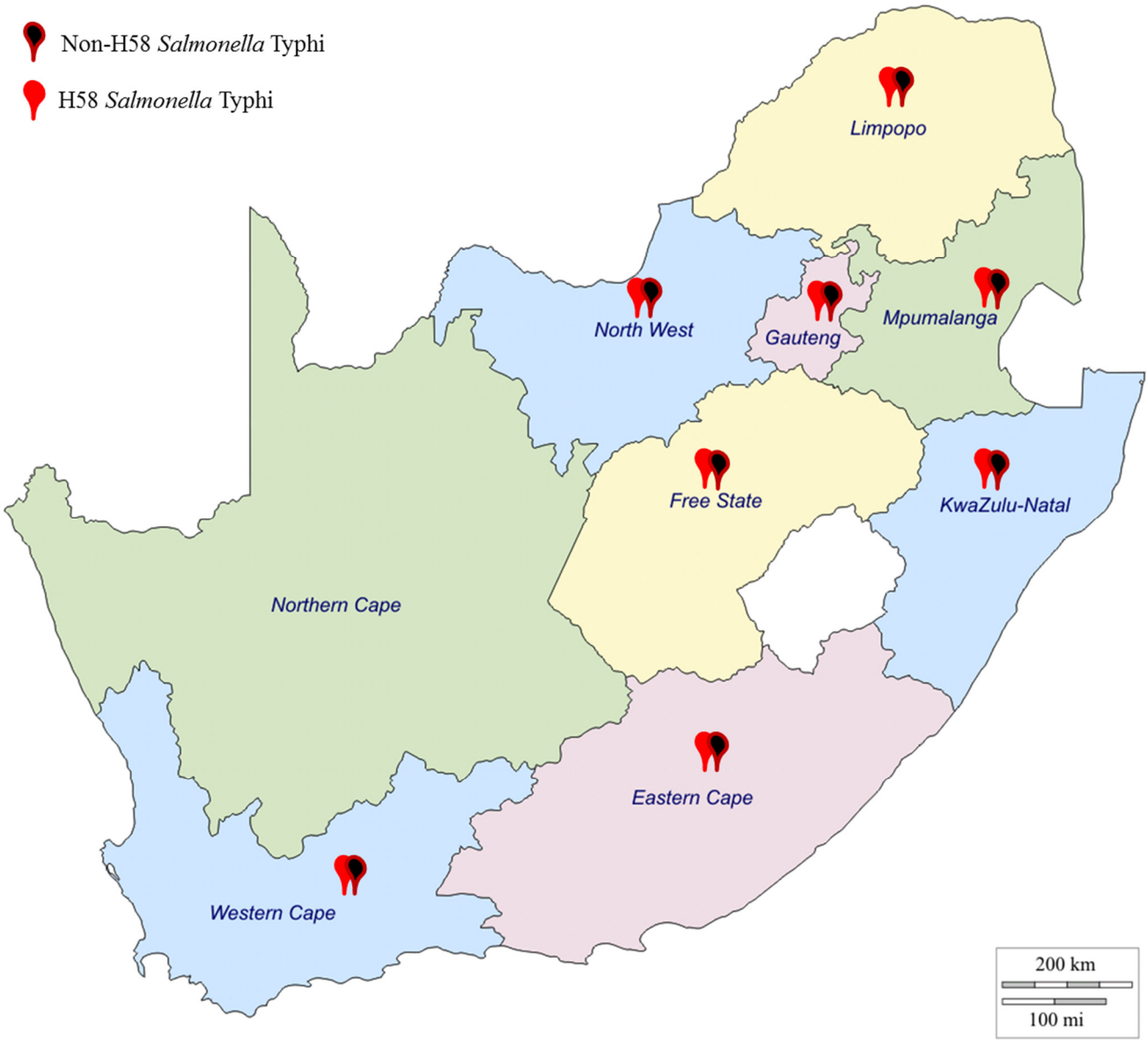


Figure 4.4 Geographical distribution of H58 and non-H58 *Salmonella* Typhi within provinces in South Africa for the years 2012 to 2014

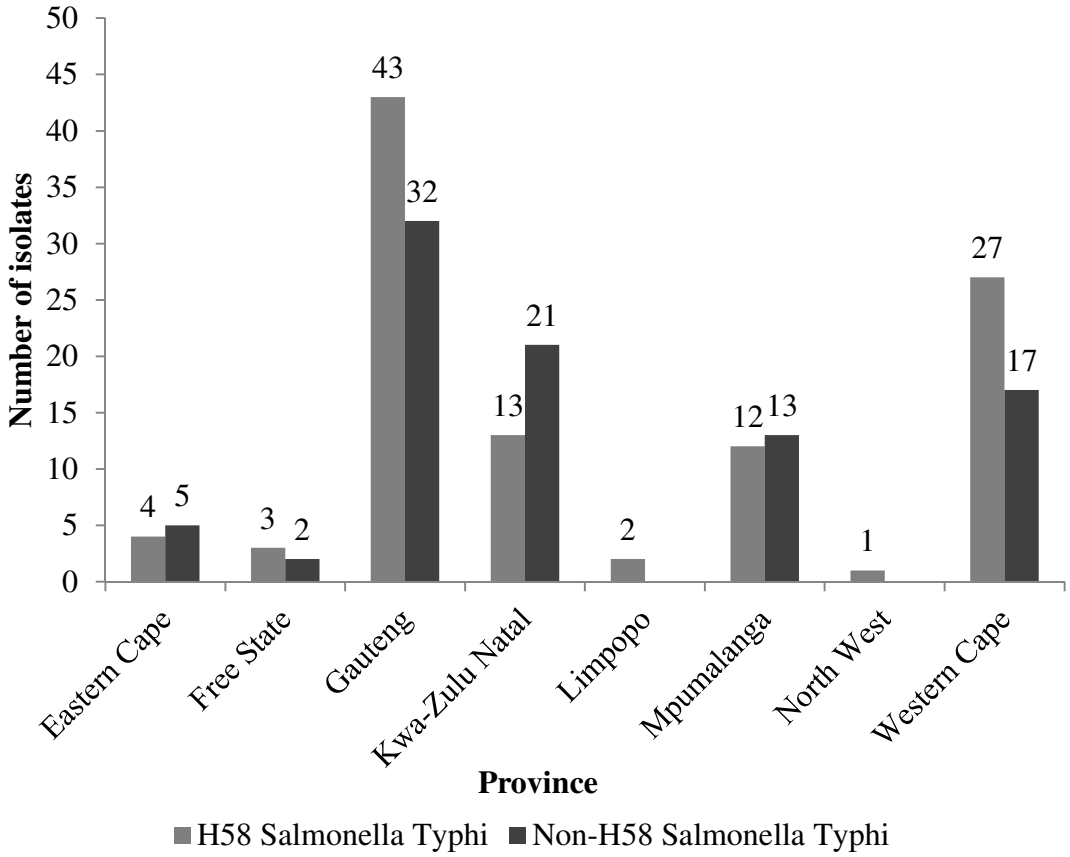


Figure 4.5 Bar graph representing *Salmonella* Typhi haplotypes identified in eight provinces in South Africa for the years 2012 to 2014.

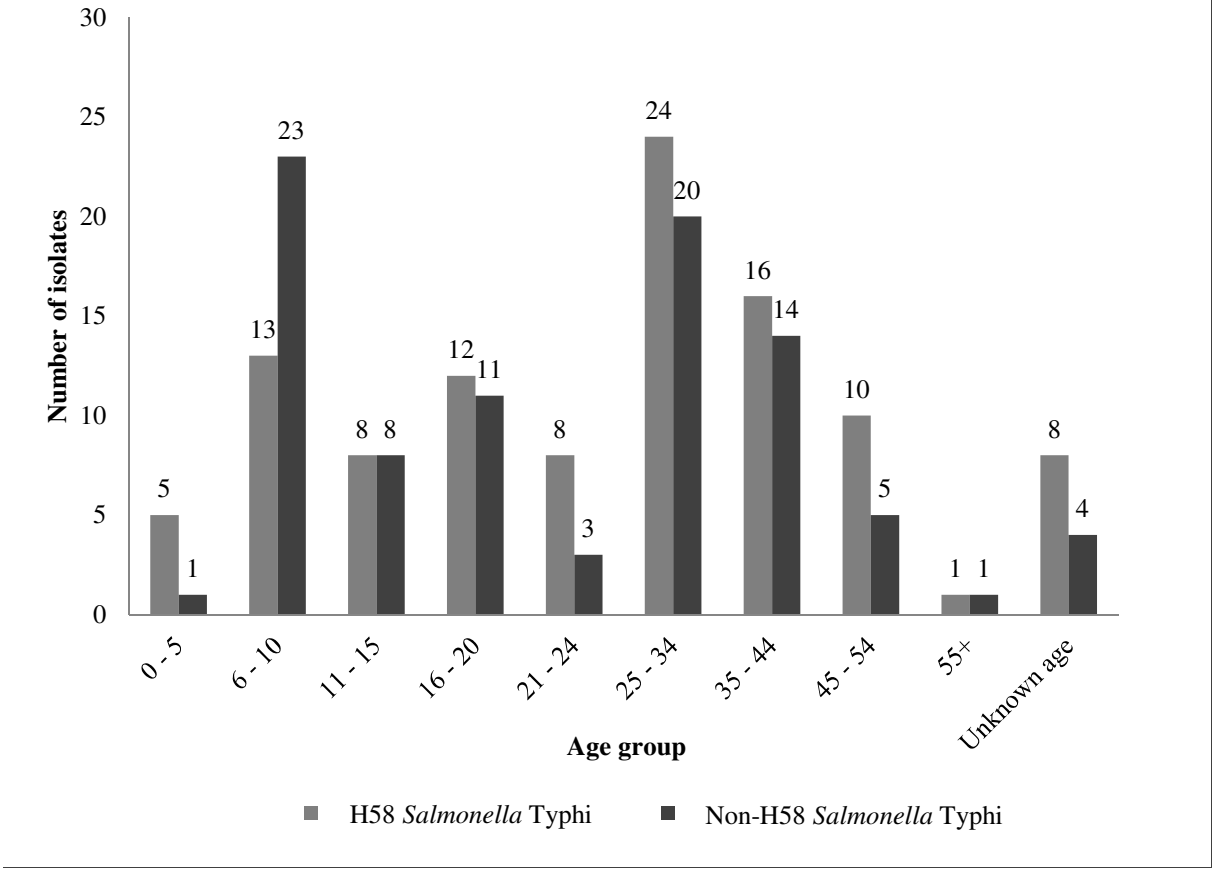


Figure 4.6 The Bar graph depicts haplotype specific typhoid fever infections by age group in South Africa, for the years 2012 to 2014.

4.3.2 Multi-Country Typhoid Fever Surveillance in sub-Saharan Africa Program (TSAP Study)

4.3.2.1 Real-time PCR confirmation of serotype

Real-time PCR assay targeting a gene unique to *Salmonella* species and an additional gene unique to *Salmonella* Typhi was used to confirm the serotype of *Salmonella* Typhi isolates collected from eight countries in the SSA. A total of 121 isolates were identified as *Salmonella* Typhi. Most of the isolates were collected from Kenya (n=56), followed by Ghana (n=34) and Burkina Faso (n=12) (Figure 4.7).

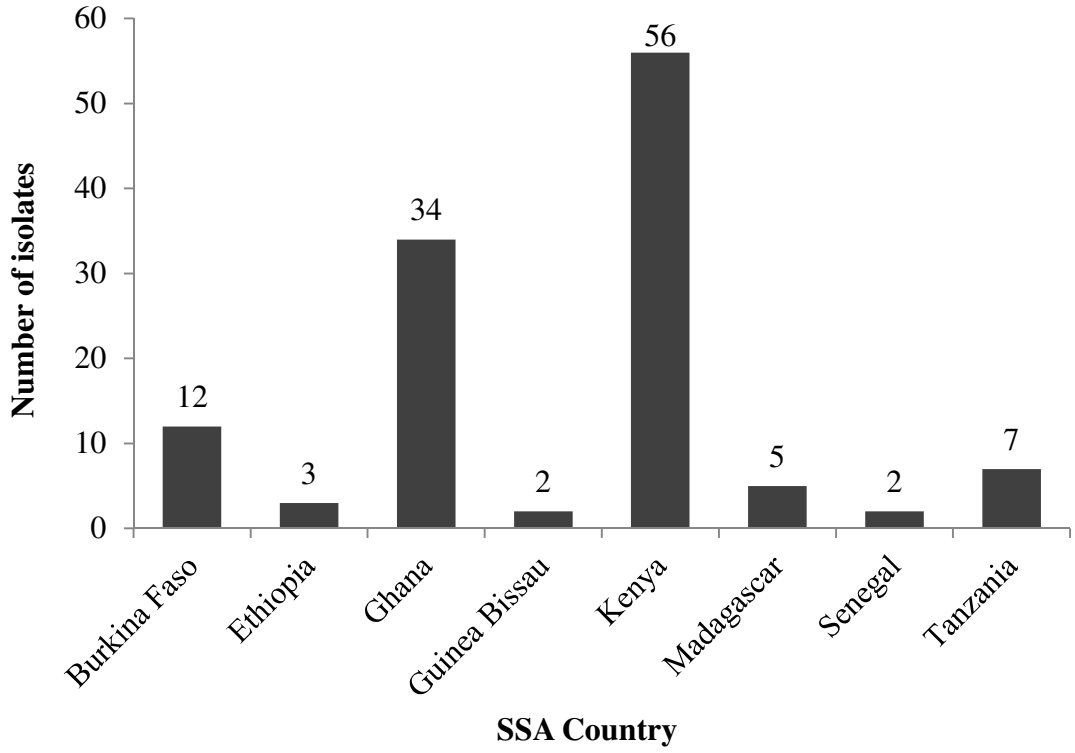


Figure 4.7 The bar graph showing the number of *Salmonella* Typhi isolates detected in SSA countries through the TSAP study for the years 2012 to 2014

4.3.2.2 PCR Screening of H58 *Salmonella* Typhi

All 121 *Salmonella* Typhi isolates were subjected to the H58 *Salmonella* Typhi conventional PCR. Of these, 75 isolates (62%) belonged to the H58 haplotype. The H58 *Salmonella* Typhi was identified in all but one of the SSA countries in the TSAP study (Figure 4.8). A total of two *Salmonella* Typhi isolates were collected from Guinea Bissau during the study period. These isolates were identified as non-H58 *Salmonella* Typhi.

The results from the conventional PCR also indicated that the H58 haplotype was predominant in both Kenya and Tanzania, as non-H58 *Salmonella* Typhi isolates were not identified. Of the two isolates collected from Senegal, only one belonged to the H58 haplotype. Three of the five isolates from Madagascar and two of three isolates from Ethiopia were identified as H58 *Salmonella* Typhi. Although the H58 haplotype was identified in Burkina Faso and Ghana, this highly resistant strain is not the predominant cause of typhoid

fever in these countries. Most of the *Salmonella* Typhi isolates in Burkina Faso (92%, n=11) and Ghana (85%, n=29) were identified as non-H58 *Salmonella* Typhi isolates (Figure 4.9).

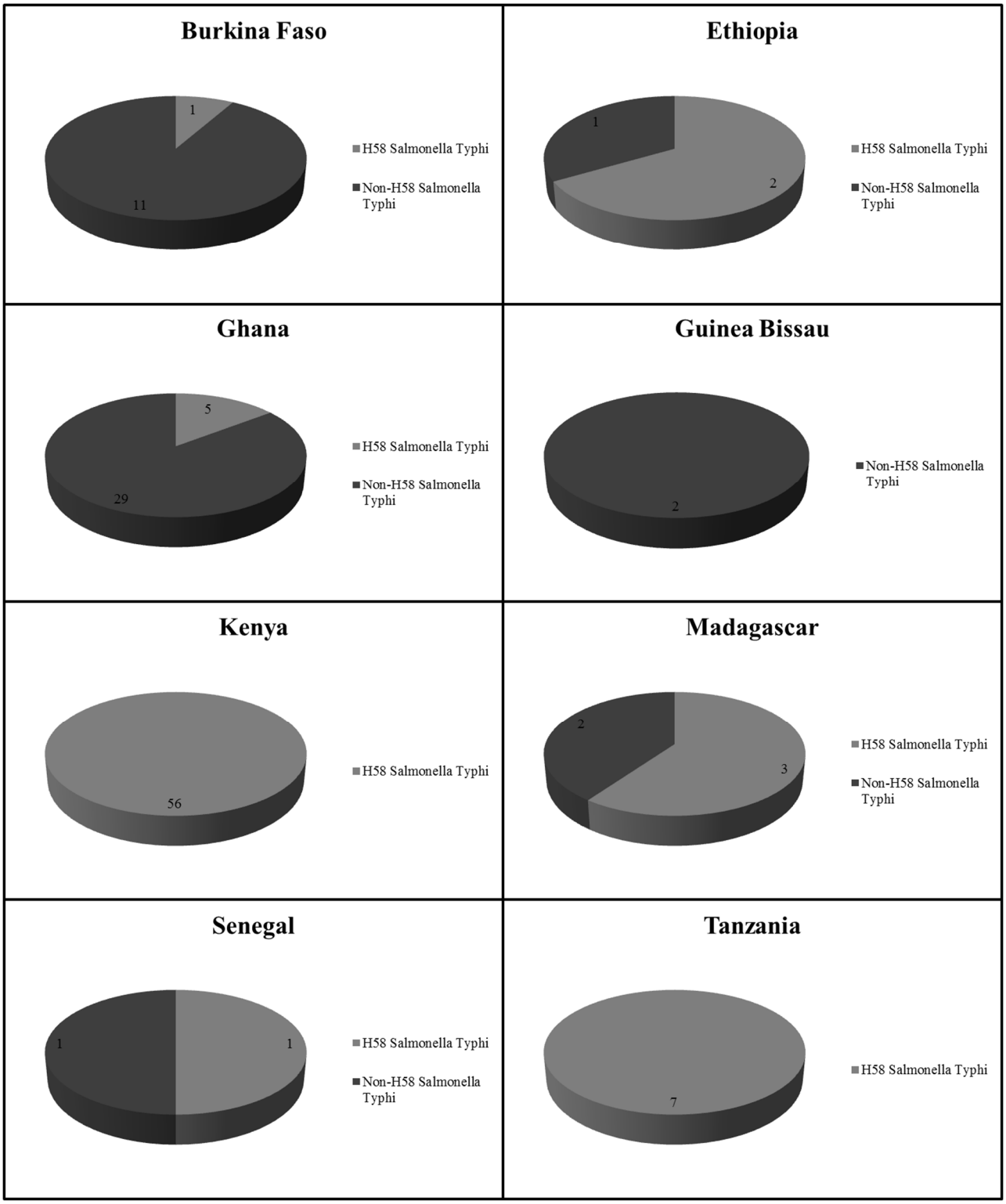


Figure 4.8 Pie charts showing the prevalence H58 *Salmonella* Typhi isolates in SSA countries.

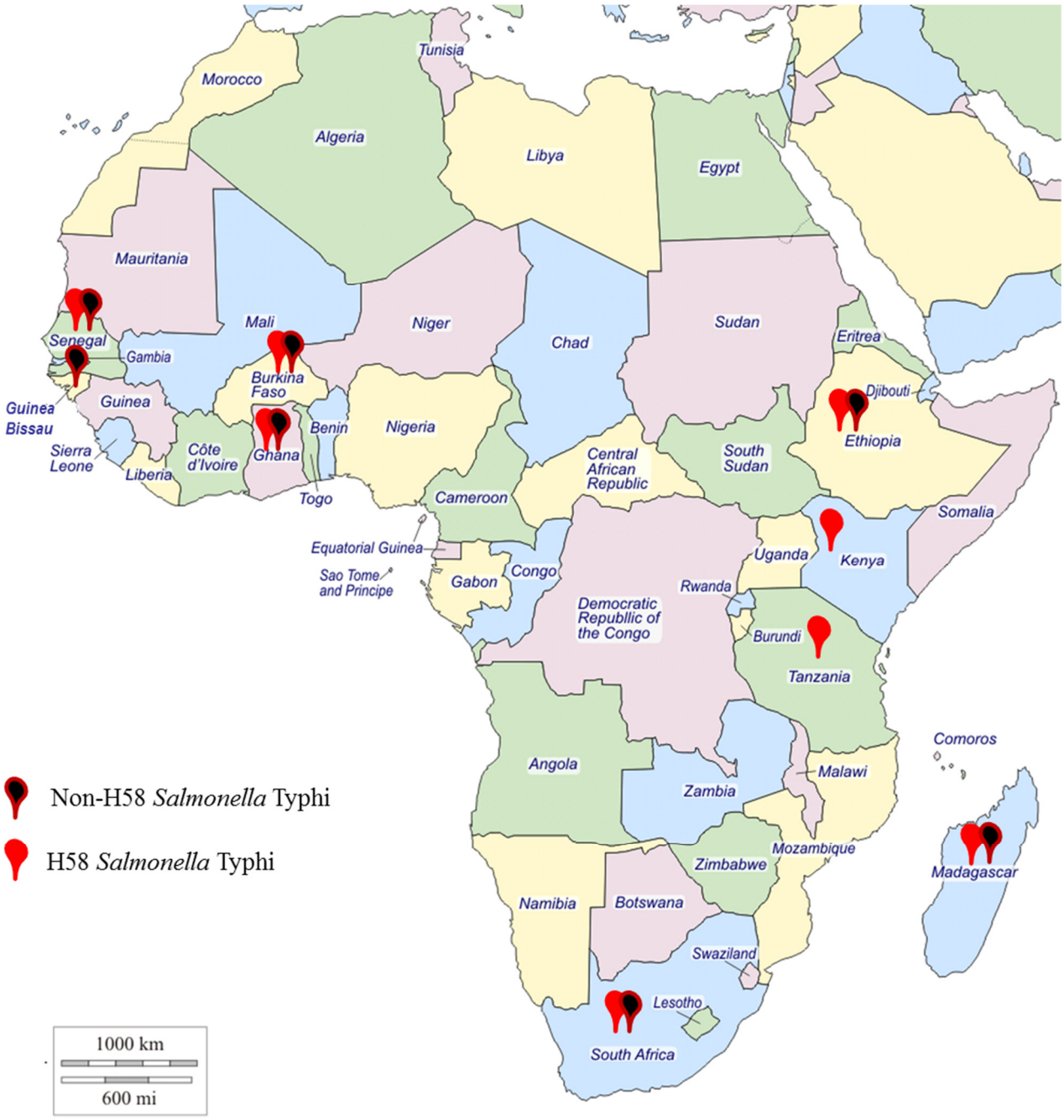


Figure 4.9 Map of Africa depicting the presence of H58 *Salmonella* Typhi in countries involved in typhoid fever surveillance through TSAP and GERMS-SA studies.

4.4 Discussion

In this study, a simple and inexpensive PCR assay was employed for the rapid identification of H58 *Salmonella* Typhi. The PCR assay targeted a 993bp deletion which is highly conserved in the H58 haplotype and was suitable for accurate identification of this haplotype. All H58 *Salmonella* Typhi isolates were identified and distinguishable from non-H58 *Salmonella* Typhi isolates. This proved to be a useful molecular tool for the preliminary screening of *Salmonella* Typhi isolates and can be accessible to most laboratories in the developing countries.

Data from SSA has suggested the presence of the H58 haplotype in some parts of the continent. Approximately 62% of the *Salmonella* Typhi isolates collected through the TSAP study belonged to H58 haplotype. Data also indicated that all isolates from Kenya (n=56) belonged to this haplotype. The emergence of the H58 haplotype in Kenya was first described in 2010 by Kariuki and colleagues, where they reported that approximately 75% of *Salmonella* Typhi isolates from Kenya belonged to the H58 haplotype [29]. It is very concerning that isolates collected from the TSAP surveillance site in Kenya were all identified H58 *Salmonella* Typhi isolates. This correlates with observations by Kariuki and colleagues that the H58 haplotype has rapidly spread in Kenya and is replacing the antimicrobial susceptible *Salmonella* Typhi [29]. Over-the-counter sales of fluoroquinolones without prescription in Kenya is a contributing factor to the misuse of antimicrobials which has resulted in emergence of antimicrobial resistant pathogens [22].

The H58 *Salmonella* Typhi seems to be spreading beyond the borders of Kenya. All isolates (n=7) collected in Tanzania for the period of the study belonged to the H58 haplotype. Additionally, 67% (n=2) of the isolates collected from Ethiopia were H58 *Salmonella* Typhi. There were H58 *Salmonella* Typhi isolates detected from Madagascar. Approximately 60% (n=3) of isolates collected from Madagascar belonged to the H58 haplotype. The majority of isolates collected in Ghana and Burkina Faso belonged to the non-H58 *Salmonella* Typhi haplotypes and accounted for approximately 85% (n=29) and 92% (n=11) of all isolates, respectively. Of the two isolates collected from Senegal during the study period, one belonged to the H58 haplotype. The H58 *Salmonella* Typhi was not detected in Guinea Bissau. Our data is in agreement with that of Wong and colleagues (2015) indicating

predominance of the H58 *Salmonella* Typhi in the eastern and southern parts of Africa; which suggests an on-going clonal replacement of the non H58 haplotypes.

In South Africa, an increase in the number of H58 *Salmonella* Typhi was observed for the years 2012 to 2014. During the period of the study, the number of H58 *Salmonella* Typhi isolates in South Africa increased from 50% in 2012 to 64% in 2014. This was particularly alarming as H58 *Salmonella* Typhi is known for its MDR characteristic coupled with reduced susceptibility to fluoroquinolones. The H58 haplotype was identified in patients of all age groups, but was predominant amongst patients aged 25-34 and 35-44 years. Interestingly, the H58 haplotype was identified as the predominant cause of infection in children aged 0-5 years. Fluoroquinolones are generally not recommended for treatment of typhoid fever in children under the age of 18 years due to major problems associated with bones, joints and join tissues [203]. Therefore, it is not unexpected that the majority of typhoid fever cases in this age group are not caused by the H58 *Salmonella* Typhi. Exposure to the H58 *Salmonella* Typhi in age groups 0-5 years, 6-10 years and 11-15 years could be as a result of contact with adult carriers and not necessarily of selection pressure. Qin and colleagues (2006) studied ciprofloxacin resistance in gram-negative bacilli from the faecal microflora of children that were not exposed to the antimicrobial agent [203]. In their study, they demonstrated that resistant bacteria belonging to species that would normally be susceptible to fluoroquinolones can be found in the intestinal tracts of individuals who were not exposed to fluoroquinolones. This suggested that bacteria that are resistant to clinically useful antimicrobial agents can be acquired by humans without experiencing selective pressure due to either direct or indirect exposure to the antimicrobial agents. Therefore, since fluoroquinolones are not used for the treatment of typhoid fever in children, the presence of the H58 haplotype in this age group could be through contact with adult typhoid carriers rather than antimicrobial selection pressure.

H58 *Salmonella* Typhi was mostly identified in Gauteng and in the Western Cape provinces (Figure 4.5). The reason for the detection of high typhoid fever incidences in these provinces could include; (a) that many of the South African population relocate to these provinces to seek employment opportunities as the unemployment rate in Gauteng and Western Cape provinces is higher than anywhere else in the country; (b) the Gauteng and Western Cape

province are two of the most popular tourist destinations in South Africa; (c) additionally, over the years, the Gauteng Province has become home to many refugees from many impoverished countries in the SSA. Even though H58 *Salmonella* Typhi was detected in Mpumalanga, Kwa-Zulu Natal and the Eastern Cape, other haplotypes of *Salmonella* Typhi were detected, accounting for more than half the isolates collected over the study period.

Study limitations

There were no cases of typhoid fever reported from the Northern Cape Province during the study period; this could be due to a couple of reasons including (1) underreporting of laboratory-confirmed cases to the GERMS-SA laboratory-based surveillance program and (2) various provinces in SA could be involved in more intense surveillance as compared to others. Similarly, the difference in the number of isolates collected from the various countries in SSA could be an effect of variation in surveillance efforts and not a reflection of the burden of typhoid fever disease in those countries.

We could not perform statistical calculations for the difference in number of H58 *Salmonella* Typhi isolates identified in various age groups as well as in different provinces in SA, as there were not enough variables. This would have been possible if we compared the number of H58 *Salmonella* Typhi isolates in age groups or provinces per year. This was not done in our study.

4.5 Conclusion

The predominance of H58 *Salmonella* Typhi in South Africa as well as other parts of the SSA region is concerning. In this country, the spread and transmission of typhoid fever is believed to be fuelled by lack of access to clean water and proper sanitation especially in countries where stringent legislations regarding over-the-counter sales of antimicrobials exists. The increase in the number of H58 *Salmonella* Typhi threatens successful treatment of typhoid fever. The inappropriate and misuse of antimicrobials, fuelled by over-the counter sale of antimicrobials, is believed to contribute greatly to antimicrobial resistance. The enforcement of legislations that restrict over-the-counter dispensation of antimicrobials without prescription could greatly reduce emergence of highly resistant strains which occur as a result of antimicrobial selection pressure.

There is evidence of intracontinental spread of H58 *Salmonella* Typhi in the SSA. Therefore, continued surveillance of typhoid fever is crucial for monitoring disease spread and to inform prevention and control strategies.

Chapter 5

Multiple-locus variable-number tandem-repeats analysis of *Salmonella* Typhi isolates collected from eight sub-Saharan African countries, 2012 - 2014

5.1 Introduction

MLVA typing has been used successfully for molecular typing of *Salmonella* Enteritidis as well as *Salmonella* Typhi [153, 154]. In this study, we used MLVA typing to study the molecular epidemiology of *Salmonella* Typhi isolates collected from eight countries from within SSA including, Burkina Faso, Ethiopia, Ghana, Guinea Bissau, Kenya, Madagascar, Senegal, South Africa (through GERMS-SA and TSAP surveillance programs) and Tanzania through the TSAP study. The TSAP is a network of 13 sentinel sites within ten countries in SSA (Figure 5.1) [187]. This program was established in SSA with the aim of introducing standardized multi-country surveillance for typhoid fever and invasive non-typhoidal *Salmonella.* Along with their collaborators, the TSAP has managed to strengthen local surveillance by improving disease surveillance, diagnosis of bacterial disease analysis and reporting systems [187].

5.1.1 The study objectives:

a) To use MLVA assay to investigate the molecular epidemiology of *Salmonella* Typhi isolates from SSA
b) To investigate whether H58 *Salmonella* Typhi is associated with a particular MLVA profile
c) To compare MLVA molecular subtyping data to the PFGE genotyping data

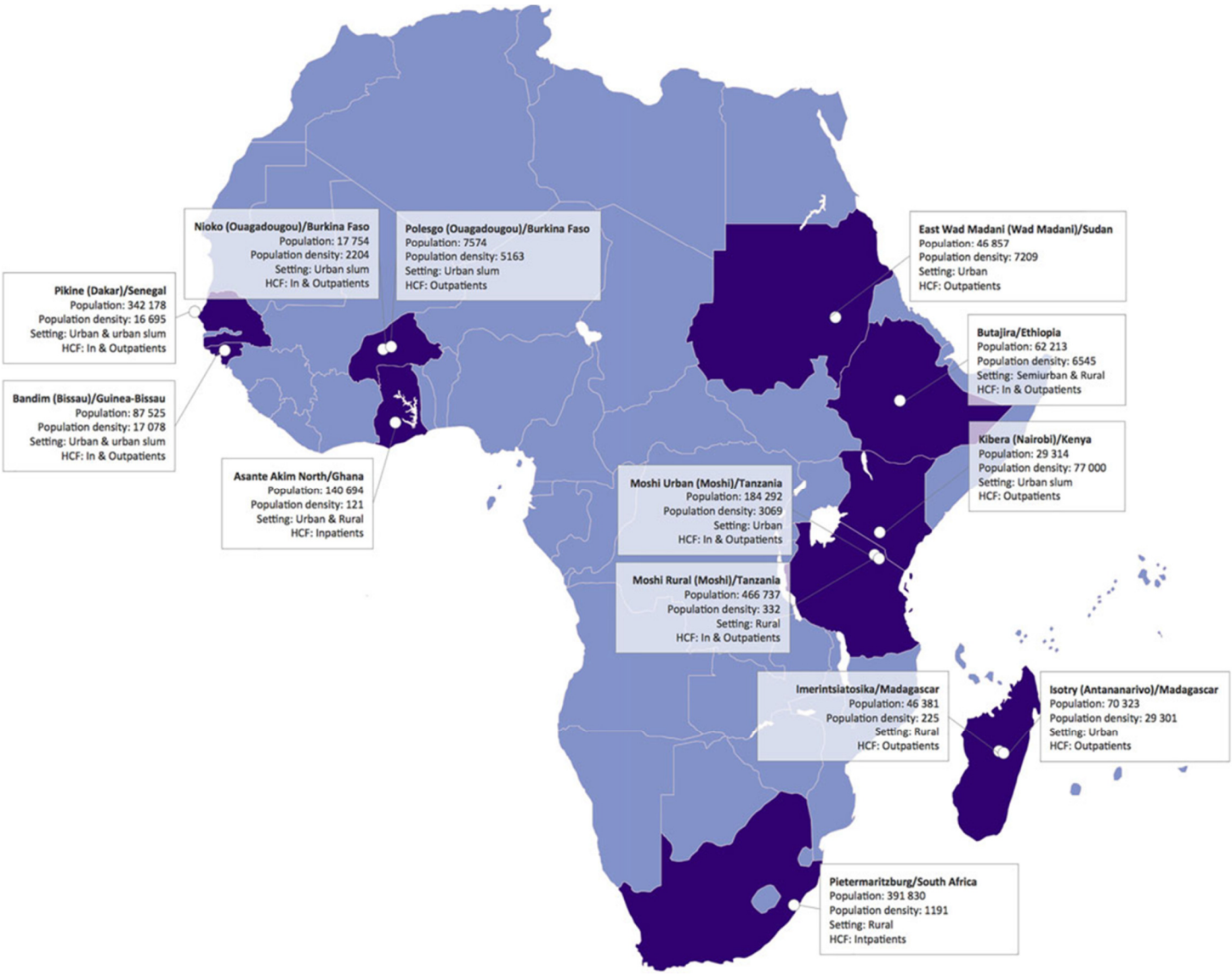


Figure 5.1 Thirteen TSAP sentinel sites[187]

5.2 Materials and Methods

5.2.1 *Salmonella* Typhi isolates

A total of 316 *Salmonella* Typhi isolates from eight countries in the SSA continent were analysed. The isolates were collected through two surveillance studies, GERMS-SA and TSAP described in Chapter 2, section 2.1.

5.2.2 Pulsed-field gel-electrophoresis (PFGE)

PFGE was performed as described in Chapter 2 (section 2.5).

5.2.3 Multiple-locus variable-number tandem-repeats analysis (MLVA)

MLVA typing was performed by multiplex PCR amplification of the five most variable VNTR loci as described in Chapter 3 (section 3.2.2.6).

A minimum spanning tree (MST) was generated in order to display the distance in relation between the *Salmonella* Typhi isolates. The distance was calculated based on the number of different VNTR loci between two MLVA profiles. For instance: if two MLVA profiles differed at one VNTR locus, the distance between them would be 'one', thus indicating close relation between the MLVA types. Furthermore, if two MLVA profiles differed at two VNTR loci, the distance between these MLVA types would be 'two', and so forth. The majority of isolates in this study differed by either one or two VNTR loci. There was no apparent major cluster in the MST; however, six small clusters consisting of four to sixteen isolates with indistinguishable MLVA profiles were identified.

5.2.4 Data analysis

The Simpson's DI and 95% confidence intervals (CIs) were calculated using an online tool available at (www.comparingpartitions.info). The congruence between the MLVA assay and PFGE analysis the Wallace coefficient was determined, via an online tool (www.comparingpartitions.info).

5.3 Results

5.3.1 Diversity of *Salmonella* Typhi isolates analysed by MLVA

Three hundred and sixteen *Salmonella* Typhi isolates from SSA were analysed using a MLVA assay consisting of five VNTR loci. All five VNTR loci were amplified in one multiplex PCR reaction using fluorescently labelled primers. The PCR amplicons were separated based on fragment size by multicolour capillary electrophoresis Genetic analyzer. The loci sizes for each VNTR locus were inferred numbers based on bin fragment size. The loci sizes for each VNTR locus were then combined into a string of integers referred to as a MLVA profile. As an example, 197-487-309-196-241 is a string of integers representing amplicon fragment sizes for VNTR loci TR1-TR2-Sal02-Sal20-TR4699 respectively. This particular string of integers or MLVA profile was assigned MLVA type STyMT-86 (*i.e. Salmonella* Typhi MLVA type-86). Altogether, 226 MLVA types were identified amongst the *Salmonella* Typhi isolates from SSA (Appendix G).

The MST revealed two major clonal complexes (clonal complex I and clonal complex II) which consisted of the majority of isolates (n/N=252/316) (Figure 5.2). It was evident that the *Salmonella* Typhi isolates from the different SSA countries were distributed throughout the MST and not limited to one clonal complex as no branching was completely dominated by isolates from the same country. Similarly, the H58 clone was not limited to a single clonal complex (Figure 5.3). This haplotype was dispersed throughout the MST indicating that heterogeneity exists within the H58 clonal group. Even so, the majority of isolates (n/N=83/135) in clonal complex I belonged to non-H58 *Salmonella* Typhi haplotype.

5.3.2 MLVA typing of isolates from South Africa

MLVA typing of 195 isolates from South Africa indicated that diversity exists amongst the *Salmonella* Typhi isolates (Figure 5.4). In total, 155 MLVA types were identified amongst the South African *Salmonella* Typhi isolates. Although no major clusters were identified, seven small clusters consisting of three to four isolates were identified. These were STyMT-121, STyMT-114, STyMT-102, STyMT-130, STyMT-132, STyMT-136 and STyMT-139. Analysis of *Salmonella* Typhi isolates using MLVA did not indicate predominance of any MLVA type in the country. There was high variability observed amongst the isolates from

South Africa with 66 MLVA types identified in Gauteng, 34 in Kwa-Zulu Natal, 30 in Western Cape, 24 in Mpumalanga, 9 in Eastern Cape, 2 in Limpopo and 1 in North West.

Several MLVA types were shared between provinces and included STyMT-100, STyMT-107, STyMT-114 and STyMT-136 identified in isolates from Gauteng and Western Cape; STyMT-112, STyMT-132 and STyMT-261 identified in Gauteng and Mpumalanga; MLVA types STyMT-102 and STyMT-139 identified in Gauteng and Eastern Cape; MLVA type STyMT-134 identified in Kwa-Zulu Natal and Limpopo; and MLVA type STyMT-111 identified in Eastern Cape and Western Cape provinces.

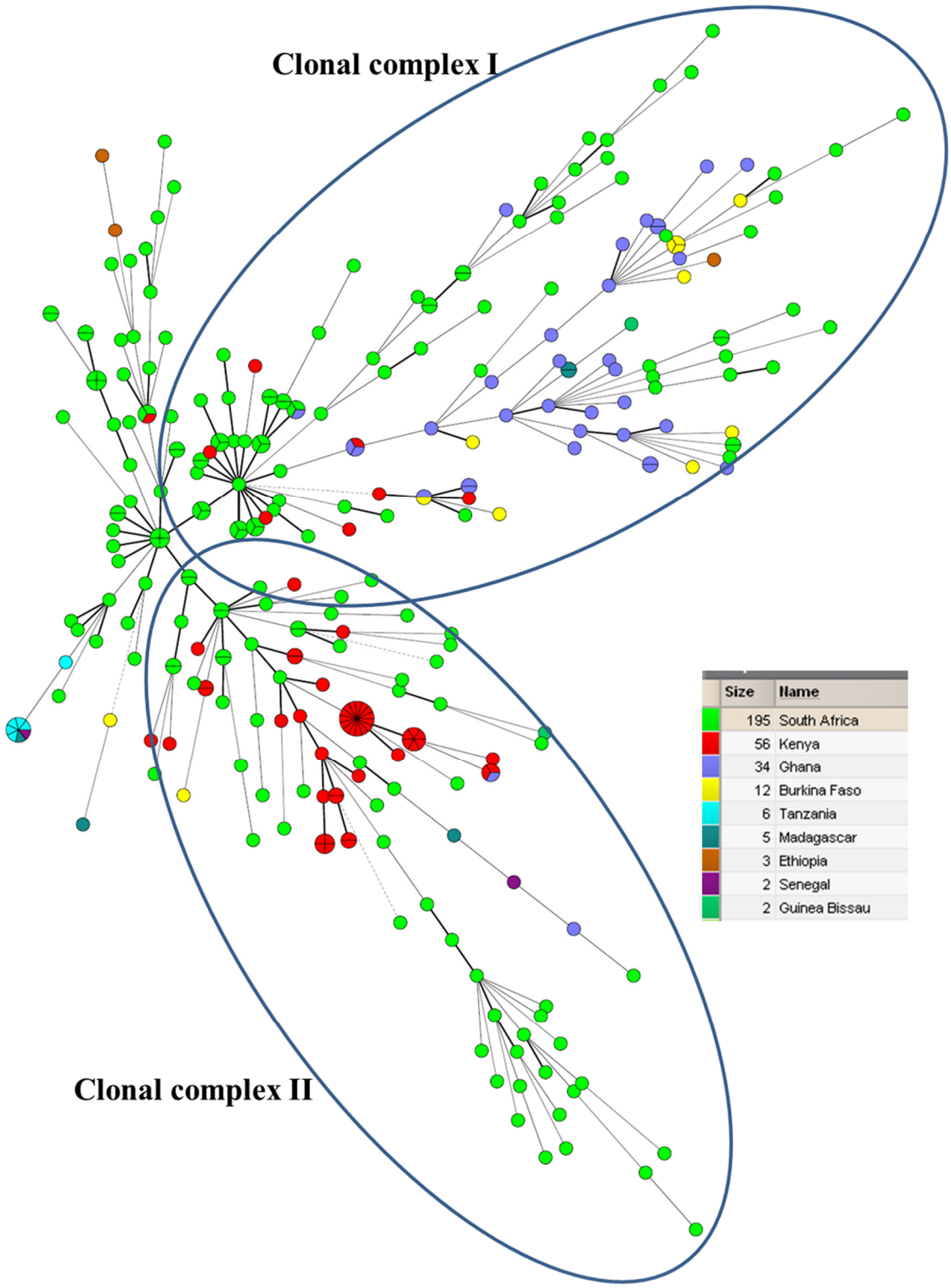


Figure 5.2 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from SSA collected over a three year period (n=316). Each node represents a different MLVA profile. Node colour denotes country of isolate origin. Branch thickness indicated difference in the number of loci between MLVA profiles. Thick solid lines connect nodes that differ at one VNTR locus, thin lines connect nodes that differ at two VNTR loci, dashed lines connect nodes that differ at three VNTR loci and dotted lines connect nodes that differ at four or more VNTR loci.

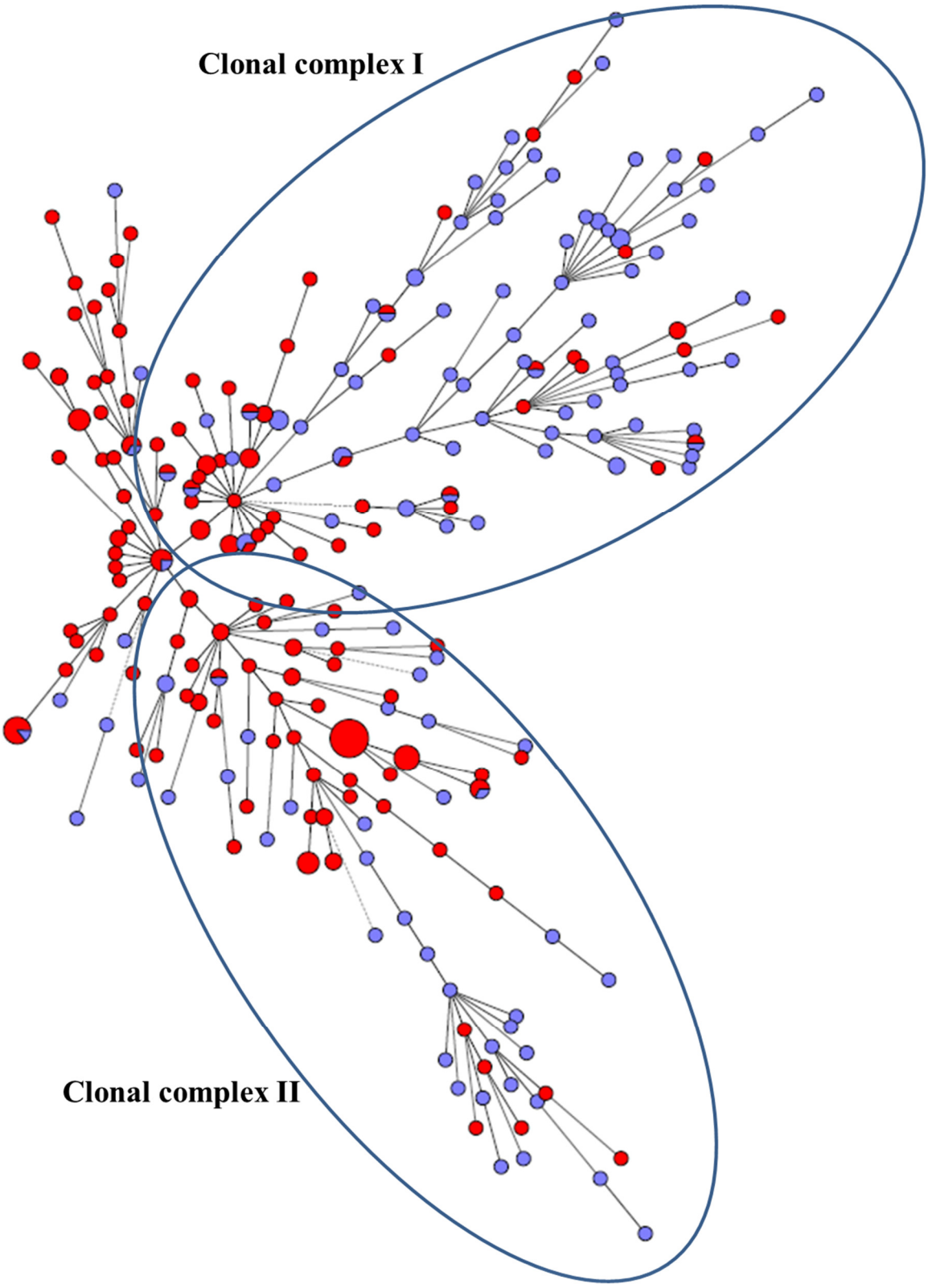


Figure 5.3 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from SSA collected over a three year period (n=316). Each node represents a different MLVA profile. Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates.

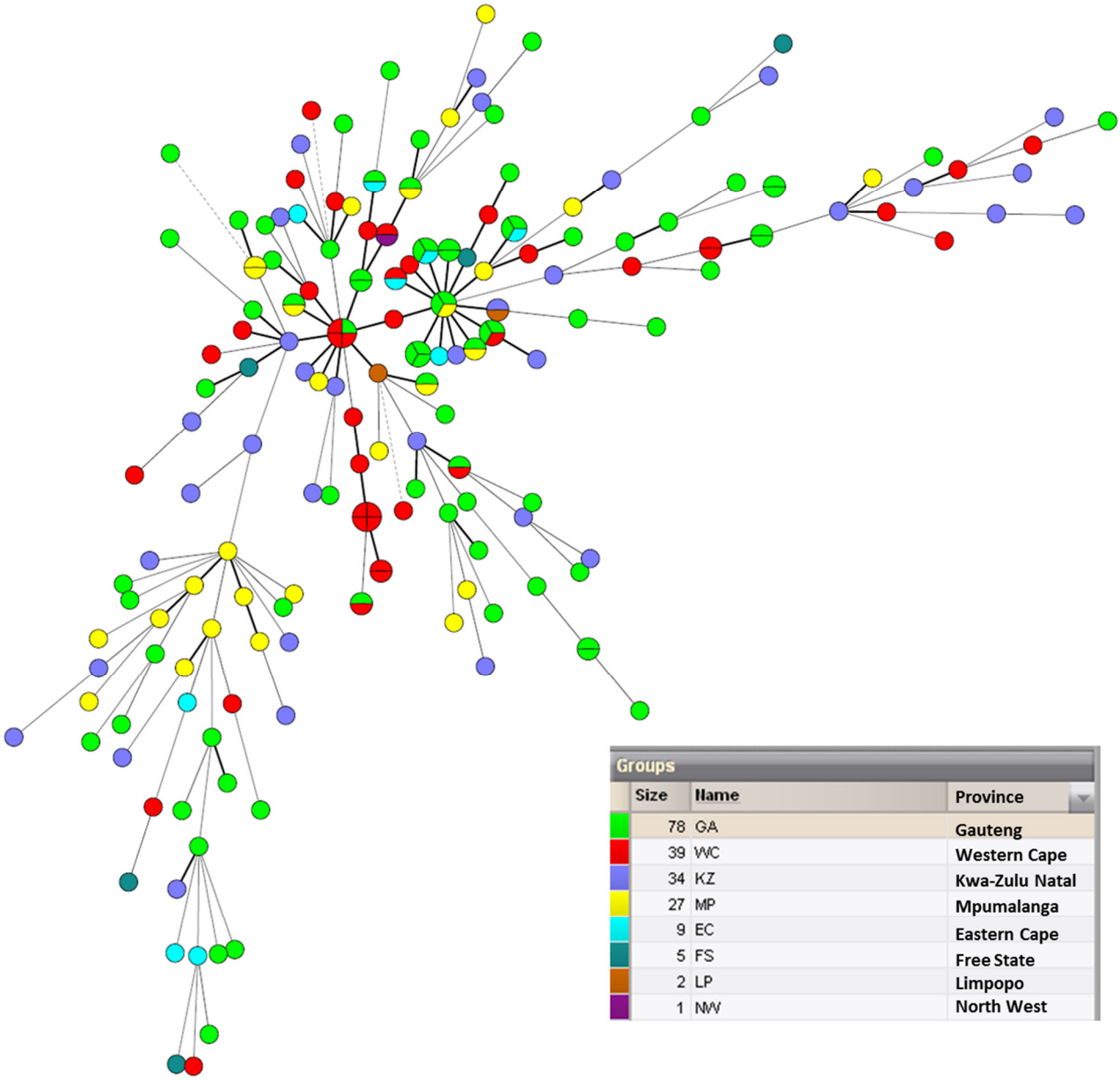


Figure 5.4 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected over a three year period (n=195). Each node represents a different MLVA profile. Node colours denote province of isolate origin.

5.3.3 Annual distribution of MLVA types amongst *Salmonella* Typhi in South Africa

A total of 57 *Salmonella* Typhi isolates collected during the year 2012 were analysed using MLVA assay. There were 45 MLVA types identified from these isolates (Figure 5.5A). There was no evidence of a predominant MLVA type; however, small clusters consisting of two to four isolates were identified and these were MLVA types STyMT-121, STyMT-112, STyMT-239, STyMT-125 and STyMT-100. MST revealed a clonal complex consisting of isolates that belong to the H58 haplotype. The H58 *Salmonella* Typhi was associated with MLVA type and these included MLVA types STyMT-239, STyMT-112, STyMT-127, STyMT-125 and STyMT-100 which consisted of two to four isolates each (Figure 5.5 B).

From the 58 *Salmonella* Typhi isolates collected during the year 2013, 52 MLVA types were identified (Figure 5.6 A). MST revealed two clonal complexes from which small clusters consisting of two isolates each were identified; these were STyMT-102, STyMT-136, STyMT-113, STyMT-56, STyMT-127, STyMT-261 and STyMT-87. The H58 clone was not limited to a single clonal complex as isolates belonging to this haplotype were dispersed throughout the MST (Figure 5.6 B). The MST revealed MLVA types consisting of both the H58 as well as non-H58 haplotype isolates. These MLVA types were identified as STyMT-261, STyMT-113 and STyMT-56. MLVA types which were associated with the H58 haplotype were STyMT-136 and STyMT-127 consisting of two isolates each.

A total of 73 MLVA types were identified from 85 *Salmonella* Typhi isolates collected during the year 2014 (Figure 5.7 A). Similar to the years 2012 and 2013, there was no evidence of a predominant MLVA type; however, small clusters consisting of two to three isolates each were identified. These MLVA type included, STyMT-132, STyMT-139, STyMT-173, STyMT-133, STyMT-111, STyMT-146, STyMT-59, STyMT-114 and STyMT-134. The MST revealed MLVA clusters consisting of both H58 and non-H58 *Salmonella* Typhi isolates, and these included STyMT-139, STyMT-173, STyMT-114 and STyMT-111 (Figure 5.7 B). MLVA types which were exclusively associated with the H58 haplotype included STyMT-132, STyMT-134 and STyMT-59.

During each year of the study, distinct MLVA types were detected; however, nine MLVA types were identified throughout the study period. MLVA types STyMT-114 and STyMT-130 were observed every year throughout the study; MLVA types STyMT-121, STyMT-132 and STyMT-159 were observed during the years 2012 and 2014; MLVA type STyMT-128 was observed during 2012 and 2013 and finally, MLVA types STyMT-102, STyMT-107 and STyMT-136 were observed during the years 2013 and 2014.

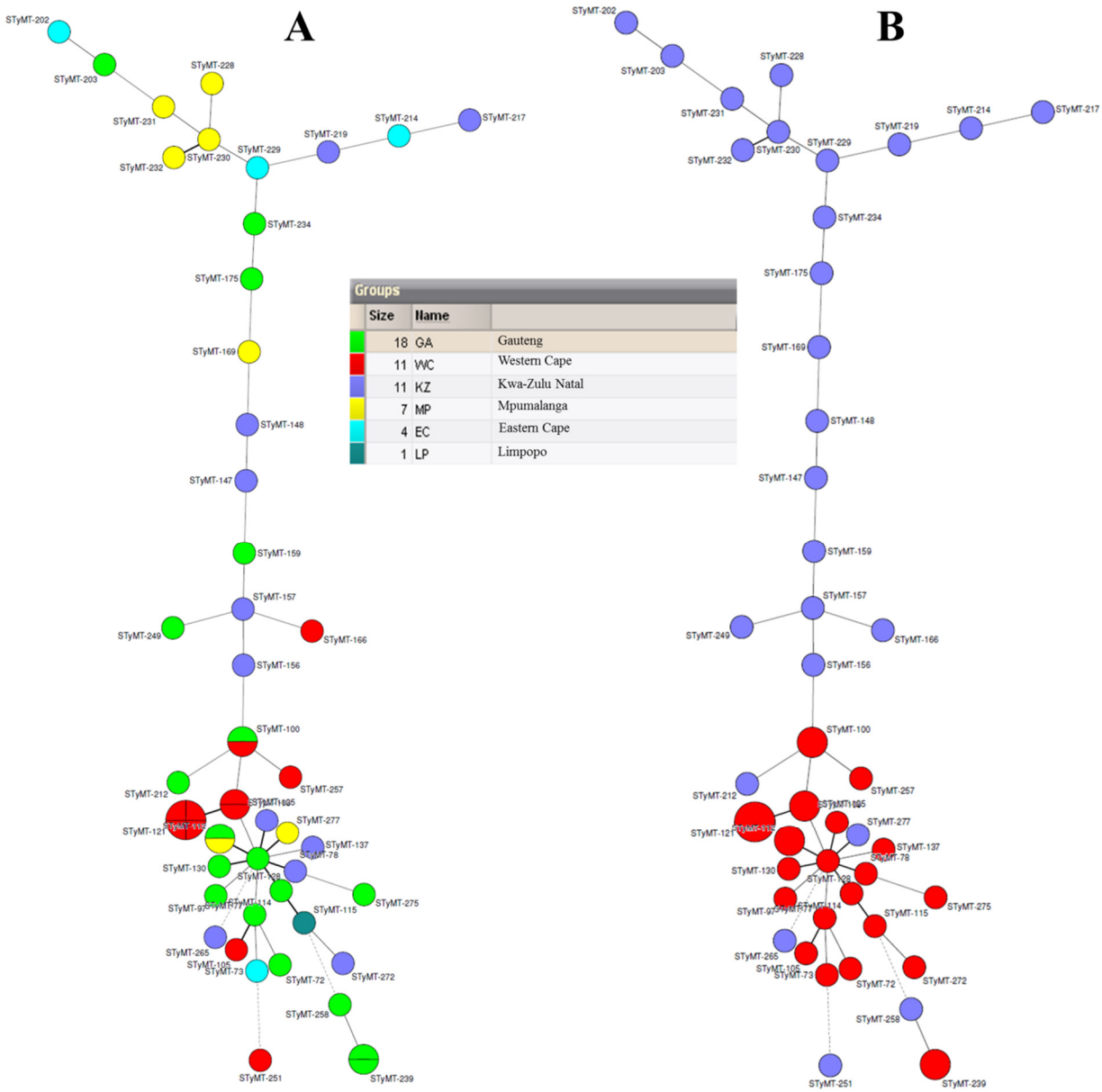


Figure 5.5 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2012 (n=57). Each node represents a different MLVA profile. (A) Node colours denote province of isolate origin and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates.

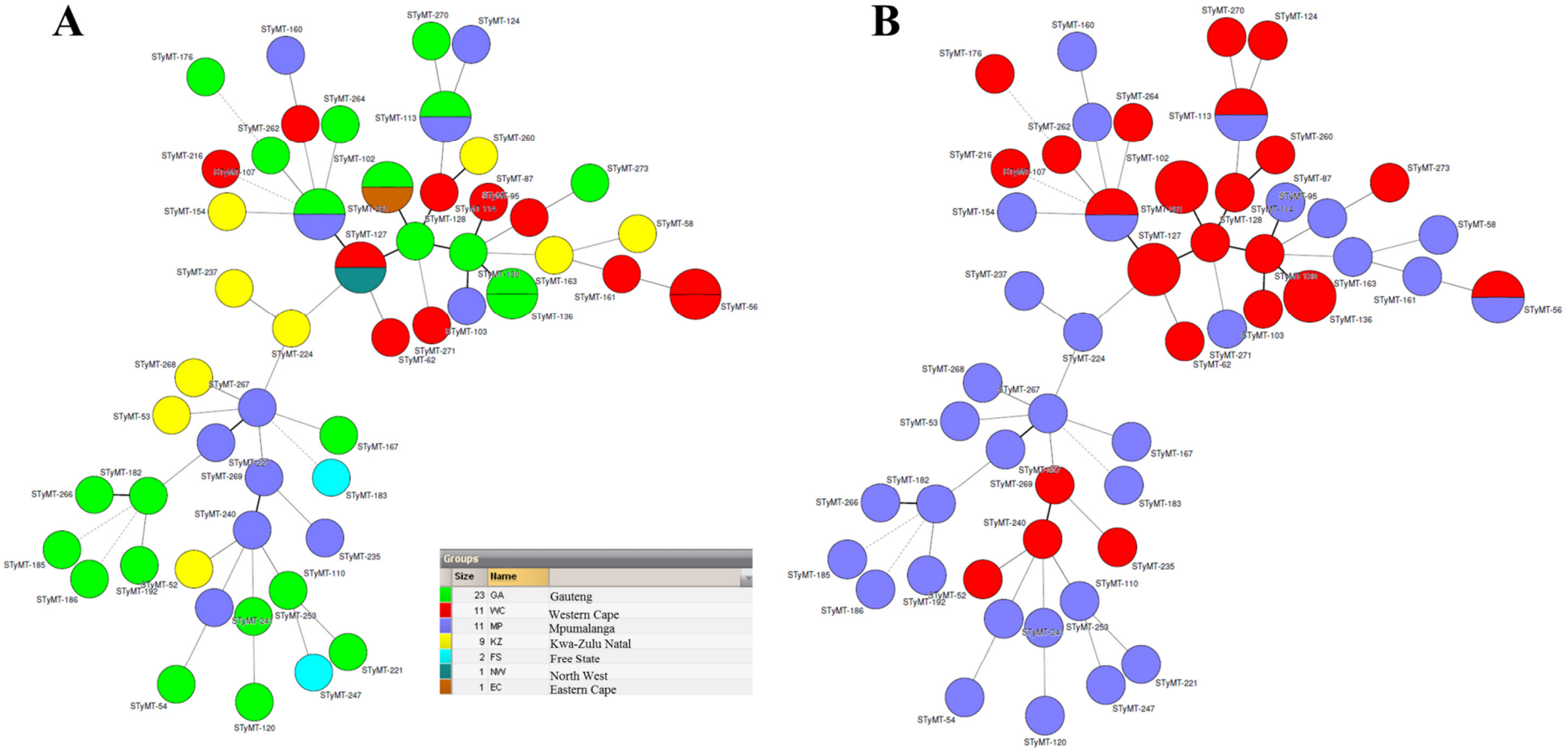


Figure 5.6 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2013 (n=58). Each node represents a different MLVA profile. (A) Node colours denote province of isolate origin and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates.

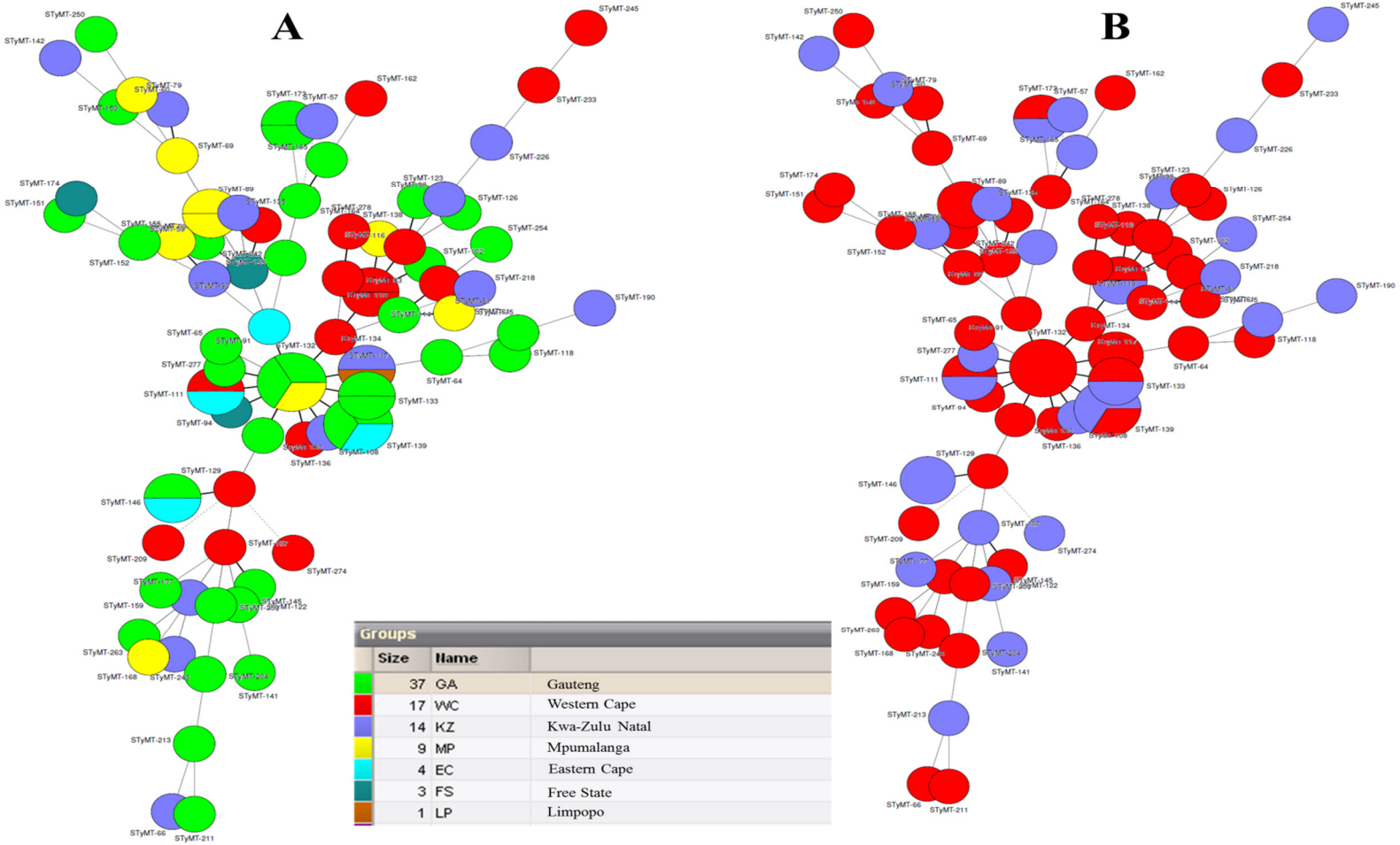


Figure 5.7 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from South Africa collected during the year 2014 (n=73). Each node represents a different MLVA profile. (A) Node colours denote province of isolate origin and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates**.**

5.3.4 MLVA typing of isolates from other sub-Saharan African Countries

5.3.4.1 Burkina Faso

MLVA typing was performed on 12 isolates collected in Burkina Faso during the years 2012 to 2014 (Figure 5.8 A). A total of ten MLVA types were identified. The MST did not reveal a predominant MLVA type. Even so, a cluster consisting of three isolates belonging to MLVA type STyMT-206 was identified. *Salmonella* Typhi isolates from this country showed heterogeneity. Only one MLVA type, STyMT-184, was associated with the H58 haplotype (Figure 5.8 B). H58 *Salmonella* Typhi was not predominant in Burkina Faso.

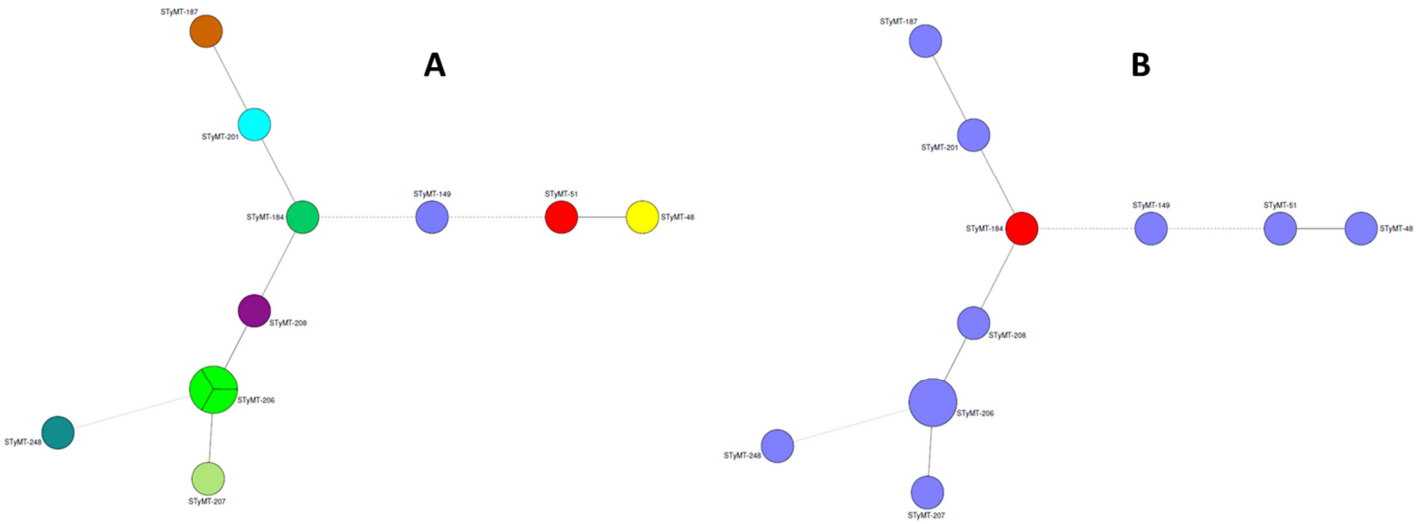


Figure 5.8 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Burkina Faso collected over a three year period (n=12). Each node represents a different MLVA profile. (A) Node colours denote MLVA profiles and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates**.**

5.3.4.2 Ethiopia

The three *Salmonella* Typhi isolates collected from Ethiopia during the study period showed diversity amongst the isolates. MLVA typing revealed three distinct MLVA types that differed at two or more VNTR loci. The H58 haplotype was associated with two MLVA types, namely STyMT-170 and 171; however, this is not irrefutable as only a small number of isolates was collected from this country.

5.3.4.3 Ghana

A total of 34 *Salmonella* Typhi isolates from Ghana were analysed using the MLVA assay (Figure 5.9 A). Of these, 30 MLVA types were detected. Even though there was no predominant MLVA type identified, small clusters consisting of two isolates each were identified. The five MLVA types consisting of small cluster of isolates included STyMT-223, STyMT-236, STyMT-195 and STyMT-49. We identified five MLVA types that were associated with the H58 clone, namely MLVA types STyMT-200, STyMT-158, STyMT-181, STyMT-180 and STyMT-220; even so, this is not absolute as only one isolate was identified for each MLVA type. Analysis of more isolates could confirm this finding or reveal non-H58 isolates belonging to these MLVA types (Figure 5.9 B). MLVA type STyMT-49 consisted of both the H58-clone and a non-H58 haplotype isolate.

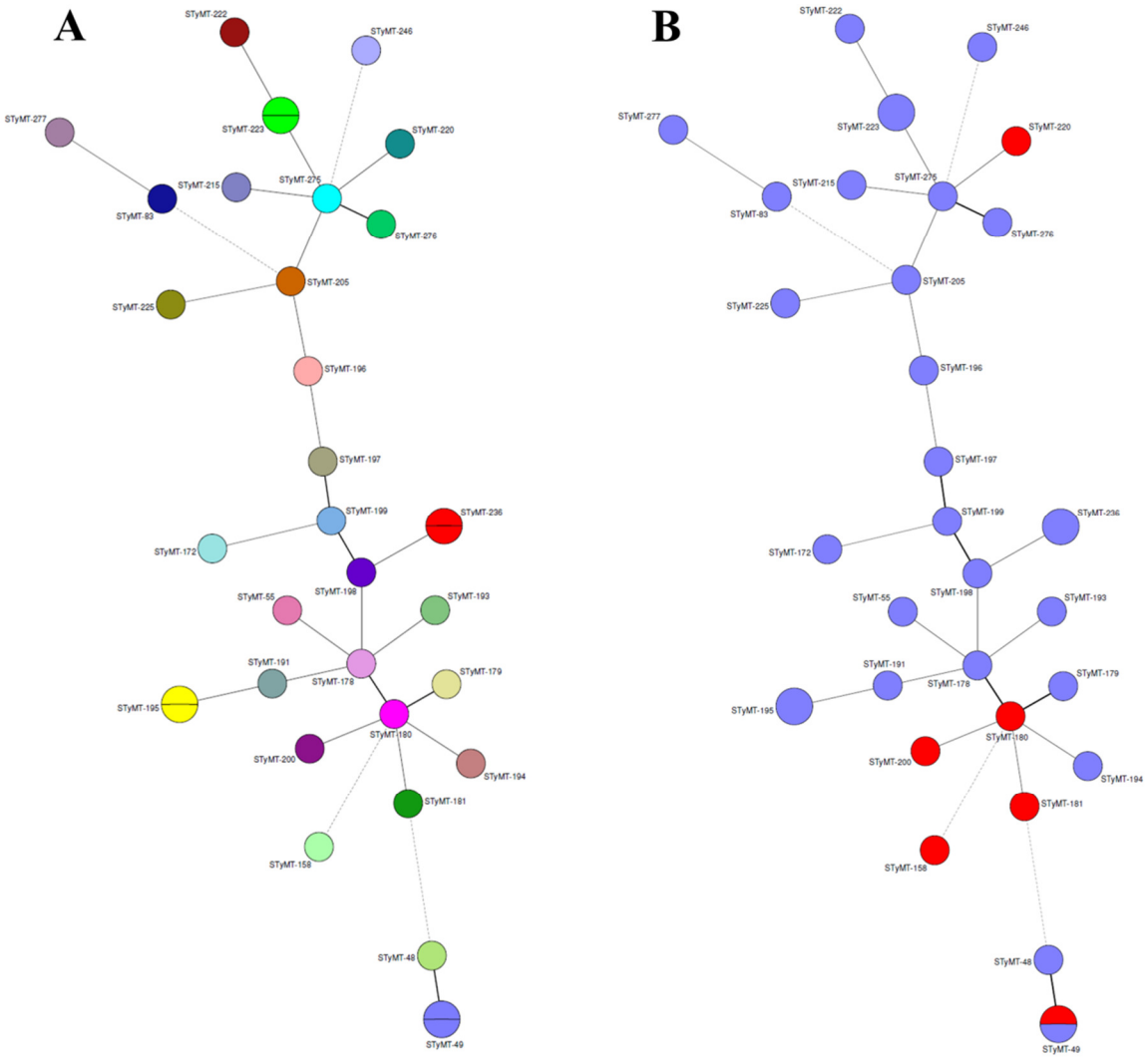


Figure 5.9 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Ghana collected over a three year period (n=34). Each node represents a different MLVA profile. (A) Node colours denote MLVA profiles and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates.

5.3.4.4 Guinea Bissau

Diversity was observed amongst the two isolates collected from Guinea Bissau during the study period. Two MLVA types were identified, namely STyMT-188 and STyMT-252, and these differed at 4 VNTR loci. The H58 clone was not associated with any MLVA type in this country as this clone was not identified in Guinea Bissau.

5.3.4.5 Kenya

MLVA typing was performed on 56 *Salmonella* Typhi isolates collected in Kenya during the year 2012 to 2014 (Figure 5.10 A). A total of 27 MLVA types were observed. The MST revealed that MLVA type STyMT-86 was predominant in Kenya. This MLVA type consisted of a cluster of 16 isolates. Several other clusters were identified amongst MLVA types STyMT-256 and STyMT-84 consisting of six and four isolates, respectively; MLVA types STyMT-80, STyMT-82, STyMT-83, STyMT-67 and STyMT-93 consisted of two isolates each. Kenyan *Salmonella* Typhi isolates appear to be more clonal that isolates from the rest of the continent (Figure 5.10 B). Given that all isolates from Kenya belong to the H58 clone, this haplotype was dispersed throughout the MST.

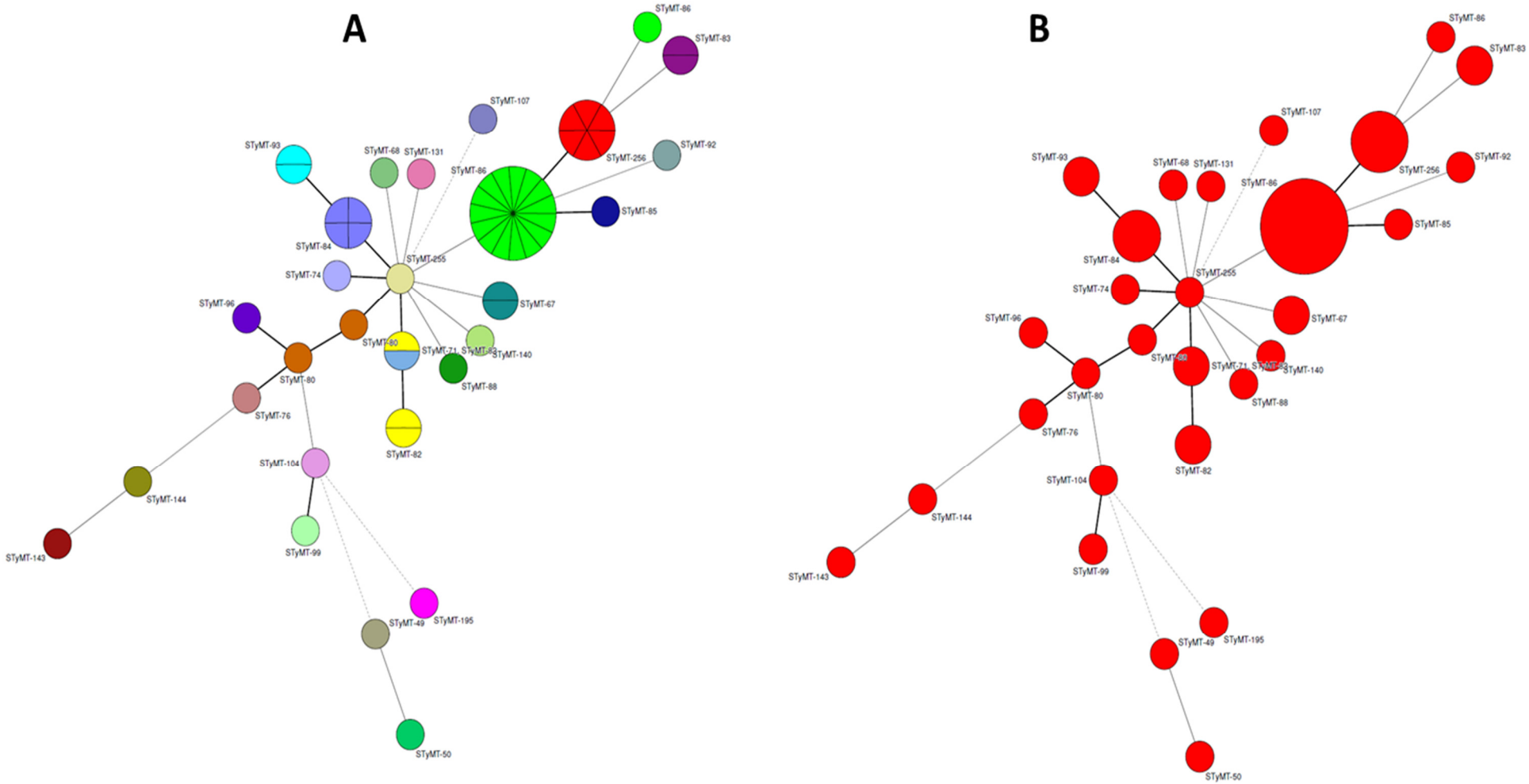


Figure 5.10 Minimum spanning tree calculated for MLVA profiles of *Salmonella* Typhi isolates from Kenya collected over a three year period (n=56). Each node represents a different MLVA profile. (A) Node colours denote MLVA profiles and (B) Blue coloured node colour denote non-H58 *Salmonella* Typhi isolates and red coloured nodes denote H58 *Salmonella* Typhi isolates.

5.3.4.6 Madagascar

Four MLVA types were identified amongst the five *Salmonella* Typhi isolates collected from Madagascar. MLVA type STyMT-189 consisted of two isolates. The H58 clone was not limited to a single MLVA type. MLVA type STyMT-189 consisted of isolates belonging to both the H58 clone as well as non-H58 haplotype.

5.3.4.7 Senegal

Diversity was observed amongst two *Salmonella* Typhi isolates collected in Senegal. These MLVA types differ at four VNTR loci. Within this group of isolates, the H58 clone was associated with MLVA type STyMT-210.

5.3.4.8 Tanzania

MLVA typing showed that *Salmonella* Typhi isolates from Tanzania were highly clonal. There were two MLVA types identified from the seven isolates subjected to MLVA typing. These two MLVA types differ at two VNTR loci. MLVA type STyMT-101 was found to be predominant in Tanzania. Isolates belonging to the H58 clonal group were observed in both MLVA types. All isolates from Tanzania belonged to this clonal group

5.3.5 Intracontinental transmission of *Salmonella* Typhi isolates

MLVA typing revealed evidence of intracontinental transmission of typhoid fever (Figure 5.11). Similar MLVA types were identified in different SSA countries. MLVA type STyMT-101 was identified in Senegal, Tanzania and Madagascar. The majority of isolates from Tanzania belong to this MLVA type. This may suggests that isolates are transmitted between Tanzania, Senegal and Madagascar. MLVA type STyMT-132 was identified in both Ghana and South Africa and MLVA type STyMT-48 in Ghana, South Africa and Burkina Faso. Isolates belonging to MLVA type STyMT-107 were identified in Kenya and in South Africa while MLVA types STyMT-49, STyMT-83 and STyMT-195 were common in Kenya and Ghana.

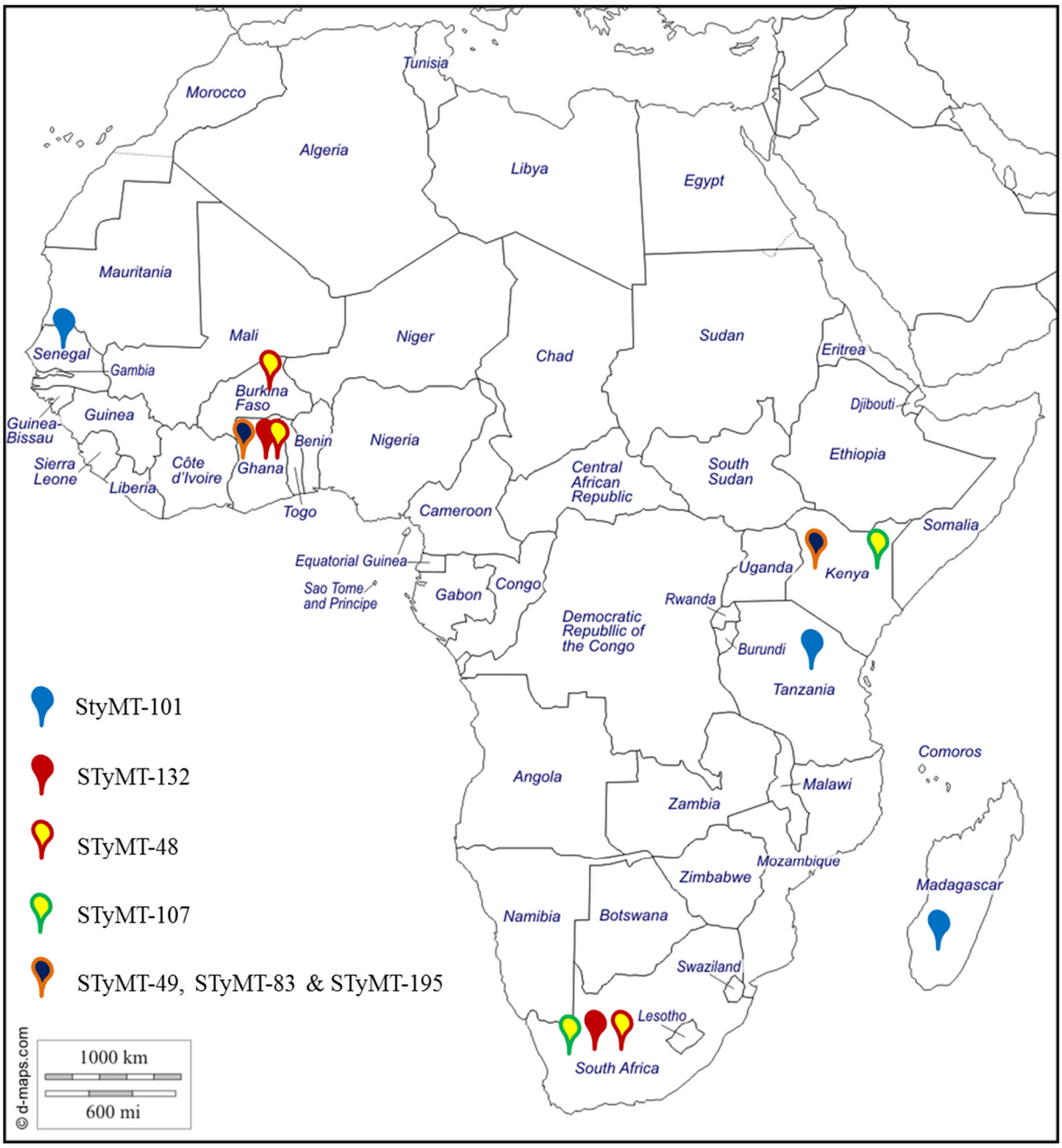


Figure 5.11 Outline map of Africa showing intracontinental distribution and transmission of *Salmonella* Typhi isolates based on MLVA assay.

5.3.6 Comparison of PFGE analysis of *Salmonella* Typhi isolates from SSA to MLVA typing

PFGE analysis of *Salmonella* Typhi isolates from SSA was performed using *Xba*I in order to study the diversity and strain relatedness of these isolates. Digestion of the genomic DNA using *Xba*I revealed 143 unique pulsotypes from which 44 clusters were identified. PFGE clusters were defined as a group of two or more isolates with indistinguishable pulsotypes. PFGE analysis revealed five major clusters consisting of nine or more *Salmonella* Typhi isolates.

5.3.6.1 PFGE Cluster I

PFGE Cluster I consisted of ten *Salmonella* Typhi isolates from South Africa (Figure 5.12). The PFGE fingerprint pattern of these isolated was designated as STyPT-10. Even though isolates within this cluster had the same pulsotype, six MLVA types were identified within this group. These included STyMT-105 (n=1), STyMT-121 (n=4), STyMT-100 (n=2), STyMT-125 (n=1), STyMT-63 (n=1), STyMT-69 (n=1). Isolates within this cluster differed at up to four VNTR loci.

5.3.6.2 PFGE Cluster II

PFGE analysis revealed high homogeneity amongst *Salmonella* Typhi isolates from Kenya (Figure 5.12 and Figure 5.13). Pulsotype STyPT14 consisted only of Kenyan *Salmonella* Typhi isolates. This was found to be the predominant PFGE fingerprint pattern in Kenya consisting of 36 of the 56 isolates (64%). Within this cluster, 20 MLVA types were identified. The most common MLVA type within PFGE cluster II was STyMT-86 (n=8) followed by STyMT-256 (n=3) and STyMT-84 (n=3). Interestingly, some isolates within this group differed at all five VNTR loci; even so, these were found to be indistinguishable by PFGE analysis.

5.3.6.3 PFGE Cluster III

PFGE analysis suggested that isolates from South Africa are highly clonal, and that an association exists between *Salmonella* Typhi isolates from South Africa, Tanzania and Madagascar (Figure 5.13 and Figure 5.14). Of the 42 isolates that belonged to PFGE Cluster III, four were from Tanzania, one from Madagascar and 37 from South Africa. This was

identified as the largest cluster and dendrogram analysis indicated that STyPT-22 was the predominant PFGE fingerprint type in South Africa. This pulsotypes consisted of isolates belonging to 30 different MLVA Types. The most common MLVA types identified within this cluster were STyMT-101 (n=4) and STyMT-132 (n=3). This cluster consisted of isolates that differed at one to all five VNTR loci.

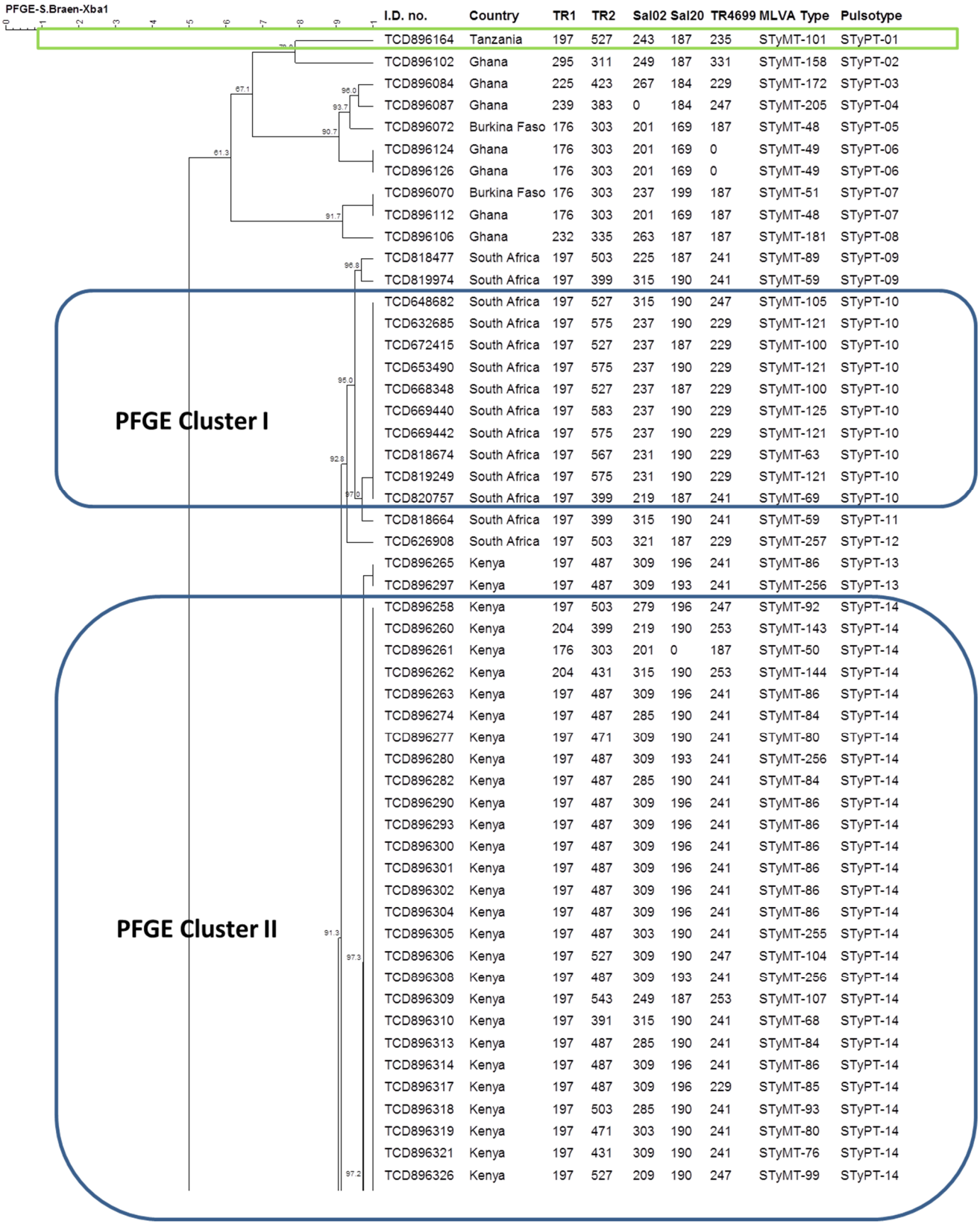
5.3.6.4 PFGE Cluster IV

PFGE Cluster IV consisted of 12 *Salmonella* Typhi isolates from South Africa (Figure 5.14). Within this group, 12 unique MLVA types were identified. Most isolates within this cluster differed at two VNTR loci; however; isolates that differed at all five VNTR loci were identified. Dendrogram analysis suggests that pulsotypes STyPT-25 is the second most common PFGE fingerprint type in South Africa.

5.3.6.5 PFGE Cluster V

PFGE Cluster V consisted of isolates from Ghana (n=2), South Africa (n=2) and Burkina Faso (n=5) (Figure 5.15). A total of 7 unique MLVA types were identified in this cluster. The most common MLVA Type within this group of isolates was STyMT-206, identified in most of the isolates from Burkina Faso (n=3). Isolates within this cluster differed at one to all five VNTR loci.

Dendrogram analysis revealed a group of isolates clustered together by MLVA typing but found to be genetically diverse by PFGE (Figure 5.12, Figure 5.13, Figure 5.14 and Figure 5.16). A total of seven isolates belonging to MLVA type STyMT-101 were dispersed throughout out the dendrogram and were found to be unrelated by MLVA.



I.D. no.	Country	TR1	TR2	Sal02	Sal20	TR4699	MLVA Type	Pulsotype
TCD896164	Tanzania	197	527	243	187	235	STyMT-101	STyPT-01
TCD896102	Ghana	295	311	249	187	331	STyMT-158	STyPT-02
TCD896084	Ghana	225	423	267	184	229	STyMT-172	STyPT-03
TCD896087	Ghana	239	383	0	184	247	STyMT-205	STyPT-04
TCD896072	Burkina Faso	176	303	201	169	187	STyMT-48	STyPT-05
TCD896124	Ghana	176	303	201	169	0	STyMT-49	STyPT-06
TCD896126	Ghana	176	303	201	169	0	STyMT-49	STyPT-06
TCD896070	Burkina Faso	176	303	237	199	187	STyMT-51	STyPT-07
TCD896112	Ghana	176	303	201	169	187	STyMT-48	STyPT-07
TCD896106	Ghana	232	335	263	187	187	STyMT-181	STyPT-08
TCD818477	South Africa	197	503	225	187	241	STyMT-89	STyPT-09
TCD819974	South Africa	197	399	315	190	241	STyMT-59	STyPT-09
TCD648682	South Africa	197	527	315	190	247	STyMT-105	STyPT-10
TCD632685	South Africa	197	575	237	190	229	STyMT-121	STyPT-10
TCD672415	South Africa	197	527	237	187	229	STyMT-100	STyPT-10
TCD653490	South Africa	197	575	237	190	229	STyMT-121	STyPT-10
TCD668348	South Africa	197	527	237	187	229	STyMT-100	STyPT-10
TCD669440	South Africa	197	583	237	190	229	STyMT-125	STyPT-10
TCD669442	South Africa	197	575	237	190	229	STyMT-121	STyPT-10
TCD818674	South Africa	197	567	231	190	229	STyMT-63	STyPT-10
TCD819249	South Africa	197	575	231	190	229	STyMT-121	STyPT-10
TCD820757	South Africa	197	399	219	187	241	STyMT-69	STyPT-10
TCD818664	South Africa	197	399	315	190	241	STyMT-59	STyPT-11
TCD626908	South Africa	197	503	321	187	229	STyMT-257	STyPT-12
TCD896265	Kenya	197	487	309	196	241	STyMT-86	STyPT-13
TCD896297	Kenya	197	487	309	193	241	STyMT-256	STyPT-13
TCD896258	Kenya	197	503	279	196	247	STyMT-92	STyPT-14
TCD896260	Kenya	204	399	219	190	253	STyMT-143	STyPT-14
TCD896261	Kenya	176	303	201	0	187	STyMT-50	STyPT-14
TCD896262	Kenya	204	431	315	190	253	STyMT-144	STyPT-14
TCD896263	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896274	Kenya	197	487	285	190	241	STyMT-84	STyPT-14
TCD896277	Kenya	197	471	309	190	241	STyMT-80	STyPT-14
TCD896280	Kenya	197	487	309	193	241	STyMT-256	STyPT-14
TCD896282	Kenya	197	487	285	190	241	STyMT-84	STyPT-14
TCD896290	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896293	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896300	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896301	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896302	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896304	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896305	Kenya	197	487	303	190	241	STyMT-255	STyPT-14
TCD896306	Kenya	197	527	309	190	247	STyMT-104	STyPT-14
TCD896308	Kenya	197	487	309	193	241	STyMT-256	STyPT-14
TCD896309	Kenya	197	543	249	187	253	STyMT-107	STyPT-14
TCD896310	Kenya	197	391	315	190	241	STyMT-68	STyPT-14
TCD896313	Kenya	197	487	285	190	241	STyMT-84	STyPT-14
TCD896314	Kenya	197	487	309	196	241	STyMT-86	STyPT-14
TCD896317	Kenya	197	487	309	196	229	STyMT-85	STyPT-14
TCD896318	Kenya	197	503	285	190	241	STyMT-93	STyPT-14
TCD896319	Kenya	197	471	303	190	241	STyMT-80	STyPT-14
TCD896321	Kenya	197	431	309	190	241	STyMT-76	STyPT-14
TCD896326	Kenya	197	527	209	190	247	STyMT-99	STyPT-14

Figure 5.12 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters I and II. The green border line indicates *Salmonella* Typhi isolates that were indistinguishable by MLVA assay but diverse by PFGE analysis

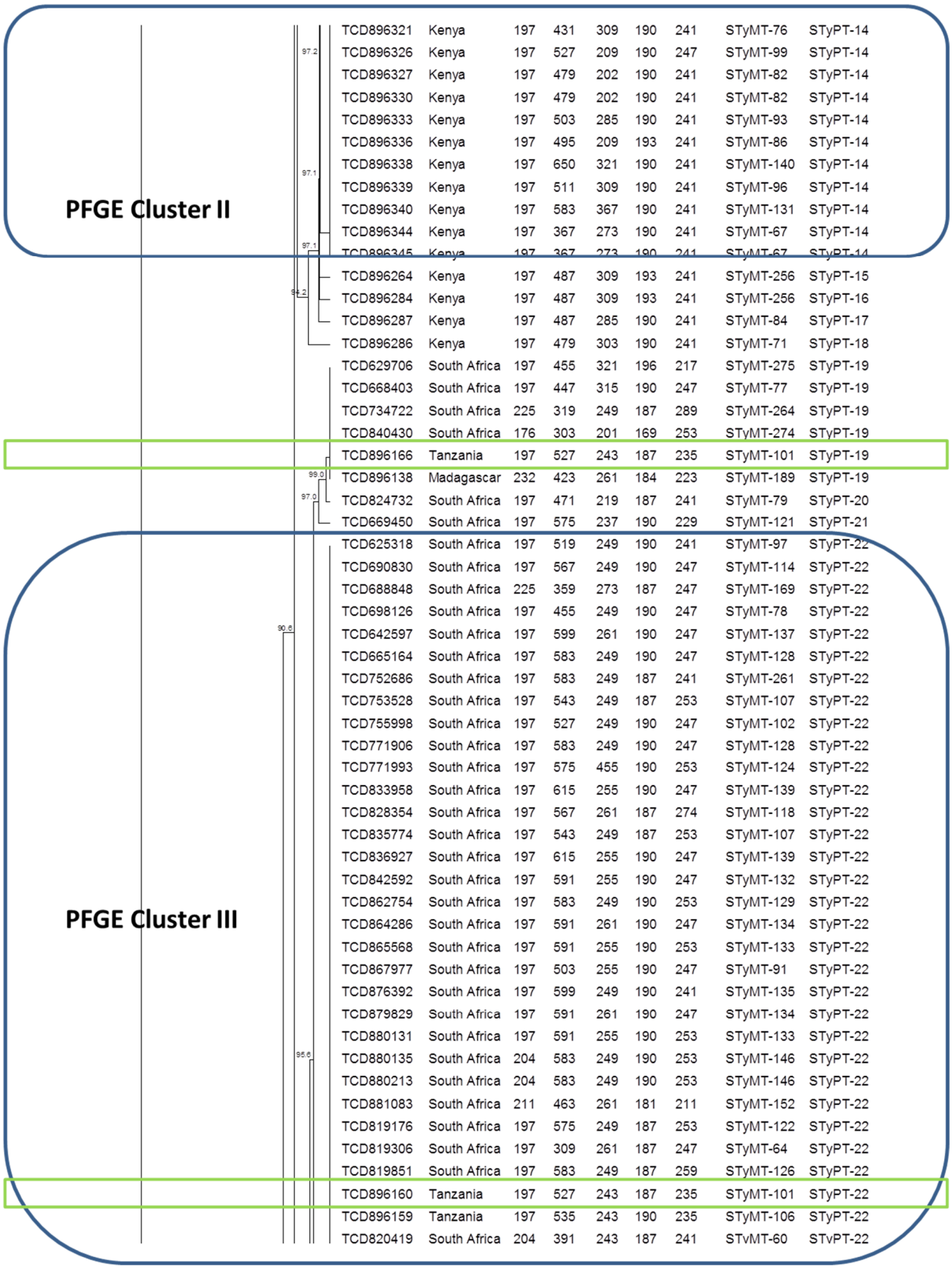


Figure 5.13 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters II and III. The green border line indicates *Salmonella* Typhi isolates that were indistinguishable by MLVA assay but diverse by PFGE analysis

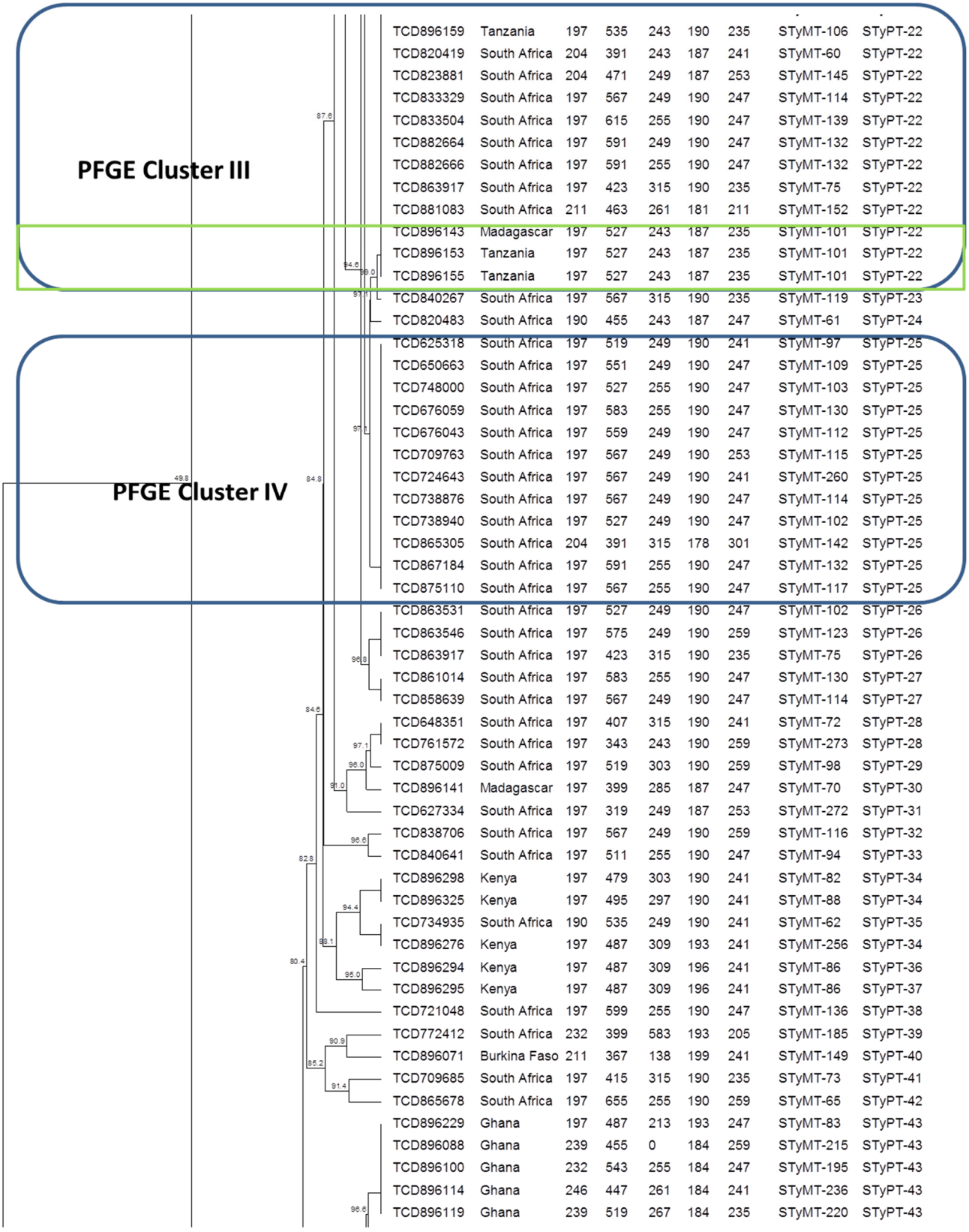


Figure 5.14 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE clusters III and IV. The green border line indicates *Salmonella* Typhi isolates that were indistinguishable by MLVA assay but diverse by PFGE analysis

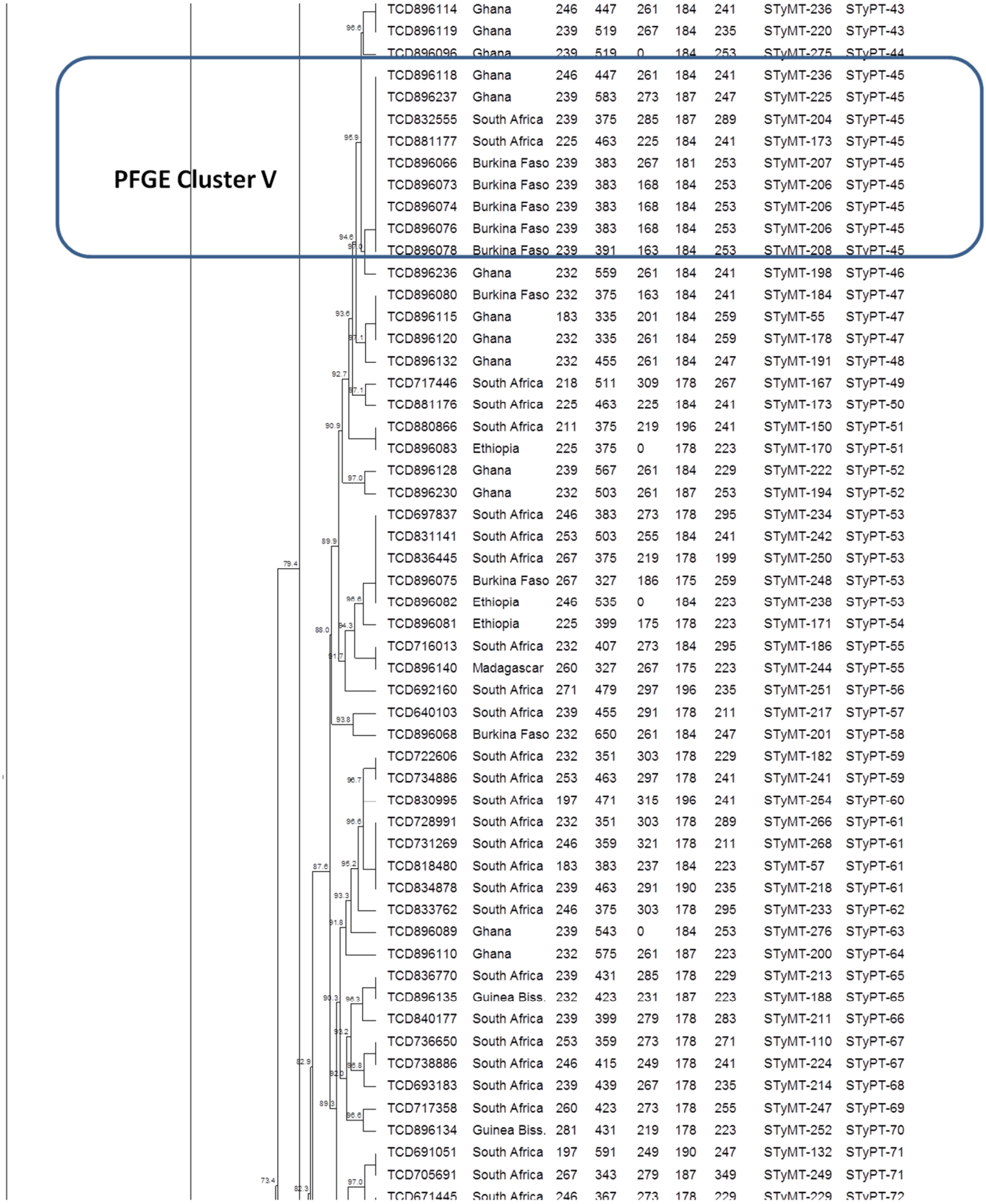


Figure 5.15 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA showing PFGE cluster V.

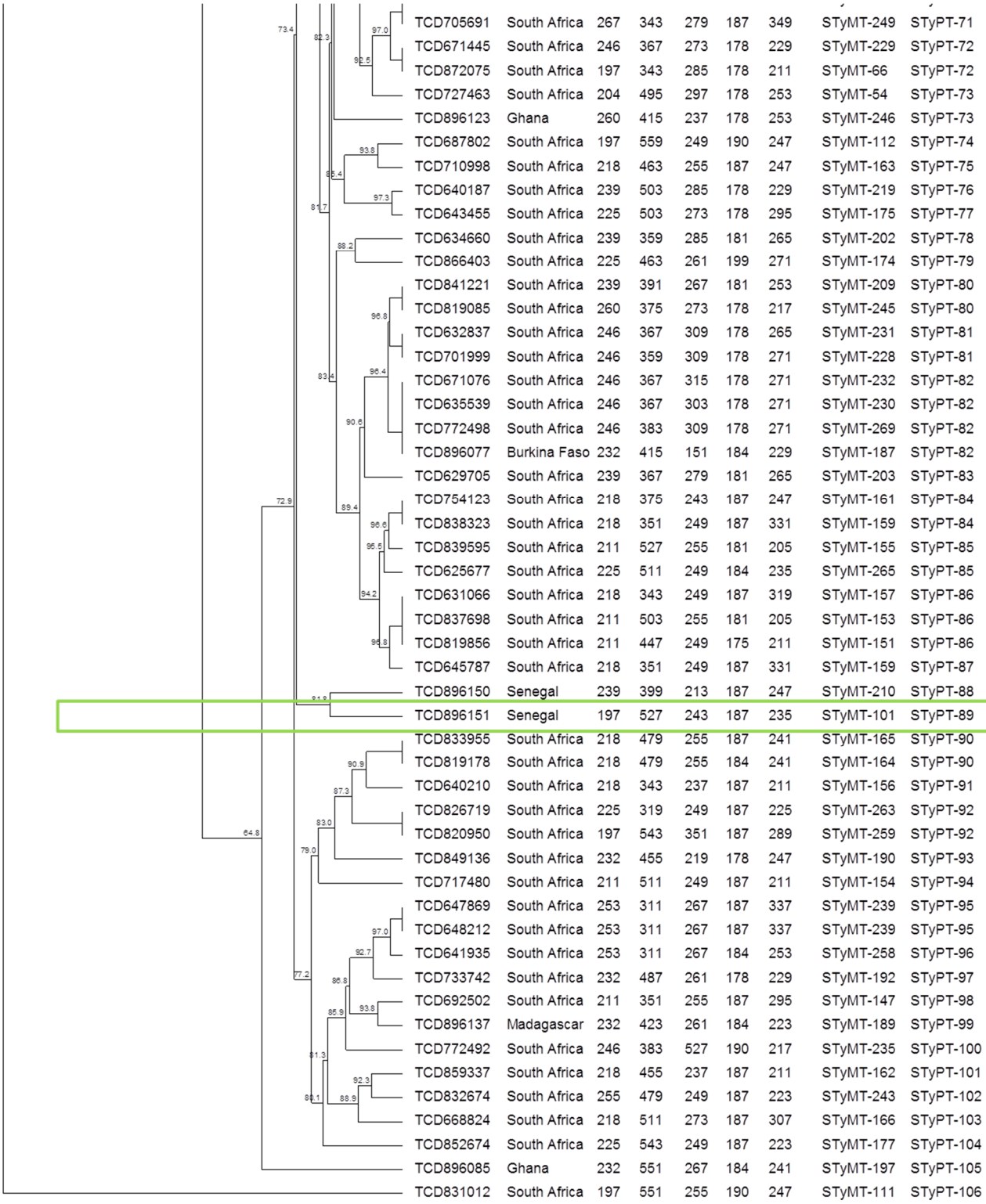


Figure 5.16 Dendrogram analysis of PFGE patterns of *Salmonella* Typhi isolates from SSA. The green border line indicates *Salmonella* Typhi isolates that were indistinguishable by MLVA assay but diverse by PFGE analysis

5.3.7 Comparison of the discriminatory power of MLVA assay with PFGE analysis

In order to estimate the discriminatory power of the two subtyping techniques, Simpson's DI was applied to the *Salmonella* Typhi isolates from SSA. A total of 226 MLVA types were identified from the 316 *Salmonella* Typhi isolates that were analysed. For the MLVA assay, Simpson's DI was calculated at 0.995 (95% CI 0.992 – 0.997). In contrast, Simpson's DI for PFGE analysis of the same isolates was calculated at 0.959 (95% CI 0.943 – 0.972). The difference in Simpson's DI for MLVA and PFGE were statistically significant as non-overlapping CIs were observed (p-value <0.001). As expected, several MLVA profiles were observed for isolates that shared the same PFGE profile; however, the reverse was observed in some cases where several pulsotypes were observed for isolates of the same MLVA type (Figures 5.6, 5.7, 5.8 and 5.10)

The congruence between the MLVA assay and PFGE analysis was assessed by calculating the Wallace coefficient. Through calculating the Wallace coefficient, the probability that 2 isolates that cluster together by one subtyping method could also be clustered together using another subtyping method was determined. The Wallace coefficient between MLVA assay and PFGE pulsotypes was determined to be 52%; however, the probability that two isolates that belong to the same MLVA type could have the same pulsotypes was determined to be 6 %. The p-value between the Wallace indices was determined to be <0.001. This indicates that the MLVA assay exhibited more discriminatory power than PFGE analysis.

Table 5.1 Simpson's Index of Diversity

Method	**Partitions**	**Simpson's ID**	**CI (95%)**	P-value
PFGE	143	0.959	(0.947 – 0.972)	<0.001
MLVA	226	0.995	(0.992 – 0.997)	

Table 5.2 Wallace coefficient and analytical 95% CI

	MLVA assay	**PFGE analysis**	**P-value**
MLVA assay		0.516 (0.353 – 0.678)	<0.001
PFGE analysis	0.062 (0.025 – 0.099)		

5.4 Discussion

In this study, MLVA assay was applied to a total of 316 *Salmonella* Typhi isolates from SSA. The MST was generated in order to study the diversity and show relations between the isolates. MLVA showed high discrimination of the isolates as 226 MLVA types were observed from the 316 isolates that were processed. The SSA community of *Salmonella* Typhi isolates were found to be diverse by MLVA analysis as no major clusters were observed. These results showed that MLVA was able to provide good discrimination even for the highly clonal *Salmonella* Typhi.

MLVA assay was able to discriminate isolates from South Africa as 155 MLVA types were identified from 195 isolates. Although no major clusters were observed, six clusters consisting of three to four isolates belonging to the same MLVA type were identified. There was evidence of inter-provincial transmission of *Salmonella* Typhi isolates as several MLVA types were observed in more than one province. Even though distinct MLVA types were observed during each year of the study, MST showed three MLVA types that were common throughout the study period. Additionally, MST revealed common MLVA types that were observed between the years 2012 and 2013, between 2012 and 2014 and between 2013 and 2014.

MLVA typing of *Salmonella* Typhi isolates from seven other SSA countries revealed heterogeneity amongst the isolates. There were no apparent major clusters observed in Burkina Faso, Ethiopia, Ghana, Guinea Bissau, Madagascar and Senegal. MLVA analysis revealed diversity amongst *Salmonella* Typhi isolates from these countries and only a few clusters consisting of two to three isolates identified.

Although high diversity was observed amongst Kenyan *Salmonella* Typhi isolates, isolates belonging to MLVA type STyMT-86 were the predominant. Approximately 30% of all isolates from Kenya belonged to STyMT-86. This could indicate that *Salmonella* Typhi STyMT-86 is a dominant clonal type responsible for most typhoid fever infections in Kenya.

MLVA indicated that *Salmonella* Typhi isolates from Tanzania were highly clonal. This was contrary to what was observed in other SSA countries where *Salmonella* Typhi isolates were found to be heterogeneous. Only two MLVA types were identified amongst isolates from Tanzania with the majority of isolates (85%, n=6) belonging to MLVA type STyMT-101.

MLVA typing revealed intra-continental transmission of isolates as common MLVA types were identified between SSA countries. Common MLVA types were identified between Senegal, Tanzania and Madagascar, between Kenya and South Africa, between South Africa and Ghana, between Ghana and Kenya; and finally between Ghana, Burkina Faso and South Africa. This correlates with observation by Wong and colleagues (2015) that there is an on-going epidemic of typhoid fever in Africa [132]. Interestingly, MLVA assay did not reveal evidence of transmission of isolates between Kenya and its neighbouring country, Tanzania.

The MST showed that H58 *Salmonella* Typhi isolates were dispersed and were not limited to any clonal complex. Even so, the majority of these isolates were found in clonal complex I. Although preliminary data suggests that correlation exists between MLVA type and the H58 haplotype, analysis of a greater number of isolates is required in order to qualify this observation. Nonetheless, MLVA typing showed that the H58 haplotype is not limited to one particular MLVA type, as MLVA types representing both H58 *Salmonella* Typhi and non-H58 *Salmonella* Typhi isolates were identified.

PFGE analysis of the 316 isolates from SSA revealed 143 pulsotypes. PFGE analysis revealed five major clusters consisting of nine isolates or more. Clustering based on PFGE revealed grouping that was dissimilar to that of MLVA. PFGE cluster I was exclusive to ten *Salmonella* Typhi isolates from South Africa. Although PFGE analysis found these isolates to be indistinguishable, six MLVA types were identified within this group. Similarly, the 12 South African *Salmonella* Typhi isolates in PFGE cluster IV were found to be highly diverse by MLVA typing as 12 MLVA types were identified within this group of isolates. PFGE cluster II consisted of 64% of the isolates from Kenya. This suggested that the Kenyan *Salmonella* Typhi isolates were highly clonal; however, MLVA typing revealed high diversity amongst the isolates as 20 MLVA types were identified within this group.

PFGE clusters III and V exhibited evidence of intracontinental transmission of *Salmonella* Typhi isolates. PFGE cluster III consisted of isolates from South Africa, Tanzania and Madagascar. This suggested transmission of isolates from Tanzania to both South African and Madagascar. On the contrary, MLVA typing did not reveal similarities between isolates from Tanzania and South Africa or Madagascar and South Africa; however PFGE analysis correlates with MLVA typing in indicating that transmission of isolates between Tanzania and Madagascar exist. Although PFGE analysis found the 42 isolates within this cluster to be indistinguishable, MLVA typing indicated high diversity as 30 MLVA types were identified. PFGE cluster V consisted of isolates from South Africa, Burkina Faso and Ghana. Although MLVA typing found isolates within this group to be highly diverse, MLVA typing correlates with PFGE in showing that there is transmission of isolates between South Africa, Burkina Faso and Ghana.

In order to qualify the degree of similarity between MLVA and PFGE in clustering of isolates, both the Simpson DI and congruence between the typing methods were calculated. The Simpson DI is the most commonly used mathematical measure to estimate the discriminatory ability of subtyping methods [138]. This measures the probability that two epidemiologically unrelated isolates will be characterized as diverse by the typing method evaluated. The Simpson's DI is subjective towards the relative abundance of each subtype in the population studied. The Simpson's DI indicated that MLVA exhibited higher discriminatory power than PFGE.

The Wallace coefficient was calculated in order to determine the probability that isolates clustered together by MLVA typing will also cluster together by PFGE analysis. The directional congruence as indicated by Wallace coefficient from MLVA to PFGE was 53% suggesting that there was a high probability that isolates assigned to a cluster by MLVA typing would be assigned to the same cluster by PFGE. On the contrary, calculation of the Wallace coefficient indicated that isolates assigned to the same cluster by PFGE analysis a low probability to be assigned to the same cluster by MLVA typing.

The discriminatory capacity of MLVA and PFGE analysis of *Salmonella* Typhi was assessed in other studies [195, 204]. Tien and colleagues (2012) compared the discriminatory capacities of MLVA and PFGE by subtyping 125 *Salmonella* Typhi isolates from Taiwan collected during the years 1996 to 2009 [195]. MLVA analysis distinguished the isolates into 109 MLVA types while PFGE analysis of the same revealed 71 pulsotypes. MLVA assay was found to be more discriminatory, exhibiting Simpson's DI of 0.996, as compared to PFGE 0.980, showing Simpson's DI [195]. In a recent study by Wang and colleagues (2016), an 8-loci MLVA assay was used to discriminate against 103 *Salmonella* Typhi isolates from China [204]. In their study, the discriminatory capacity was evaluated against that of PFGE using XbaI restriction enzyme. MLVA assay characterized the 103 isolates into 93 MLVA types while PFGE characterized the same isolates into 88 pulsotypes. MLVA exhibited high Simpson's DI values of 0.9981 as compared to Simpson's DI value of 0.9968 exhibited by PFGE analysis [204]

5.5 Conclusion

In this study, MLVA typing showed high levels of diversity amongst *Salmonella* Typhi isolates from SSA. MLVA typing clearly discerned closely related and epidemiologically unrelated isolates. This molecular epidemiological method showed higher discriminatory power that PFGE analysis. Clear differences between the distributions of isolates by MLVA typing and PFGE analysis were observed. While PFGE analysis showed that *Salmonella* Typhi isolates from SSA were highly clonal while MLVA revealed that heterogeneity existed amongst these isolates. PFGE was found to be suitable for establishing the major phylogenetic relationships between isolates, while MLVA typing was found to be more appropriate in determining the relationships of very closely related isolates. Rapid identification and characterization of typhoid fever outbreaks using MLVA analysis can improve effective initiation of public health intervention measures.

For epidemiological purposes, the following MLVA interpretations could be implemented: Isolates with identical MLVA profiles are considered to be highly related; isolates exhibiting MLVA profiles that differ at 1 locus are considered to be related; and isolates exhibiting MLVA profiles that differ at 2 or more loci are considered to be unrelated. We suggest using MLVA in conjunction with PFGE in typhoid fever surveillance and outbreak investigations for the accumulation of data regarding phylogenetic relationships and for further characterization of epidemiologically related isolates.

Chapter 6

General Discussions and Conclusions

This study focused on the use of MLVA subtyping tool for the molecular characterization of *Salmonella* Typhi isolates, particularly in the SSA context. *Salmonella* Typhi, the aetiological agent of typhoid fever, is endemic in this region and remains a public health problem. Very little data exists about the molecular epidemiology of *Salmonella* Typhi in SSA. The lack of genetic diversity amongst *Salmonella* Typhi isolates has presented a significant challenge in developing suitable molecular subtyping methodology for characterizing this pathogen.

In this study, we explored the use of VNTRs as molecular markers for the differentiation of *Salmonella* Typhi isolates by developing a MLVA assay consisting of five VNTR loci. We also used a simple conventional PCR to study the prevalence of a highly resistant *Salmonella* Typhi strain in the SSA region. Furthermore, the developed MLVA assay was used to discriminate *Salmonella* Typhi isolates from SSA and the discriminatory capacity of MLVA was compared to that of PFGE analysis of the same isolates.

6.1 Development of a five loci MLVA assay

PFGE has been commonly used for molecular subtyping of bacterial pathogens; however, it has presented with a series of limitations including labour intensity as well as suboptimal discrimination of pathogens, especially with the highly homogenous *Salmonella* Typhi isolates. Newer technologies, such as WGS applications provide remarkable discrimination of *Salmonella* Typhi isolates, however, using WGS applications as a routine methodology in the SSA region is still many years away; the high costs of WGS including expensive equipment and expensive running costs, and the requirements for specialized laboratory personnel with technical knowhow, makes this an unrealistic option at present. Conversely MLVA has attracted intense interest due to its high discriminatory capabilities combined with the advantage of being rapid, cost effective and a reliable method that is easy to implement.

A total of 13 potential VNTRs were selected and evaluated using a panel of 50 diverse *Salmonella* Typhi isolates. Of these, six were monomorphic and showed no variation among the isolates. The remaining seven were subjected to nucleotide sequencing and only five were found to have consistent and conserved flanking regions. These five VNTR loci (TR1, TR2, Sal02, Sal20 and TR4699) were included and used in the development of the MLVA assay. This MLVA assay was used to type the panel of 50 *Salmonella* Typhi isolates. MLVA showed great discrimination of the isolates as 47 MLVA and showed high Simpson DI of 0,998 (CI 0,994-1,000).

6.2 The use of conventional PCR to study the prevalence of H58 *Salmonella* Typhi

We explored the use of a low-cost conventional PCR assay that did not require specialized equipment and was suitable for the rapid identification of H58 *Salmonella* Typhi isolates. The H58 *Salmonella* Typhi haplotype has been associated with MDR and resistance to the current typhoid fever treatment regime, ciprofloxacin. This strain is wide-spread in South East Asia and is spreading into SSA. The ability for laboratories in the SSA region to rapidly identify this highly resistant *Salmonella* Typhi strain is crucial for successful treatment of typhoid fever. The conventional PCR targeted a deletion in the STY1507 gene that is present in all H58 haplotype isolates. This PCR was used to screen *Salmonella* Typhi isolates for the prevalence of the H58 haplotype.

Approximately 54% (105 of 195) *Salmonella* Typhi isolates from South Africa were identified as H58 *Salmonella* Typhi. From the other eight countries (Burkina Faso, Ethiopia, Ghana, Guinea Bissau, Kenya, Madagascar, Senegal, South Africa and Tanzania) in the SSA region, approximately 62% (75 of 121) *Salmonella* Typhi isolates were identified as H58 *Salmonella* Typhi. Results from the PCR assay indicate that the H58 *Salmonella* Typhi is now widespread in SSA and seems to be replacing the antimicrobial susceptible *Salmonella* Typhi.

6.3 MLVA offers higher discriminatory power than PFGE analysis for molecular subtyping *Salmonella* Typhi isolates from SSA

The MLVA assay developed in this study was used for the molecular subtyping of *Salmonella* Typhi isolates from nine countries in the SSA region, including Burkina Faso, Ethiopia, Ghana, Guinea Bissau, Kenya, Madagascar, Senegal, South Africa and Tanzania. These isolates were also subjected to PFGE analysis and discriminatory power of MLVA and PFGE compared. From a total of 316 *Salmonella* Typhi isolates from SSA, 226 MLVA types were identified; while PFGE analysis of the same isolates revealed 143 PFGE fingerprint patterns. MLVA was able to resolve all isolates from SSA and indicated intracontinental transmission of *Salmonella* Typhi isolates. Throughout the SSA region, MLVA showed no evidence of geographical association with MLVA type.

For the South African *Salmonella* Typhi isolates, MLVA indicated interprovincial transmission of isolates. Additionally, MLVA showed a number of isolates that were present throughout the study period. MLVA typing also provided insight into the province related epidemiology of isolates in South Africa. There was evidence of interprovincial transmission of typhoid fever. However, no single MLVA type was predominant around the country; suggesting that different *Salmonella* Typhi populations exist and are circulating throughout the country.

MLVA types that were exclusively associated the H58 *Salmonella* Typhi were identified, and MLVA types representing both H58 *Salmonella* Typhi and non-H58 *Salmonella* Typhi isolates were also observed. The H58 *Salmonella* Typhi isolates were diverse and not limited to one clonal complex. The diversity amongst the H58 *Salmonella* Typhi isolates suggest that the emergence of the H58 haplotype could occur as a result of clonal transmission events of the pathogen as well as antimicrobial selective pressure.

There was a significant difference in the discriminatory power of MLVA as compared to PFGE. The Simpson's DI was for the MLVA assay was calculated at 0.995 (95% CI 0.992 – 0.997) as compared to 0.959 (95% CI 0.943 – 0.972) for PFGE analysis. The MLVA assay showed higher discriminatory of *Salmonella* Typhi isolates than PFGE and could be better

suited for would as an effective molecular epidemiological investigation tool for *Salmonella* Typhi outbreaks in SSA

6.4 Concluding remarks

Typhoid fever remains a dreadful disease in the SSA region. In order to obtain important epidemiological information such as transmission routes of typhoid fever outbreaks, the application of molecular subtyping tools is essential. This dissertation represents one of the first studies to explore the use of a MLVA assay for molecular subtyping of the SSA *Salmonella* Typhi population. We have shown that MLVA assay consisting of five highly polymorphic VNTR has provided an excellent discrimination of *Salmonella* Typhi isolates.

The MLVA assay as well as the simple, low-cost and effective protocol used to identify the H58 haplotype in this study could be applied in laboratories across the continent. Laboratories in SSA are resource poor and cannot afford to implement established methods such as PFGE and WGS. MLVA serves as a more affordable method which involves the use of equipment that most laboratories have and can be used for outbreak investigation. To the best of our knowledge, capillary electrophoresis equipment is available in Kenya, Gambia, Ivory Coast and Uganda to which the MLVA assay can be implemented. The adoption of these molecular methodologies by reference laboratories in the SSA region could be helpful in monitoring the spread of typhoid fever across the SSA and to emphasise the role of specific clones as the cause of typhoid fever infection. The MLVA assay could also assist in understanding the role of imported cases of typhoid fever into SSA countries and their contribution to the burden of disease. This could promote effective and appropriate disease control strategies to prevent the emergence and spread of this pathogen.

The provision of clean water coupled with improvements in sanitation in the SSA countries could curb the spread of typhoid fever on the continent. Additionally, given that typhoid carriers play a huge role in the transmission of typhoid fever, it is then evident that a huge gap exists within the surveillance system, as molecular characterization and discrimination of isolates is based on isolates collected only from patients reporting to the hospitals as well as laboratory confirmed cases that are submitted for laboratory-based surveillance. It would be of

great interest and benefit to use MLVA typing to study the association between *Salmonella* Typhi isolates collected from typhoid carriers and those collected from hospital admitted patients. It is important that MLVA typing data should be coupled with epidemiological data in order to make meaningful interpretation that could be used to implement effective control measures.

List of references

1. **Todd EC**. Epidemiology of foodborne diseases: a worldwide review. *World Health Statistics Quarterly*. 1996;50(1-2):30-50.
2. **Foley SL, Lynne AM, Nayak R**. Molecular typing methodologies for microbial source tracking and epidemiological investigations of Gram-negative bacterial foodborne pathogens. *Infection, Genetics & Evolution*. 2009;9(4):430-40.
3. **Newell DG, Koopmans M, Verhoef L, Duizer E, Aidara-Kane A*, et al.*** Food-borne diseases of the challenges of 20years ago still persist while new ones continue to emerge. *International journal of food microbiology*. 2010;139:S3-S15.
4. **Liebana E**. Molecular tools for epidemiological investigations of *Salmonella enterica* subspecies *enterica* infections. *Research in Veterinary Science*. 2002;72(3):169-75.
5. **Bean NH, Griffin PM, Goulding JS, Ivey CB**. Foodborne disease outbreaks, 5-year summary, 1983-1987. *Journal of Food Protection*. 1990;53(8):711-28.
6. **Esrey SA, Feachem RG, Hughes JM**. Interventions for the control of diarrhoeal diseases among young children: improving water supplies and excreta disposal facilities. *Bulletin of the World Health Organization*. 1985;63(4):757.
7. **Swaminathan B, Barrett TJ, Hunter SB, Tauxe RV, Force CPT**. PulseNet: the molecular subtyping network for foodborne bacterial disease surveillance, United States. *Emerging infectious diseases*. 2001;7(3):382.
8. **Wattiau P, Boland C, Bertrand S**. Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: gold standards and alternatives. *Applied & Environmental Microbiology*. 2011;77(22):7877-85.
9. **Lindstedt BA**. Multiple–locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis*. 2005;26(13):2567-82.
10. **Torpdahl M, Skov MN, Sandvang D, Baggesen DL**. Genotypic characterization of *Salmonella* by multilocus sequence typing, pulsed-field gel electrophoresis and amplified fragment length polymorphism. *Journal of Microbiological Methods*. 2005;63(2):173-84.
11. **Liu Y, Lee MA, Ooi EE, Mavis Y, Tan AL*, et al.*** Molecular typing of *Salmonella enterica* serovar Typhi isolates from various countries in Asia by a multiplex PCR assay on variable-number tandem repeats. *Journal of Clinical Microbiology*. 2003;41(9):4388-94.
12. **Crump JA, Mintz ED**. Global trends in typhoid and paratyphoid fever. *Clinical Infectious Diseases*. 2010;50(2):241-6.
13. **Buckle GC, Walker CLF, Black RE**. Typhoid fever and paratyphoid fever: Systematic review to estimate global morbidity and mortality for 2010. *Journal of Global Health*. 2012;2(1).
14. **Foxman B, Riley L**. Molecular epidemiology: focus on infection. *American Journal of Epidemiology*. 2001;153(12):1135-41.
15. **Ochman H, Jones IB**. Evolutionary dynamics of full genome content in Escherichia coli. *The EMBO journal*. 2000;19(24):6637-43.
16. **Molbak K, Olsen JE, Wegener HC**. *Salmonella* infections. *Foodborne Infections and Intoxications*. 2006:57-136.
17. **Achtman M, Wain J, Weill FX, Nair S, Zhou Z*, et al.*** Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathogens*. 2012;8(6):e1002776.

18. **Rowe B, Hall MLM**. Kauffman-White scheme. *Public Health Laboratory Service, London, UK*. 1989.
19. **Kidgell C, Reichard U, Wain J, Linz B, Torpdahl M*, et al.*** *Salmonella* Typhi, the causative agent of typhoid fever, is approximately 50,000 years old. *Infection, Genetics & Evolution*. 2002;2(1):39-45.
20. **Crump JA, Luby SP, Mintz ED**. The global burden of typhoid fever. *Bulletin of the World Health Organization*. 2004;82(5):346-53.
21. **Raffatellu M, Wilson RP, Winter SE, Baumler AJ**. Clinical pathogenesis of typhoid fever. *The Journal of Infection in Developing Countries*. 2008;2(04):260-6.
22. **Kariuki S**. Typhoid fever in sub-Saharan Africa: challenges of diagnosis and management of infections. *The Journal of Infection in Developing Countries*. 2008;2(06):443-7.
23. **Feasey NA, Archer BN, Heyderman RS, Sooka A, Dennis B, *et al.*** Typhoid fever and invasive nontyphoid salmonellosis, Malawi and South Africa. *Emerging Infectious Diseases*. 2010;16(9):1448.
24. **Kanungo S, Dutta S, Sur D**. Epidemiology of typhoid and paratyphoid fever in India. *The Journal of Infection in Developing Countries*. 2008;2(06):454-60.
25. **Mweu E, English M**. Typhoid fever in children in Africa. *Tropical Medicine and International Health*. 2008;13(4):532-40.
26. **Beyene G, Asrat D, Mengistu Y, Aseffa A, Wain J**. Typhoid fever in Ethiopia. *The Journal of Infection in Developing Countries*. 2008;2(06):448-53.
27. Mogasale V, Maskery B, Ochiai RL, Lee JS, Mogasale VV, *et al.* Burden of typhoid fever in low-income and middle-income countries: a systematic, literature-based update with risk-factor adjustment. The Lancet Global Health: Elsevier. p. e570-e80.
28. **Lynch MF, Blanton EM, Bulens S, Polyak C, Vojdani J, *et al.*** Typhoid fever in the United States, 1999-2006. *Jama*. 2009;302(8):859-65.
29. **Kariuki S, Revathi G, Kiiru J, Mengo DM, Mwituria J, *et al.*** Typhoid in Kenya is associated with a dominant multidrug-resistant *Salmonella enterica* serovar Typhi haplotype that is also widespread in Southeast Asia. *Journal of Clinical Microbiology*. 2010;48(6):2171-6.
30. **Dougan G, Baker S**. *Salmonella enterica* serovar Typhi and the pathogenesis of typhoid fever. *Annual Review of Microbiology*. 2014;68:317-36.
31. **Wain J, Hendriksen RS, Mikoleit ML, Keddy KH, Ochiai RL**. Typhoid fever. *The Lancet*. 2014(0).
32. Parry CM, Hien TT, Dougan G, White NJ, Farrar JJ. Typhoid Fever. New England Journal of Medicine; 2002: Massachusetts Medical Society; 2002. p. 1770-82.
33. **Jones BD, Falkow S**. SALMONELLOSIS: Host Immune Responses and Bacterial Virulence Determinants1. *Annual Review of Immunology*. 1996;14(1):533-61.
34. **Gonzalez-Escobedo G, Marshall JM, Gunn JS**. Chronic and acute infection of the gall bladder by *Salmonella* Typhi: understanding the carrier state. *Nature Reviews Microbiology*. 2011;9(1):9-14.
35. **Parry C, Wain J, Chinh NT, Vinh H, Farrar JJ**. Quinolone-resistant *Salmonella* Typhi in Vietnam. *The Lancet*. 1998;351(9111):1289.
36. **Crawford RW, Rosales-Reyes R, Ramírez-Aguilar MdlL, Chapa-Azuela O, Alpuche-Aranda C, *et al.*** Gallstones play a significant role in *Salmonella* spp. gallbladder colonization and carriage. *Proceedings of the National Academy of Sciences*. 2010;107(9):4353-8.
37. **Caygill CPJ, Braddick M, Hill MJ, Knowles RL, Sharp JCM**. The association between typhoid carriage, typhoid infection and subsequent cancer at a number of sites. *European Journal of Cancer Prevention*. 1995;4(2):187-94.

38. **Caygill CPJ, Hill MJ, Braddick M, Sharp JCM**. Cancer mortality in chronic typhoid and paratyphoid carriers. *The Lancet*. 1994;343(8889):83-4.
39. **Nath G, Singh YK, Maurya P, Gulati AK, Srivastava RC*, et al.*** Does *Salmonella* Typhi primarily reside in the liver of chronic typhoid carriers. *J Infect Dev Ctries*. 2010;4(4):259-61.
40. **Nath G, Mauryal P, Gulati AK, Singh TB, Srivastava R*, et al.*** Comparison of Vi serology and nested PCR in diagnosis of chronic typhoid carriers in two different study populations in typhoid endemic area of India. *The Southeast Asian Journal of Tropical Medicine and Public Health*. 2010;41(3):636-40.
41. **Yee Chan K, Reidpath DD**. "Typhoid Mary" and "HIV Jane": Responsibility, Agency and Disease Prevention. *Reproductive Health Matters*. 2003;11(22):40-50.
42. **Lin FY, Ho VA, Khiem HB, Trach DD, Bay PV*, et al.*** The efficacy of a *Salmonella* Typhi Vi conjugate vaccine in two-to-five-year-old children. *New England Journal of Medicine*. 2001;344(17):1263-9.
43. **Slayton R, Date K, Mintz E**. Vaccination for typhoid fever in sub-Saharan Africa. *Human Vaccines and Immunotherapeutics*. 2013;9(4):903-6.
44. **Grabenstein JD, Pittman PR, Greenwood JT, Engler RJ**. Immunization to protect the US Armed Forces: heritage, current practice, and prospects. *Epidemiologic Reviews*. 2006;28(1):3-26.
45. **Engels EA, Falagas ME, Lau J, Bennish ML**. Typhoid fever vaccines: a meta-analysis of studies on efficacy and toxicity. *British Medical Journal*. 1998;316(7125):110-6.
46. **Guzman CA, Borsutzky S, Griot-Wenk M, Metcalfe IC, Pearman J*, et al.*** Vaccines against typhoid fever. *Vaccine*. 2006;24(18):3804-11.
47. **Anwar E, Goldberg E, Fraser A, Acosta CJ, Paul M*, et al.*** Vaccines for preventing typhoid fever. *The Cochrane Library*. 2014.
48. **Klugman KP, Koornhof HJ, Robbins JB, Le Cam NN**. Immunogenicity, efficacy and serological correlate of protection of *Salmonella* Typhi Vi capsular polysaccharide vaccine three years after immunization. *Vaccine*. 1996;14(5):435-8.
49. **Ghenghesh KS, Franka E, Tawil K, Wasfy M, Ahmed SF*, et al.*** Enteric fever in Mediterranean north Africa. *The Journal of Infection in Developing Countries*. 2009;3(10):753-61.
50. **El Nageh MM**. *Salmonella* isolations from human faeces in Tripoli, Libya. *Transactions of the Royal Society of Tropical Medicine & Hygiene*. 1988;82(2):324-6.
51. **Khemiri F, Hassine M, Jekov S, Bel HA, Bel HN**. Evaluation of *Salmonella* 1977 to 1982. *Archives de l'Institut Pasteur de Tunis*. 1982;60(1-2):1-12.
52. **Aissa RB, Al-Gallas N, Troudi H, Belhadj N, Belhadj A**. Trends in *Salmonella enterica* serotypes isolated from human, food, animal, and environment in Tunisia, 1994ΓÇô2004. *Journal of Infection*. 2007;55(4):324-39.
53. **Saida NB, Mhalla S, Bouzouïa N, Boukadida J**. Genotypic analysis of *Salmonella enterica* serovar Typhi collected during two successive autumnal typhoid outbreaks in southeast Tunisia. *Pathologie Biologie*. 2007;55(7):336-9.
54. **Connor BA, Schwartz E**. Typhoid and paratyphoid fever in travellers. *The Lancet Infectious Diseases*. 2005;5(10):623-8.
55. **Guechi Z, Hamza A**. Epidemiology of *Salmonella* infections in Algeria. Evolution of the *Salmonella* serovars isolated from 1986 to 1990. *Archives de l'Institut Pasteur d'Algerie Institut Pasteur d'Algerie*. 1991;58:7-16.
56. **ProMed-mail**. Typhoid fever - Algeria. *ProMed-mail*. 1996.
57. **ProMed-mail**. Typhoid fever - Algeria (Batna). *ProMed- mail*. 2004.

58. **Wasfy MO, Frenck R, Ismail TF, Mansour H, Malone JL*, et al.*** Trends of multiple-drug resistance among *Salmonella* serotype Typhi isolates during a 14-year period in Egypt. *Clinical Infectious Diseases*. 2002;35(10):1265-8.
59. **Srikantiah P, Girgis FY, Luby SP, Jennings G, Wasfy MO*, et al.*** Population-based surveillance of typhoid fever in Egypt. *The American Journal of Tropical Medicine and Hygiene*. 2006;74(1):114-9.
60. **Smith SI, Alao F, Goodluck HT, Fowora M, Bamidele M*, et al.*** Prevalence of *Salmonella* Typhi among food handlers from bukkas in Nigeria. *British Journal of Biomedical Science*. 2008;65(3):158-60.
61. **Mbuh FA, Galadima M, Ogbadu L**. Rate of co-infection with malaria parasites and *Salmonella* Typhi in Zaria, Kaduna State, Nigeria. 2003.
62. **Oboegbulam SI, Oguike JU, Gugnani HC**. Microbiological studies on cases diagnosed as typhoid/enteric fever in south-east Nigeria. *The Journal of Communicable Diseases*. 1995;27(2):97-100.
63. **Adabara NU, Ezugwu BU, Momojimoh A, Madzu A, Hashiimu Z*, et al.*** The prevalence and antibiotic susceptibility pattern of *Salmonella* Typhi among patients attending a military hospital in Minna, Nigeria. *Advances In Preventive Medicine*. 2012;2012.
64. **Akinyemi KO, Coker AO**. Trends of antibiotic resistance in *Salmonella enterica* serovar Typhi isolated from hospitalized patients from 1997 to 2004 in Lagos, Nigeria. *Indian Journal of Medical Microbiology*. 2007;25(4):436-7.
65. **Kalu IG, Ogbulie TE, Opara FN**. Pattern of multi-drug resistant *Salmonella enterica* serovar Typhi isolates in Nigeria. *African Journal of Biotechnology*. 2008;7(21).
66. **Akinyemi KO, Coker AO, Olukoya DK, Oyefolu AO, Amorighoye EP*, et al.*** Prevalence of multi-drug resistant *Salmonella* Typhi among clinically diagnosed typhoid fever patients in Lagos, Nigeria. *Zeitschrift für Naturforschung C*. 2000;55(5-6):489-93.
67. **Akinyemi KO, Smith SI, Oyefolu AB, Coker AO**. Multidrug resistance in *Salmonella enterica* serovar Typhi isolated from patients with typhoid fever complications in Lagos, Nigeria. *Public Health*. 2005;119(4):321-7.
68. **Marks F, Adu-Sarkodie Y, Hünger F, Sarpong N, Ekuban S*, et al.*** Typhoid fever among children, Ghana. *Emerging Infectious Diseases*. 2010;16(11):1796.
69. **Lunguya O, Lejon V, Phoba MF, Bertrand S, Vanhoof R*, et al.*** *Salmonella* Typhi in the democratic republic of the congo: fluoroquinolone decreased susceptibility on the rise. 2012.
70. **Muyembe-Tamfum JJ, Veyi J, Kaswa M, Lunguya O, Verhaegen J*, et al.*** An outbreak of peritonitis caused by multidrug-resistant *Salmonella* Typhi in Kinshasa, Democratic Republic of Congo. *Travel Medicine & Infectious Disease*. 2009;7(1):40-3.
71. **Carmeli Y, Raz R, Schapiro JM, Alkan M**. Typhoid fever in Ethiopian immigrants to Israel and native-born Israelis: a comparative study. *Clinical Infectious Diseases*. 1993;16(2):213-5.
72. **Worku B**. Typhoid fever in an Ethiopian children's hospital: 1984-1995. *Ethiopian Journal of Health Development*. 2000;14(3):311-6.
73. **Lutterloh E, Likaka A, Sejvar J, Manda R, Naiene J*, et al.*** Multidrug-resistant typhoid fever with neurologic findings on the Malawi-Mozambique border. *Clinical Infectious Diseases*. 2012:cis012.
74. **Gordon MA, Graham SM, Walsh AL, Wilson L, Phiri A*, et al.*** Epidemics of invasive *Salmonella enterica* serovar enteritidis and *Salmonella enterica* Serovar

Typhimurium infection associated with multidrug resistance among adults and children in Malawi. *Clinical Infectious Diseases*. 2008;46(7):963-9.

75. **Gordon MA, Walsh AL, Chaponda M, Soko D, Mbvwinji M*, et al.*** Bacteraemia and mortality among adult medical admissions in Malawi--predominance of non-Typhi *Salmonellae* and Streptococcus pneumoniae. *J Infect*. 2001;42(1):44-9.
76. **Mandomando I, Sigaúque B, Morais L, Espasa M, Vallès X*, et al.*** Antimicrobial drug resistance trends of bacteremia isolates in a rural hospital in southern Mozambique. *The American Journal of Tropical Medicine and Hygiene*. 2010;83(1):152-7.
77. **Nkemngu NJ, Asonganyi ED, Njunda AL**. Treatment failure in a typhoid patient infected with nalidixic acid resistant *Salmonella enterica* serovar Typhi with reduced susceptibility to ciprofloxacin: a case report from Cameroon. *BMC Infectious Diseases*. 2005;5(1):49.
78. **Neil KP, Sodha SV, Lukwago L, Shikanga O, Mikoleit M*, et al.*** A large outbreak of typhoid fever associated with a high rate of intestinal perforation in Kasese District, Uganda, 2008 – 2009. *Clinical Infectious Diseases*. 2012:cis025.
79. **Breiman RF, Cosmas L, Njuguna H, Audi A, Olack B*, et al.*** Population-based incidence of typhoid fever in an urban informal settlement and a rural area in Kenya: implications for typhoid vaccine use in Africa. *PloS One*. 2012;7(1):e29119.
80. **Nesbitt A, Mirza NB**. *Salmonella* septicaemias in Kenyan children. *Journal of Tropical Pediatrics*. 1989;35(1):35-9.
81. **Kariuki S, Gilks C, Revathi G, Hart CA**. Genotypic analysis of multidrug-resistant *Salmonella enterica* Serovar Typhi, Kenya. *Emerging Infectious Diseases*. 2000;6(6):649.
82. **Kariuki S, Revathi G, Muyodi J, Mwituria J, Munyalo A*, et al.*** Characterization of multidrug-resistant typhoid outbreaks in Kenya. *Journal of Clinical Microbiology*. 2004;42(4):1477-82.
83. **Mengo DM, Kariuki S, Muigai A, Revathi G**. Trends in *Salmonella* enteric serovar Typhi in Nairobi, Kenya from 2004 to 2006. *The Journal of Infection in Developing Countries*. 2010;4(06):393-6.
84. **Thriemer K, Ley BB, Ame SS, Deen JL, Pak GD*, et al.*** Clinical and epidemiological features of typhoid fever in Pemba, Zanzibar: assessment of the performance of the WHO case definitions. *PloS One*. 2012;7(12):e51823.
85. **Coovadia YM, Gathiram V, Bhamjee A, Garratt RM, Mlisana K*, et al.*** An outbreak of multiresistant *Salmonella* Typhi in South Africa. *An International Journal of Medicine*. 1992;82(2):91-100.
86. **Smith AM, Govender N, Keddy KH**. Quinolone-resistant *Salmonella* Typhi in South Africa, 2003 – 2007. *Epidemiology and Infection*. 2010;138(01):86-90.
87. **Keddy KH, Smith AM, Sooka A, Ismail H, Oliver S**. Fluoroquinolone-resistant typhoid, South Africa. *Emerging Infectious Diseases*. 2010;16(5):879.
88. **Keddy KH, Sooka A, Ismail H, Smith AM, Weber I*, et al.*** Molecular epidemiological investigation of a typhoid fever outbreak in South Africa, 2005: the relationship to a previous epidemic in 1993. *Epidemiology and Infection*. 2011;139(08):1239-45.
89. **Smith AM, Keddy KH, Ismail H, Thomas J, van der Grÿp R*, et al.*** International collaboration tracks typhoid fever cases over two continents from South Africa to Australia. *Journal of Medical Microbiology*. 2011;60(9):1405-7.
90. **Waner S, Kfir R, Idema G, Coetzee D, Rasmussen K*, et al.*** Waterborne outbreak of typhoid fever in Delmas. *SOUTHERN AFRICAN JOURNAL OF EPIDEMIOLOGY AND INFECTION*. 1998;13:53-7.

91. **Bhutta ZA**. Current concepts in the diagnosis and treatment of typhoid fever. *British Medical Journal*. 2006;333(7558):78.
92. **Bhutta ZA, Mansurali N**. Rapid serologic diagnosis of pediatric typhoid fever in an endemic area: a prospective comparative evaluation of two dot-enzyme immunoassays and the Widal test. *The American Journal of Tropical Medicine and Hygiene*. 1999;61(4):654-7.
93. **Wasihun AG, Wlekidan LN, Gebremariam SA, Welderufael AL, Muthupandian S, *et al.*** Diagnosis and Treatment of Typhoid Fever and Associated Prevailing Drug Resistance in Northern Ethiopia. *International Journal of Infectious Diseases*. 2015;35:96-102. doi: 10.1016/j.ijid.2015.04.014. Epub@2015 Apr 27.:96-102.
94. **Levine MM, Grados O, Gilman RH, Woodward WE, Solis-Plaza R, *et al.*** Diagnostic value of the Widal test in areas endemic for typhoid fever. *The American Journal of Tropical Medicine and Hygiene*. 1978;27(4):795-800.
95. **Pang TIKK, Puthucheary SD**. Significance and value of the Widal test in the diagnosis of typhoid fever in an endemic area. *Journal of Clinical Pathology*. 1983;36(4):471-5.
96. **Parry CM, Hoa NTT, Diep TS, Wain J, Chinh NT, *et al.*** Value of a single-tube Widal test in diagnosis of typhoid fever in Vietnam. *Journal of Clinical Microbiology*. 1999;37(9):2882-6.
97. **Keddy KH, Sooka A, Letsoalo ME, Hoyland G, Chaignat CL, *et al.*** Sensitivity and specificity of typhoid fever rapid antibody tests for laboratory diagnosis at two sub-Saharan African sites. *Bulletin of the World Health Organization*. 2011;89(9):640-7.
98. **Hashimoto Y, Itho Y, Fujinaga Y, Khan AQ, Sultana F, *et al.*** Development of nested PCR based on the ViaB sequence to detect *Salmonella* Typhi. *Journal of Clinical Microbiology*. 1995;33(3):775-7.
99. **Prakash P, Mishra OP, Singh AK, Gulati AK, Nath G**. Evaluation of nested PCR in diagnosis of typhoid fever. *Journal of Clinical Microbiology*. 2005;43(1):431-2.
100. **Song JH, Cho H, Park MY, Na DS, Moon HB, *et al.*** Detection of *Salmonella* Typhi in the blood of patients with typhoid fever by polymerase chain reaction. *Journal of Clinical Microbiology*. 1993;31(6):1439-43.
101. **Ranjbar R, Karami A, Farshad S, Giammanco GM, Mammina C**. Typing methods used in the molecular epidemiology of microbial pathogens: a how-to guide. *New Microbiology*. 2014;37(1):1-15.
102. **Nga TVT, Karkey A, Dongol S, Thuy HN, Dunstan S, *et al.*** The sensitivity of real-time PCR amplification targeting invasive *Salmonella* serovars in biological specimens. *BMC Infectious Diseases*. 2010;10(1):1.
103. **Malorny B, Paccassoni E, Fach P, Bunge C, Martin A, *et al.*** Diagnostic real-time PCR for detection of *Salmonella* in food. *Applied & Environmental Microbiology*. 2004;70(12):7046-52.
104. **Grimont PA, Weill F-X**. Antigenic formulae of the *Salmonella* serovars. *WHO Collaborating Centre For Reference And Research on Salmonella*. 2007;9.
105. **Brenner F, Villar R, Angulo F, Tauxe R, Swaminathan B**. *Salmonella* nomenclature. *Journal of Clinical Microbiology*. 2000;38(7):2465-7.
106. **Craigie J, Yen CH**. The Demonstration of Types of B. Typhosus by Means of Preparations of Type II Vi Phage: I. Principles and Technique. *Canadian Public Health Journal*. 1938:448-63.
107. **Anderson ES, Williams REO**. Bacteriophage Typing of Enteric Pathogens and Staphylococci and its Use in Epidemiology: A Review. *Journal of Clinical Pathology*. 1956;9(2):94.

108. **Rowe B, Ward LR, Threlfall EJ**. Multidrug-resistant *Salmonella* Typhi: a worldwide epidemic. *Clinical Infectious Diseases*. 1997;24(Supplement 1):S106-S9.
109. **Byarugaba DK**. Antimicrobial resistance in developing countries and responsible risk factors. *International journal of antimicrobial agents*. 2004;24(2):105-10.
110. **Cooke FJ, Wain J**. The emergence of antibiotic resistance in typhoid fever. *Travel Medicine & Infectious Disease*. 2004;2(2):67-74.
111. **Murgatroyd F**. Typhoid treated with Chloromycetin. *British Medical Journal*. 1949;1(4610):851.
112. **Good RA, Mackenzie RD**. Chloramphenicol in typhoid fever. *The Lancet*. 1950;255(6605):611-5.
113. **Lomax W**. Chloramphenicol in Typhoid Fever. *British Medical Journal*. 1949;2(4633):911.
114. **Woodward TE, Smadel JE, LEY HL, Green R, Mankikar DS**. Preliminary report on the beneficial effect of chloromycetin in the treatment of typhoid fever. *Annals of Internal Medicine*. 1948;29(1):131-4.
115. **Woodward TE, Smadel JE**. Management of typhoid fever and its complications. *Annals of Internal Medicine*. 1964;60(1):144-57.
116. **Wisseman Jr CL, Smadel JE, Hahn FE, Hopps HE**. MODE OF ACTION OF CHLORAMPHENICOL I.: Action of Chloramphenicol on Assimilation of Ammonia and on Synthesis of Proteins and Nucleic Acids in Escherichia coli. *Journal of Bacteriology*. 1954;67(6):662.
117. **Colquhoun J, Weetch RS**. Resistance to chloramphenicol developing during treatment of typhoid fever. *The Lancet*. 1950;256(6639):621-3.
118. **Paniker CKJ, Vimala KN**. Transferable chloramphenicol resistance in *Salmonella* Typhi. *Nature*. 1972;239:109-10.
119. **Anderson ES, Smith HR**. Chloramphenicol resistance in the typhoid bacillus. *British Medical Journal*. 1972;3(5822):329-31.
120. **Olarte J, Galindo E**. *Salmonella* Typhi resistant to chloramphenicol, ampicillin, and other antimicrobial agents: strains isolated during an extensive typhoid fever epidemic in Mexico. *Antimicrobial Agents and Chemotherapy*. 1973;4(6):597-601.
121. **Mirza SH, N.J. B, Hart CA**. Multi-drug resistant typhoid: a global problem. *Journal of Medical Microbiology*. 1996;44:317-9.
122. **Datta N, Richards H, Datta C**. *Salmonella* Typhi in vivo acquires resistance to both chloramphenicol and co-trimoxazole. *The Lancet*. 1981;317(8231):1181-3.
123. **Jesudasan M, John TJ, Gupta BL, Bhujwala RA**. Multiresistant *Salmonella* Typhi in India. *The Lancet*. 1990;336(8709):252.
124. **Hoa NTT, Diep TS, Wain J, Parry CM, Hien TT, *et al.*** Community-acquired septicaemia in southern Viet Nam: the importance of multidrug-resistant *Salmonella* Typhi. *Transactions of the Royal Society of Tropical Medicine & Hygiene*. 1998;92(5):503-8.
125. **Mourad AS, Metwally M, El Deen AN, Threlfall EJ, Rowe B, *et al.*** Multiple-drug-resistant *Salmonella* Typhi. *Clinical Infectious Diseases*. 1993;17(1):135-6.
126. **Wain J, Nga LD, Kidgell C, James K, Fortune S, *et al.*** Molecular analysis of incHI1 antimicrobial resistance plasmids from *Salmonella* serovar Typhi strains associated with typhoid fever. *Antimicrobial Agents and Chemotherapy*. 2003;47(9):2732-9.
127. **Chau TT, Campbell JI, Galindo CM, Hoang NVM, Diep TS, *et al.*** Antimicrobial drug resistance of *Salmonella enterica* serovar Typhi in Asia and molecular mechanism of reduced susceptibility to the fluoroquinolones. *Antimicrobial Agents and Chemotherapy*. 2007;51(12):4315-23.

128. **Holt KE, Parkhill J, Mazzoni CJ, Roumagnac P, Weill FX*, et al.*** High-throughput sequencing provides insights into genome variation and evolution in *Salmonella* Typhi. *Nature Genetics*. 2008;40(8):987-93.
129. **Wain J, Hoa NT, Chinh NT, Vinh H, Everett MJ*, et al.*** Quinolone-resistant *Salmonella* Typhi in Viet Nam: molecular basis of resistance and clinical response to treatment. *Clinical Infectious Diseases*. 1997;25(6):1404-10.
130. **Baker S, Holt K, van de Vosse E, Roumagnac P, Whitehead S*, et al.*** High-throughput genotyping of *Salmonella enterica* serovar Typhi allowing geographical assignment of haplotypes and pathotypes within an urban District of Jakarta, Indonesia. *Journal of Clinical Microbiology*. 2008;46(5):1741-6.
131. **Ray P, Sharma J, Marak RS, Garg RK**. Predictive efficacy of nalidixic acid resistance as a marker of fluoroquinolone resistance in *Salmonella enterica* var Typhi. *Indian Journal of Medical Research*. 2006;124(1):105.
132. **Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ*, et al.*** Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter-and intracontinental transmission events. *Nature Genetics*. 2015;47(6):632-9.
133. **Feasey NA, Gaskell K, Wong V, Msefula C, Selemani G*, et al.*** Rapid Emergence of Multidrug Resistant, H58-Lineage *Salmonella* Typhi in Blantyre, Malawi. 2015.
134. **Roumagnac P, Weill FX, Dolecek C, Baker S, Brisse S*, et al.*** Evolutionary history of *Salmonella* Typhi. *Science*. 2006;314(5803):1301-4.
135. **Baker S, Hanage WP, Holt KE**. Navigating the future of bacterial molecular epidemiology. *Current Opinion in Microbiology*. 2010;13(5):640-5.
136. **Baker S, Dougan G**. The genome of *Salmonella enterica* serovar Typhi. *Clinical Infectious Diseases*. 2007;45(Supplement 1):S29-S33.
137. **Sabat AJ, Budimir A, Nashev D, Sa-Leao R, van Dijl JM*, et al.*** Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Eurosurveillance*. 2013;18(4):20380.
138. **Boxrud D, Pederson-Gulrud K, Wotton J, Medus C, Lyszkowicz E*, et al.*** Comparison of multiple-locus variable-number tandem repeat analysis, pulsed-field gel electrophoresis, and phage typing for subtype analysis of *Salmonella enterica* serotype Enteritidis. *Journal of Clinical Microbiology*. 2007;45(2):536-43.
139. **Thong KL, Cheong YM, Puthucheary S, Koh CL, Pang T**. Epidemiologic analysis of sporadic *Salmonella* Typhi isolates and those from outbreaks by pulsed-field gel electrophoresis. *Journal of Clinical Microbiology*. 1994;32(5):1135-41.
140. **Nair S, Poh CL, Lim YS, Tay L, Goh KT**. Genome fingerprinting of *Salmonella* Typhi by pulsed-field gel electrophoresis for subtyping common phage types. *Epidemiology and Infection*. 1994;113(03):391-402.
141. **Kubota K, Barrett TJ, Ackers ML, Brachman PS, Mintz ED**. Analysis of *Salmonella enterica* Serotype Typhi Pulsed-Field Gel Electrophoresis Patterns Associated with International Travel. *Journal of Clinical Microbiology*. 2005;43(3):1205-9.
142. **Thong KL, Puthucheary S, Yassin RM, Sudarmono P, Padmidewi M*, et al.*** Analysis of *Salmonella* Typhi isolates from Southeast Asia by pulsed-field gel electrophoresis. *Journal of Clinical Microbiology*. 1995;33(7):1938-41.
143. **Fakhr MK, Nolan LK, Logue CM**. Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing *Salmonella enterica* serovar Typhimurium. *Journal of Clinical Microbiology*. 2005;43(5):2215-9.
144. **Pourshafie MR, Saifi M, Mousavi SF, Sedaghat M, Nikbakht GH*, et al.*** Clonal diversity of *Salmonella enterica* serotype Typhi isolated from patients with typhoid fever in Tehran. *Scandinavian Journal of Infectious Diseases*. 2008;40(1):18-23.

145. **Echeita MA, Usera MA**. Chromosomal rearrangements in *Salmonella enterica* serotype Typhi affecting molecular typing in outbreak investigations. *Journal of Clinical Microbiology*. 1998;36(7):2123-6.
146. **Liu SL, Sanderson KE**. Highly plastic chromosomal organization in *Salmonella* Typhi. *Proceedings of the National Academy of Sciences*. 1996;93(19):10303-8.
147. **Maiden MC, Bygraves JA, Feil E, Morelli G, Russell JE*, et al.*** Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences*. 1998;95(6):3140-5.
148. **Hughes LA, Wigley P, Bennett M, Chantrey J, Williams N**. Multi-locus sequence typing of *Salmonella enterica* serovar Typhimurium isolates from wild birds in northern England suggests host-adapted strain. *Letters in Applied Microbiology*. 2010;51(4):477-9.
149. **Dahiya S, Kapil A, Kumar R, Das BK, Sood S*, et al.*** Multiple locus sequence typing of *Salmonella* Typhi, isolated in north India-a preliminary study. *The Indian Journal of Medical Research*. 2013;137(5):957.
150. **Techaruvichit P, Takahashi H, Vesaratchavest M, Keeratipibul S, Kuda T*, et al.*** Development of Multiple-Locus Variable Number Tandem Repeat Analysis for Molecular Subtyping of *Campylobacter* jejuni by Using Capillary Electrophoresis. *Applied & Environmental Microbiology*. 2015:AEM-01151.
151. **Malorny B, Junker E, Helmuth R**. Multi-locus variable-number tandem repeat analysis for outbreak studies of *Salmonella enterica* serotype Enteritidis. *BMC Microbiology*. 2008;8(1):84.
152. **Lindstedt BA, Vardund T, Aas L, Kapperud G**. Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. *Journal of Microbiological Methods*. 2004;59(2):163-72.
153. **Larsson JT, Torpdahl M, Petersen RF, Sørensen G, Lindsted BA*, et al.*** Development of a new nomenclature for *Salmonella* Typhimurium multilocus variable number of tandem repeats analysis (MLVA). *Eurosurveillance (Online Edition)*. 2009;14(15).
154. **Hopkins KL, Peters TM, De Pinna E, Wain J**. Standardisation of multilocus variable-number tandem-repeat analysis (MLVA) for subtyping of *Salmonella enterica* serovar Enteritidis. *Eurosurveillance*. 2011;16(32).
155. **Van Cuyck H, Farbos-Granger A, Leroy P, Yith V, Guillard B*, et al.*** MLVA polymorphism of *Salmonella enterica* subspecies isolated from humans, animals, and food in Cambodia. *BMC research notes*. 2011;4(1):306.
156. **Gilbert GL**. Using MLVA to type strains of *Salmonella* Typhimurium in New South Wales. *New South Wales Public Health Bulletin*. 2008;19(2):29-31.
157. **Hill V, Zozio T, Sadikalay S, Viegas S, Streit E*, et al.*** MLVA based classification of Mycobacterium tuberculosis complex lineages for a robust phylogeographic snapshot of its worldwide molecular diversity. *PloS One*. 2012;7(9):e41991.
158. **Li Y, Cui Y, Hauck Y, Platonov ME, Dai E*, et al.*** Genotyping and phylogenetic analysis of Yersinia pestis by MLVA: insights into the worldwide expansion of Central Asia plague foci. *PloS One*. 2009;4(6):e6000.
159. **Tong CX, ZHAO Xq, MA Jj, LIU Zg, CAO Wj*, et al.*** Characterization of Mycobacterium tuberculosis isolates from patients in Gansu by multiple locus variable numbers of tandem repeats. *Chinese Journal of Zoonoses*. 2010;26:627-30.

160. **Fasanella A, Van Ert M, Altamura SA, Garofolo G, Buonavoglia C*, et al.*** Molecular diversity of Bacillus anthracis in Italy. *Journal of Clinical Microbiology*. 2005;43(7):3398-401.
161. **Liu Y, Shi X, Li Y, Chen Q, Jiang M*, et al.*** The evaluation and application of multilocus variable number tandem repeat analysis (MLVA) for the molecular epidemiological study of *Salmonella enterica* subsp. *enterica* serovar Enteritidis infection. *Annals Of Clinical Microbiology and Antimicrobials*. 2016;15(1):1.
162. **Wuyts V, Mattheus W, de Bex GDL, Wildemauwe C, Roosens NH*, et al.*** MLVA as a tool for public health surveillance of human *Salmonella* Typhimurium: prospective study in Belgium and evaluation of MLVA loci stability. *PLoS One*. 2013;8(12):e84055.
163. **Joensen KG, Scheutz F, Lund O, Hasman H, Kaas RS*, et al.*** Real-time whole-genome sequencing for routine typing, surveillance, and outbreak detection of verotoxigenic Escherichia coli. *Journal of Clinical Microbiology*. 2014;52(5):1501-10.
164. **Deng X, Shariat N, Driebe EM, Roe CC, Tolar B*, et al.*** Comparative Analysis of Subtyping Methods against a Whole-Genome-Sequencing Standard for *Salmonella enterica* Serotype Enteritidis. *Journal of Clinical Microbiology*. 2015;53(1):212-8.
165. **Zhang S, Yin Y, Jones MB, Zhang Z, Kaiser BLD*, et al.*** *Salmonella* Serotype Determination Utilizing High-throughput Genome Sequencing Data. *Journal of Clinical Microbiology*. 2015:JCM-00323.
166. **Ashton PM, Peters T, Ameh L, McAleer R, Petrie S*, et al.*** Whole Genome Sequencing for the Retrospective Investigation of an Outbreak of *Salmonella* Typhimurium DT 8. *PLoS currents*. 2014;7.
167. **Ansorge WJ**. Next-generation DNA sequencing techniques. *New Biotechnology*. 2009;25(4):195-203.
168. **Donkor ES**. Sequencing of bacterial genomes: principles and insights into pathogenesis and development of antibiotics. *Genes*. 2013;4(4):556-72.
169. **Octavia S, Wang Q, Tanaka MM, Kaur S, Sintchenko V, *et al.*** Delineating community outbreaks of *Salmonella enterica* serovar Typhimurium using whole genome sequencing: insights into genomic variability within an outbreak. *Journal of Clinical Microbiology*. 2015:JCM-03235.
170. **Morozova O, Marra MA**. Applications of next-generation sequencing technologies in functional genomics. *Genomics*. 2008;92(5):255-64.
171. **Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H*, et al.*** Multilocus sequence typing of total genome sequenced bacteria. *Journal of Clinical Microbiology*. 2012:JCM. 06094-11.
172. **Taylor AJ, Lappi V, Wolfgang WJ, Lapierre P, Palumbo MJ*, et al.*** Characterization of Foodborne Outbreaks of *Salmonella enterica* Serovar Enteritidis with Whole-Genome Sequencing Single Nucleotide Polymorphism-Based Analysis for Surveillance and Outbreak Detection. *Journal of Clinical Microbiology*. 2015;53(10):3334-40.
173. **Kwong JC, Mercoulia K, Tomita T, Easton M, Li HY*, et al.*** Prospective whole-genome sequencing enhances national surveillance of Listeria monocytogenes. *Journal of Clinical Microbiology*. 2016;54(2):333-42.
174. **Wuyts V, Denayer S, Roosens NH, Mattheus W, Bertrand S*, et al.*** Whole Genome Sequence Analysis of *Salmonella* Enteritidis PT4 Outbreaks from a National Reference Laboratory's Viewpoint. *PLoS Currents*. 2015;7.

175. **Tan K-K, Tan Y-C, Chang L-Y, Lee KW, Nore SS*, et al.*** Full genome SNP-based phylogenetic analysis reveals the origin and global spread of Brucella melitensis. *BMC Genomics*. 2015;16(1):1.
176. **Leekitcharoenphon P, Nielsen EM, Kaas RS, Lund O, Aarestrup FM**. Evaluation of whole genome sequencing for outbreak detection of *Salmonella enterica*. *PloS One*. 2014;9(2):e87991.
177. **Byrne L, Fisher I, Peters T, Mather A, Thomson N*, et al.*** A multi-country outbreak of *Salmonella* Newport gastroenteritis in Europe associated with watermelon from Brazil, confirmed by whole genome sequencing: October 2011 to January 2012. 2014.
178. **den Bakker HC, Allard MW, Bopp D, Brown EW, Fontana J*, et al.*** Rapid whole-genome sequencing for surveillance of *Salmonella enterica* serovar enteritidis. *Emerging Infectious Diseases*. 2014;20(8):1306.
179. **Cao G, Meng J, Strain E, Stones R, Pettengill J*, et al.*** Phylogenetics and differentiation of *Salmonella* Newport lineages by whole genome sequencing. *PloS One*. 2013;8(2):e55687.
180. **Crump JA, Kretsinger K, Gay K, Hoekstra RM, Vugia DJ*, et al.*** Clinical response and outcome of infection with *Salmonella enterica* serotype Typhi with decreased susceptibility to fluoroquinolones: a United States foodnet multicenter retrospective cohort study. *Antimicrobial Agents and Chemotherapy*. 2008;52(4):1278-84.
181. **Effa EE, Lassi ZS, Critchley JA, Garner P, Sinclair D*, et al.*** Fluoroquinolones for treating typhoid and paratyphoid fever (enteric fever). *The Cochrane Library*. 2011.
182. **Girgis NI, Butler T, Frenck RW, Sultan Y, Brown FM*, et al.*** Azithromycin versus ciprofloxacin for treatment of uncomplicated typhoid fever in a randomized trial in Egypt that included patients with multidrug resistance. *Antimicrobial Agents and Chemotherapy*. 1999;43(6):1441-4.
183. **Frenck RW, Mansour A, Nakhla I, Sultan Y, Putnam S*, et al.*** Short-course azithromycin for the treatment of uncomplicated typhoid fever in children and adolescents. *Clinical Infectious Diseases*. 2004;38(7):951-7.
184. **Kobayashi T, Hayakawa K, Mawatari M, Mezaki K, Takeshita N*, et al.*** Case report: failure under azithromycin treatment in a case of bacteremia due to *Salmonella enterica* Paratyphi A. *BMC Infectious Diseases*. 2014;14(1):404.
185. **Capoor MR, Rawat D, Nair D, Hasan AS, Deb M*, et al.*** In vitro activity of azithromycin, newer quinolones and cephalosporins in ciprofloxacin-resistant *Salmonella* causing enteric fever. *J Med Microbiol*. 2007;56:1490-4.
186. Govender N, Quan V, Crowther P, du Plessis D, Keddy K*, et al. GERMS-SA annual report, National Institute for Communicable Diseases, a Division of the NHLS*. 2010 2010. Report No.
187. **Von Kalckreuth V, Konings F, Aaby P, Adu-Sarkodie Y, Ali M*, et al.*** The Typhoid Fever Surveillance in Africa Program (TSAP): clinical, diagnostic, and epidemiological methodologies. *Clinical Infectious Diseases*. 2016;62(suppl 1):S9-S16.
188. **Al-Emran HM, Hahn A, Baum J, Espinoza LMC, Deerin J*, et al.*** Diagnosing *Salmonella enterica* Serovar Typhi Infections by Polymerase Chain Reaction Using EDTA Blood Samples of Febrile Patients From Burkina Faso. *Clinical Infectious Diseases*. 2016;62(suppl 1):S37-S41.
189. Panzner U, Pak GD, Im J, Deerin J, Soura A*, et al.* Typhoid fever surveillance in Africa program: Healthcare patterns in febrile study populations. International Journal of Infectious Diseases: Elsevier. p. 246-7.
190. **Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB*, et al.*** Standardization of pulsed-field gel electrophoresis protocols for the subtyping of Escherichia coli

O157: H7, *Salmonella*, and Shigella for PulseNet. *Foodbourne Pathogens and Disease*. 2006;3(1):59-67.

191. **Hunter SB, Vauterin P, Lambert-Fair MA, Van Duyne MS, Kubota K, *et al.*** Establishment of a universal size standard strain for use with the PulseNet standardized pulsed-field gel electrophoresis protocols: converting the national databases to the new size standard. *Journal of Clinical Microbiology*. 2005;43(3):1045-50.
192. **Lindstedt BA, Torpdahl M, Vergnaud G, Le Hello S, Weill FX, *et al.*** Use of multilocus variable-number tandem repeat analysis (MLVA) in eight European countries, 2012. *Eurosurveillance*. 2013.
193. **Octavia S, Lan R**. Multiple-locus variable-number tandem-repeat analysis of *Salmonella enterica* serovar Typhi. *Journal of Clinical Microbiology*. 2009;47(8):2369-76.
194. **Ramisse V, Houssu P, Hernandez E, Denoeud F, Hilaire V, *et al.*** Variable number of tandem repeats in *Salmonella enterica* subsp. *enterica* for typing purposes. *Journal of Clinical Microbiology*. 2004;42(12):5722-30.
195. **Tien YY, Ushijima H, Mizuguchi M, Liang SY, Chiou CS**. Use of multilocus variable-number tandem repeat analysis in molecular subtyping of *Salmonella enterica* serovar Typhi isolates. *Journal of Medical Microbiology*. 2012;61(2):223-32.
196. **Simpson EH**. Measurement of diversity. *Nature*. 1949.
197. **Hunter PR, Gaston MA**. Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *Journal of Clinical Microbiology*. 1988;26(11):2465-6.
198. **Nadon CA, Trees E, Ng LK, Møller Nielsen E, Reimer A, *et al.*** Development and application of MLVA methods as a tool for inter-laboratory surveillance. *Eurosurveillance*. 2013;18(35):20565.
199. Scientific TF. DNA Fragment Analysis by Capillary Electrophoresis. 2014.
200. **Wallace DL**. A method for comparing two hierarchical clustering: Comment. *Journal of the American Statistical Association*. 1983;78(383):569-76.
201. **Crump JA, Youssef FG, Luby SP, Wasfy MO, Rangel JM, *et al.*** Estimating the incidence of typhoid fever and other febrile illnesses in developing countries. *Emerging Infectious Diseases*. 2003;9(5):539.
202. **Murgia M, Rubino S, Wain J, Gaind R, Paglietti B**. A novel broadly applicable PCR-RFLP method for rapid identification and subtyping of H58 *Salmonella* Typhi. *Journal of Microbiological Methods*. 2016;127:219-23.
203. **Qin X, Razia Y, Johnson JR, Stapp JR, Boster DR, *et al.*** Ciprofloxacin-resistant gram-negative bacilli in the fecal microflora of children. *Antimicrobial Agents and Chemotherapy*. 2006;50(10):3325-9.
204. **Wang H, Diao B, Cui Z, Yan M, Kan B**. Genotyping of *Salmonella* Typhi using 8-loci multi locus VNTR analysis. *Gut Pathogens*. 2016;8(1):14.

Appendix A

Preparation of reagents for PFGE

A.1 Cell suspension buffer

100 mM Tris, pH 8.0 (Merck Chemical Ltd., Nottingham, England and Wales)

100 mM EDTA, pH 8.0 (Sigma-Aldrich, Inc.)

A.2 Tris: EDTA (TE) Buffer

10 mM Tris, pH 8.0 (Merck)

1 mM EDTA, pH 8.0 (Sigma)

A.3 Proteinase K (20mg/ml stock)

20mg Proteinase K powder (Roche Diagnostics GmbH, Mannheim, Germany)

1ml TE buffer

A.4 10% Sodium dodecyl sulphate

10 g sodium dodecyl sulphate (SDS) (Merck)

100 ml deionized water (dH_2O)

A.5 1% SeaKem® Gold agarose: 1% SDS (for agarose plugs)

0.25 g SeaKem® Gold agarose (Lonza, Rockland, USA)

22.5 ml TE buffer

2.5 ml 10% SDS (Merck)

Boil to dissolve. Keep at 55-60°C while in use

A.6 Cell Lysis Buffer

50 mM Tris, pH 8.0 (Merck)

50 mM EDTA, pH 8.0 (Sigma)

1% N-Lauroylsarcosine, Sodium salt (Sarcosyl)

0.1 mg/ml Proteinase K (Roche)

A.7 *Xba*I Restriction Enzyme buffer µl/plug slice

20 µl of 10X H restriction buffer (Roche)

180µl of sterile dH_2O

A.8 *Xba*I Restriction Enzyme Master Mix: (50U/plug slice)

20 µl of 10X H restriction Buffer (Roche)

5 µl of *Xba*I Enzyme (10U/µl) (Roche)

175µl of sterile dH_2O

A.9 0.5X Tris-borate EDTA (TBE) buffer

50 ml 10X TBE Buffer (Merck)

900 ml dH_2O

A.10 1% SeaKem® Gold Agarose in 0.5X TBE

1.5 g of SeaKem® Gold Agarose (Lonza)

150 ml of 0.5X TBE

Boil to dissolve

A.11 Ethidium bromide solution (10 mg/ml)

1 g Ethidium bromide

100 ml distilled H2O

A.12 Ethidium bromide (EtBr) staining solution

25 µl of ethidium bromide stock solution (10 mg/ml) (Merck)

250 ml of TBE buffer (Merck)

Appendix B

PCR amplification of VNTR loci

B.1 Preparation of 20 µM primer working solutions from 100 µM stock solutions

c1v1=c2v2

100 µM x v1=20 µM X 100 µl

v1=20 µl (Add 20 µl of primer stock solution to 80 µl of distilled H2O)

B.2 Simplex PCR setup for VNTR loci screening

PCR reaction mix	
Qiagen multiplex PCR kit (Qiagen, Hilden, Germany)	
Distilled H_2O	7 µl
Qiagen master mix Q solution Forward primer (20 µM) Reverse Primer (20 µM)	12.5 µl 2.5 µl 1 µl 1 µl
DNA template	1 µl
Final volume	25 µl

B.3 Multiplex PCR setup for MLVA typing

B.3.1 The forward and reverse multiplex primer mix

VNTR primers	**Forward primer mix (Volume)**	**Reverse primer mix (Volume)**
TR1	6 µl	6 µl
TR2	6 µl	6 µl
Sal02	5 µl	5 µl
Sal20	3.5 µl	3.5 µl
TR4699	4 µl	4 µl
Distilled H2O	75.5 µl	75.5 µl
Final volume	100 µl	100 µl

B.3.2 Multiplex PCR setup for MLVA typing

PCR reaction mix	
Qiagen multiplex PCR kit (Qiagen, Hilden, Germany)	
Distilled H_2O	7 µl
Qiagen master mix Q solution Forward primer mix Reverse Primer mix	12.5 µl 2.5 µl 1 µl 1 µl
DNA template	1 µl
Final volume	25 µl

B.4 Determination of primer concentration in PCR reaction

Primer concentration in primer mix

c1v1=c2v2

(Concentration in working solution)(Volume added to primer mix) = (Concentration in primer mix) (Primer mix final volume)

Example: VNTR TR1

20 µM x 6 µl= c2 X 100 µl

c2= 1.2 µM

VNTR TR1 primer concentration in primer mix = 1.2 µM

Appendix C

Sequencing of VNTR loci using the ABI Prism®BigDye® Terminator cycle sequencing kit

C.1 Nucleotide sequencing of VNTR loci

PCR reaction mix	
ABI Prism®BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems)	
Distilled H_2O	7.5 µl
5X Sequencing buffer BigDye terminator reaction mix Forward primer (5 µM) Reverse Primer (5 µM)	1.5 µl 3 µl 1 µl 1 µl
DNA template	1 µl
Final volume	15 µl

C.2 Cycle sequencing PCR conditions

PCR cycle	**Temperature**	**Time**	**Number of cycles**
Initial denaturation	95 °C	2 min	1
Denaturation Annealing Extension	95 °C 50 °C 60 °C	30 sec 30 sec 4 min	25
	6 °C	∞	1

Appendix D

VNTR sequence alignments

D.1 Sequence alignment: VNTR locus TR1

Multiple sequence alignment of the VNTR locus TR1 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat region.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TCD105056    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD461595    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD458017    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD409778    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD488042    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD316823    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD281699    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD241722    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD240006    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA
TCD537014    GCCAACGATC GCTACTTTTT TCATCATTCA TCCTGCATCA GAAGAAAGAA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                         60         70         80         90        100
TCD105056    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAATT GTGGCATTAA
TCD461595    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
TCD458017    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAATT GTGGCATTAA
TCD409778    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
TCD488042    GAAAGAAGAA AGAAGAAAGA AGAATTGTGG CATTAAAGCC CTGACGCCGC
TCD316823    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
TCD281699    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
TCD241722    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
TCD240006    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG
```

```
TCD537014    GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA AAGAAGAAAG AAGAAAGAAG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     110        120        130        140        150
TCD105056    AGCCCTGACG CCGCCAGCCT TTCCAGGCGC TGGCCGCAAT TTGATCCAGT
TCD461595    AAAGAAGAAA GAAGAAAGAA GAAAGAAGAA TTGTGGCATT AAAGCCCTGA
TCD458017    AGCCCTGACG CCGCCAGCCT TTCCAGGCGC TGGCCGCAAT TTGATCCAGT
TCD409778    AATTGTGGCA TTAAAGCCCT GACGCCGCCA GCCTTTCCAG GCGCTGGCCG
TCD488042    CAGCCTTTCC AGGCGCTGGC CGCAATTTGA TCCAGTCGGT ATTACCGCTT
TCD316823    AAAGAAGAAA GAAGAAAGAA GAAAGAAGAA TTGTGGCATT AAAGCCCTGA
TCD281699    AAAGAAGAAA GAAGAATTGT GGCATTAAAG CCCTGACGCC GCCAGCCTTT
TCD241722    AAAGAAGAAA GAAGAAAGAA GAAAGAAGAA AGAAGAAAGA AGAAAGAAGA
TCD240006    AAAGAAGAAA GAAGAATTGT GGCATTAAAG CCCTGACGCC GCCAGCCTTT
TCD537014    AAAGAAGAAA GAAGAAAGAA GAAAGAAGAA AGAAGAATTG TGGCATTAAA

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     160        170        180        190        200
TCD105056    CGGTATTACC GCTTATTAGG TGTAGGCAAA TTCTGTGATC GGTGTAGTAT
TCD461595    CGCCGCCAGC CTTTCCAGGC GCTGGCCGCA ATTTGATCCA GTCGGTATTA
TCD458017    CGGTATTACC GCTTATTAGG TGTAGGCAAA TTCTGTGATC GGTGTAGTAT
TCD409778    CAATTTGATC CAGTCGGTAT TACCGCTTAT TAGGTGTAGG CAAATTCTGT
TCD488042    ATTAGGTGTA GGCAAATTCT GTGATCGGTG TAGTATGCGC ACTTCTTGA.
TCD316823    CGCCGCCAGC CTTTCCAGGC GCTGGCCGCA ATTTGATCCA GTCGGTATTA
TCD281699    CCAGGCGCTG GCCGCAATTT GATCCAGTCG GTATTACCGC TTATTAGGTG
TCD241722    ATTGTGGCAT TAAAGCCCTG ACGCCGCCAG CCTTTCCAGG CGCTGGCCGC
TCD240006    CCAGGCGCTG GCCGCAATTT GATCCAGTCG GTATTACCGC TTATTAGGTG
TCD537014    GCCCTGACGC CGCCAGCCTT TCCAGGCGCT GGCCGCAATT TGATCCAGTC

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     210        220        230        240        250
TCD105056    GCGCACTTCT TGA..... .......... .......... ..........
TCD461595    CCGCTTATTA GGTGTAGGCA AATTCTGTGA TCGGTGTAGT ATGCGCACTT
TCD458017    GCGCACTTCT TGA..... .......... .......... ..........
TCD409778    GATCGGTGTA GTATGCGCAC TTCTTGA... .......... ..........
TCD488042    .......... .......... .......... .......... ..........
TCD316823    CCGCTTATTA GGTGTAGGCA AATTCTGTGA TCGGTGTAGT ATGCGCACTT
TCD281699    TAGGCAAATT CTGTGATCGG TGTAGTATGC GCACTTCTTG AA........
TCD241722    AATTTGATCC AGTCGGTATT ACCGCTTATT AGGTGTAGGC AAATTCTGTG
```

```
TCD240006    TAGGCAAATT CTGTGATCGG TGTAGTATGC GCACTTCTTG A.........
TCD537014    GGTATTACCG CTTATTAGGT GTAGGCAAAT TCTGTGATCG GTGTAGTATG

                ....|....| ....|....| ....|..
                       260        270
TCD105056    .......... .......... .......
TCD461595    CTTGA..... .......... .......
TCD458017    .......... .......... .......
TCD409778    .......... .......... .......
TCD488042    .......... .......... .......
TCD316823    CTTGA..... .......... .......
TCD281699    .......... .......... .......
TCD241722    ATCGGTGTAG TATGCGCACT TCTTGA.
TCD240006    .......... .......... .......
TCD537014    CGCACTTCTT GA........ .......
```

D.2 Sequence Alignment: VNTR locus TR2

Multiple sequence alignment of the VNTR locus TR2 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat region.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TCD139882    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD723719    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TMI87197     TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TMI87199     TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD298424    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD316823    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD422011    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD671445    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD723719    TCCCTGTTTT TCGTGCTGAT ACGCCGCCGC TTCCCGCTGA AGCCGCGCCC
TCD674832    TCCCTGTTTT TCGTGCTGAT ACGTCGCCGC TTCCCGCTGA AGCCGCGCCC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
TCD139882       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD723719       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TMI87197        GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TMI87199        GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD298424       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD316823       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD422011       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD671445       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD723719       GGAATAAGCT CCCGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC
TCD674832       GGAATAAGCT CCTGGCCAGC CTGATACGAG GTGTCGGGCA CAAAAAAGGC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
TCD139882       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD723719       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TMI87197        GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TMI87199        GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD298424       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD316823       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD422011       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD671445       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD723719       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT
TCD674832       GACTTTCGTT GAGTCGCCTT TTCTTATCCC CTATGGGAGC GCGGTGCCTT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
TCD139882       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD723719       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TMI87197        CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TMI87199        CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD298424       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD316823       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD422011       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD671445       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD723719       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
TCD674832       CCAGGCATTT ATTTACGAAG CATGACTTCG ATAAAATCTT TCCAGTTCCC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        210        220        230        240        250
TCD139882       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCACGT
TCD723719       CAGTTCCCCA GTTCACGTTC AATCATAATA GCCTCTCTTA TTATTATGGG
TMI87197        CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TMI87199        CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TCD298424       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TCD316823       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TCD422011       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TCD671445       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA
TCD723719       CAGTTCCCCA GTTCACGTTC AATCATAATA GCCTCTCTTA TTATTATGGG
TCD674832       CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        260        270        280        290        300
TCD139882       TCAATCATAA TAGCCTCTCT TATTATTATG GGTATTCTAC GTAGTTAGCG
TCD723719       TATTCTACGT AGTTAGCGGT ATAGAGAGAA GTTCATTTAA CCGATTGTTG
TMI87197        GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT
TMI87199        GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT
TCD298424       GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCACGT TCAATCATAA
TCD316823       GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCACGT TCAATCATAA
TCD422011       GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT
TCD671445       GTTCCCCAGT TCCCCAGTTC ACGTTCAATC ATAATAGCCT CTCTTATTAT
TCD723719       TATTCTACGT AGTTAGCGGT ATAGAGAGAA GTTCATTTAA CCGATTGTTG
TCD674832       GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        310        320        330        340        350
TCD139882       GTATAGAGAG AAGTTCATTT AACCGATTGT TGCGATATCC TCTGA.....
TCD723719       CGATATCCTC TGA..... .......... .......... .......... ..........
TMI87197        TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCACGTTC AATCATAATA
TMI87199        TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCACGTTC AATCATAATA
TCD298424       TAGCCTCTCT TATTATTATG GGTATTCTAC GTAGTTAGCG GTATAGAGAA
TCD316823       TAGCCTCTCT TATTATTATG GGTATTCTAC GTAGTTAGCG GTATAGAGAG
TCD422011       TCACGTTCAA TCATAATAGC CTCTCTTATT ATTATGGGTA TTCTACGTAG
TCD671445       TATGGGTATT CTACGTAGTT AGCGGTATAG AGAGAAGTTC ATTTAACCGA
TCD723719       CGATATCCAC TGA..... .......... .......... .......... ..........
TCD674832       TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        360        370        380        390        400
TCD139882       .......... .......... .......... .......... ..........
```

```
TCD723719    .......... .......... .......... .......... ..........
TMI87197     GCCTCTCTTA TTATTATGGG TATTCTACGT AGTTAGCGGT ATAGAGAGAA
TMI87199     GCCTCTCTTA TTATTATGGG TATTCTACGT AGTTAGCGGT ATAGAGAGAA
TCD298424    GAAGTTCATT TAACCGATTG TTGCGATATC CTCTGA.... ..........
TCD316823    AAGTTCATTT AACCGATTGT TGCGATATCC TCTGA..... ..........
TCD422011    TTAGCGGTAT AGAGAGAAGT TCATTTAACC GATTGTTGCG ATATCCTCTG
TCD671445    TTGTTGCGAT ATCCTCTGAA A......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TCD674832    CCCAGTTCCC CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     410        420        430        440        450
TCD139882    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TMI87197     GTTCATTTAA CCGATTGTTG CGATATCCTC TGA....... ..........
TMI87199     GTTCATTTAA CCGATTGTTG CGATATCCTC TGA....... ..........
TCD298424    .......... .......... .......... .......... ..........
TCD316823    .......... .......... .......... .......... ..........
TCD422011    A......... .......... .......... .......... ..........
TCD671445    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TCD674832    CAGTTCCCCA GTTCCCCAGT TCCCCAGTTC CCCAGTTCCC CAGTTCCCCA

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     460        470        480        490        500
TCD139882    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TMI87197     .......... .......... .......... .......... ..........
TMI87199     .......... .......... .......... .......... ..........
TCD298424    .......... .......... .......... .......... ..........
TCD316823    .......... .......... .......... .......... ..........
TCD422011    .......... .......... .......... .......... ..........
TCD671445    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TCD674832    GTTCCCCAGT TCCCCAGTTC ACGTTCAATC ATAATAGCCT CTCTTATTAT

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     510        520        530        540        550
TCD139882    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TMI87197     .......... .......... .......... .......... ..........
```

```
TMI87199     .......... .......... .......... .......... ..........
TCD298424    .......... .......... .......... .......... ..........
TCD316823    .......... .......... .......... .......... ..........
TCD422011    .......... .......... .......... .......... ..........
TCD671445    .......... .......... .......... .......... ..........
TCD723719    .......... .......... .......... .......... ..........
TCD674832    TATGGGTATT CTACGTAGTT AGCGGTATAG AGAGAAGTTC ATTTAACCGA

             ....|....| ....|...
                    560
TCD139882    .......... ........
TCD723719    .......... ........
TMI87197     .......... ........
TMI87199     .......... ........
TCD298424    .......... ........
TCD316823    .......... ........
TCD422011    .......... ........
TCD671445    .......... ........
TCD723719    .......... ........
TCD674832    TTGTTGCGAT ATCCACTG
```

D.3 Sequence Alignment: VNTR locus Sal02

Multiple sequence alignment of the VNTR locus Sal02 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat region.

```
             ....|....| ....|....| ....|....| ....|....| ....|....|
                      10         20         30         40         50
TCD152229    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD186374    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD402284    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD428468    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD671445    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD537014    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD146495    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD281699    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
TCD316823    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
```

```
TCD530747    TCGATAGACA GCACCAGCAG AATACACAGT TGTTCCATCG GAACGCCCGG
             ....|....| ....|....| ....|....| ....|....| ....|....|
                      60         70         80         90        100
TCD152229    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD186374    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD402284    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD428468    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD671445    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD537014    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD146495    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD281699    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD316823    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG
TCD530747    AATGCCTTTA CCCACCGCCA GAATGACCGG CAGCATGGTC GCGGTGTGAG

             ....|....| ....|....| ....|....| ....|....| ....|....|
                      110        120        130        140        150
TCD152229    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD186374    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD402284    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD428468    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD671445    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD537014    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD146495    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD281699    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD316823    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT
TCD530747    CGGAAAGACT GGCGAACAAA TAGTGCGCGA AATAGAATAC CAGTACCAGT

             ....|....| ....|....| ....|....| ....|....| ....|....|
                      160        170        180        190        200
TCD152229    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD186374    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD402284    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD428468    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD671445    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD537014    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD146495    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD281699    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
TCD316823    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
```

```
TCD530747    ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA GTACCAGTAC
             ....|....| ....|....| ....|....| ....|....| ....|....|
                     210        220        230        240        250
TCD152229    CAGTACCAGT ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA
TCD186374    CAGTACCAGT ACCAGTACCA GTACCAGTAC AATGACAGTT GCATCCGGCG
TCD402284    CAGTACCAGT ACCAGTACCA GTACCAGTAC CAGTACCAGT ACCAGTACCA
TCD428468    CAGTACCAGT ACCAGTACCA GTACCAGTAC AATGACAGTT GCATCCGGCG
TCD671445    CAGTACCAGT ACCAGTACCA GTACAATGAC GGTTGCATCC GGCGAGAAAC
TCD537014    CAGTACCAGT ACCAGTACCA GTACCAGTAC AATGACGGTT GCATCCGGCG
TCD146495    CAGTACCAGT ACCAGTACAA TGACGGTTGC ATCCGGCGAG AAACCTTCCA
TCD281699    CAGTACCAGT ACCAGTACCA GTACCAGTAC CAGTACAATG ACAGTTGCAT
TCD316823    CAGTACCAGT ACCAGTACAA TGACGGTTGC ATCCGGCGAG AAACCTTCCA
TCD530747    CAGTACCAGT ACCAGTACAA TGACGGTTGC ATCCGGCGAG AAACCTTCCA

             ....|....| ....|....| ....|....| ....|....| ....|....|
                     260        270        280        290        300
TCD152229    GTACCAGTAC CAGTACCAGT ACAATGACGG TTGCATCCGG CGAGAAACCT
TCD186374    AGAAACCTTC CAGATGCGTA CTCATGGTAT TGGCGAA... ..........
TCD402284    GTACCAGTAC CAGTACAATG ACGGTTGCAT CCGGCGAGAA ACCTTCCAGA
TCD428468    AGAAACCTTC CAGATGCGTA CTCATGGTAT TGGCGAA... ..........
TCD671445    CTTCCAGATG CGTACTCATG GTATTGGCGA A......... ..........
TCD537014    AGAAACCTTC CAGATGCGTA CTCATGGTAT TGGCGAA... ..........
TCD146495    GATGCGTACT CATGGTATTG GCGAA..... .......... ..........
TCD281699    CCSGCGAGAA ACCTTCCAGA TGCGTACTCA TGGTATTGGC GAA.......
TCD316823    GATGCGTACT CATGGTATTG GCGAA..... .......... ..........
TCD530747    GATGCGTACT CATGGTATTG GCGAA..... .......... ..........

             ....|....| ....|....| ....|....
                     310        320
TCD152229    TCCAGATGCG TACTCATGGT ATTGGCGAA
TCD186374    .......... .......... .........
TCD402284    TGCGTACTCA TGGTATTGGC GAA......
TCD428468    .......... .......... .........
TCD671445    .......... .......... .........
TCD537014    .......... .......... .........
TCD146495    .......... .......... .........
TCD281699    .......... .......... .........
TCD316823    .......... .......... .........
TCD530747    .......... .......... .........
```

D.4 Sequence Alignment: VNTR locus TR4699

Multiple sequence alignment of the VNTR locus TR4699 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat region.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TCD105056       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD279436       TTGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD632685       TTGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD678093       TTGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD139882       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD422011       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD558851       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD375103       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD428468       TTGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT
TCD537014       TCGGGCAATT CGAGATAGGT AGAGAACAAA TTGCTGATCT TGGTGTTGGT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                         60         70         80         90        100
TCD105056       GTTGGTGTTG GTGTTGGTGT TGGGGTTCTT GATGTTAGCA TTTATGAAAA
TCD279436       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
TCD632685       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGGGT
TCD678093       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
TCD139882       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
TCD422011       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGGGTTC TTGATGTTAG
TCD558851       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
TCD375103       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGGGTTC TTGATGTTAG
TCD428468       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
TCD537014       GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
TCD105056       AAATGGGCAG GTCCAAAACT ATACAGTGCC ATATTCAACT CCTGTATTAT
TCD279436       TGGTGTTGGT GTTGGTGTTG GTGTTGGGGT TCTTGATGTT AGCATTTATG
TCD632685       TCTTGATGTT AGCATTTATG AAAAAAATGG GCAGGTCCAA AACTATACAG
TCD678093       TGGTGTTGGT GTTGGGGTTC TTGATGTTAG CATTTATGAA AAAAATGGGC
TCD139882       TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT TGGGGTTCTT GATGTTAGCA
TCD422011       CATTTATGAA AAAAATGGGC AGGTCCAAAA CTATACAGTG CCATATTCAA
TCD558851       TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT TGGTGTTGGT GTTGGTGTTG
TCD375103       CATTTATGAA AAAAATGGGC AGGTCCAAAA CTATACAGTG CCATATTCAA
TCD428468       TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT TGGGGTTCTT GATGTTAGCA
TCD537014       TGGTGTTGGT GTTGGTGTTG GTGTTGGTGT TGGTGTTGGG GTTCTTGATG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
TCD105056       CTTTGCCTGA TGGATATTCT AAATATAGTG TAACTATTGG TAGATACAGG
TCD279436       AAAAAAATGG GCAGGTCCAA AACTATACAG TGCCATATTC AACTCCTGTA
TCD632685       TGCCATATTC AACTCCTGTA TTATCTTTGC CTGATGGATA TTCTAAATAT
TCD678093       AGGTCCAAAA CTATACAGTG CCATATTCAA CTCCTGTATT ATCTTTGCCT
TCD139882       TTTATGAAAA AAATGGGCAG GTCCAAAACT ATACAGTGCC ATATTCAACT
TCD422011       CTCCTGTATT ATCTTTGCCT GATGGATATT CTAAATATAG TGTAACTATT
TCD558851       GTGTTGGTGT TGGTGTTGGT GTTGGGGTTC TTGATGTTAG CATTTATGAA
TCD375103       CTCCTGTATT ATCTTTGCCT GATGGATATT CTAAATATAG TGTAACTATT
TCD428468       TTTATGAAAA AAATGGGCAG GTCCAAAACT ATACAGTGCC ATATTCAACT
TCD537014       TTAGCATTTA TGAAAAAAAT GGGCAGGTCC AAAACTATAC AGTGCCATAT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       210        220        230        240        250
TCD105056       GAGGTTA... .......... .......... .......... ..........
TCD279436       TTATCTTTGC CTGATGGATA TTCTAAATAT AGTGTAACTA TTGGTAGATA
TCD632685       AGTGTAACTA TTGGTAGATA CAGGGAGGTT A......... ..........
TCD678093       GATGGATATT CTAAATATAG TGTAACTATT GGTAGATACA GGGAGGTTA.
TCD139882       CCTGTATTAT CTTTGCCTGA TGGATATTCT AAATATAGTG TAACTATTGG
TCD422011       GGTAGATACA GGGAGGTTA. .......... .......... ..........
TCD558851       AAAAATGGGC AGGTCCAAAA CTATACAGTG CCATATTCAA CTCCTGTATT
TCD375103       GGTAGATACA GGGAGGTTA. .......... .......... ..........
TCD428468       CCTGTATTAT CTTTGCCTGA TGGATATTCT AAATATAGTG TAACTATTGG
TCD537014       TCAACTCCTG TATTATCTTT GCCTGATGGA TATTCTAAAT ATAGTGTAAC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        260        270        280        290        300
TCD105056       .......... .......... .......... .......... ..........
TCD279436       CAGGGAGGTT A......... .......... .......... ..........
TCD632685       .......... .......... .......... .......... ..........
TCD678093       .......... .......... .......... .......... ..........
TCD139882       TAGATACAGG GAGGTTA... .......... .......... ..........
TCD422011       .......... .......... .......... .......... ..........
TCD558851       ATCTTTGCCT GATGGATATT CTAAATATAG TGTAACTATT GGTAGATACA
TCD375103       .......... .......... .......... .......... ..........
TCD428468       TAGATACAGG GAGGTTA... .......... .......... ..........
TCD537014       TATTGGTAGA TACAGGGAGG TTA....... .......... ..........


                ....|....

TCD105056       .........
TCD279436       .........
TCD632685       .........
TCD678093       .........
TCD139882       .........
TCD422011       .........
TCD558851       GGGAGGTTA
TCD375103       .........
TCD428468       .........
TCD537014       .........
```

D.5 Sequence Alignment: VNTR locus Sal16

Multiple sequence alignment of the VNTR locus Sal16 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple and in light Blue indicate the tandem repeat regions.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TCD105056       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD179358       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD281699       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD717480       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TMI87513        TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD193597       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD279436       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD678877       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD421690       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT
TCD671445       TTGCAGTTAA TTTCTGCGAT CATTATTCTG AGCACCGCGT GCTGGATGTT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
TCD105056       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD179358       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD281699       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD717480       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TMI87513        CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD193597       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD279436       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD678877       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD421690       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC
TCD671445       CTGGCGGACA TGACGAGGCG AGCAGCAGTG GCTGGCGGGA AACCACCATC

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
TCD105056       ACGACCACCA TCACGACCAC CATTACGACC ATGACCATGA CCATGACCAT
TCD179358       ACGACCACCA TCACGACCAC CATTACGACC ATGACCATGA CCATGACCAT
TCD281699       ACGACCACCA TCACGACCAC CATCACGACC ACCATTACGA CCATGACCAT
TCD717480       ACGACCACCA TCACGACCAC CATTACGACC ATGACCATGA CCATGACCAT
TMI87513        ACGACCACCA TCACGACCAC CATCACGACC ACCATTACGA CCATGACCAT
TCD193597       ACGACCACCA TCACGACCAC CATCACGACC ACCATTACGA CCATGACCAT
TCD279436       ACGACCACCA TCACGACCAC CATCACGACC ACCATTACGA CCATGACCAT
TCD678877       ACGACCACCA TCACGACCAC CATTACGACC ATGACCATGA CCATGACCAT
TCD421690       ACGACCACCA TCACGACCAC CATTACGACC ATGACCATGA CCATGACCAT
TCD671445       ACGACCACCA TCACGACCAC CATCACGACC ACCATCACGA CCACCATCAC
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        160        170        180        190        200
TCD105056       GACCATGACC ATCATGGTCA CATACATCCG GAAGGCGCAA CGTCAAAAGC
TCD179358       GACCATGACC ATCATGGTCA CATACATCCG GAAGGCGCAA CGTCAAAAGC
TCD281699       GACCATGACC ATGACCATGA CCATGACCAT GACCATGACC ATCATGGTCA
TCD717480       GACCATGACC ATGACCATCA TGGTCACATA CATCCGGAAG GCGCAACGTC
TMI87513        GACCATGACC ATGACCATGA CCATGACCAT GACCATGACC ATGACCATGA
TCD193597       GACCATGACC ATGACCATGA CCATGACCAT GACCATGACC ATCATGGTCA
TCD279436       GACCATGACC ATGACCATGA CCATGACCAT CATGGTCACA TACATCCGGA
TCD678877       GACCATGACC ATGACCATGA CCATCATGGT CACATACATC CGGAAGGCGC
TCD421690       GACCATGACC ATCATGGTCA CATACATCCG GAAGGCGCAA CGTCAAAAGC
TCD671445       GACCACCATT ACGACCATGA CCATGACCAT GACCATGACC ATGACCATGA

                ....|....| ....|....| ....|....| ....|....| ....|....|
                        210        220        230        240        250
TCD105056       GTATCAA... .......... .......... .......... ..........
TCD179358       GTATCAA... .......... .......... .......... ..........
TCD281699       CATACATCCG GAAGGCGCAA CGTCAAAAGC GTATCAA... ..........
TCD717480       AAAAGCGTAT CAA..... .......... .......... ..........
TMI87513        CCATGACCAT CATGGTCACA TACATCCGGA AGGCGCAACG TCAAAAGCGT
TCD193597       CATACATCCG GAAGGCGCAA CGTCAAAAGC GTATCAA... ..........
TCD279436       AGGCGCAACG TCAAAAGCGT ATCAA..... .......... ..........
TCD678877       AACGTCAAAA GCGTATCAA. .......... .......... ..........
TCD421690       GTATCA.... .......... .......... .......... ..........
TCD671445       CCATGACCAT GACCATGACC ATCATGGTCA CATACATCCG GAAGGCGCAA

                ....|....| ....|..
                        260
TCD105056       .......... .......
TCD179358       .......... .......
TCD281699       .......... .......
TCD717480       .......... .......
TMI87513        ATCAA..... .......
TCD193597       .......... .......
TCD279436       .......... .......
TCD678877       .......... .......
TCD421690       .......... .......
TCD671445       CGTCAAAAGC GTATCAA
```

D.6 Sequence Alignment: VNTR locus TR4500

Multiple sequence alignment of the VNTR locus TR4500 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat regions. Sequences highlighted in light Blue indicate the 7bp insertion sequence identified in some *Salmonella* Typhi isolates.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                         10         20         30         40         50
TMI1647764      TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD537014       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD197403       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD185048       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD173225       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD678093       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD619382       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD409778       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TCD179358       TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT
TMI83959        TCGTTGCTGC TCCGAAATAT TCTGTTTGTA TACTGTTTTC CGCTTCGGTT


                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
TMI1647764      TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD537014       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD197403       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD185048       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD173225       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD678093       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD619382       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD409778       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TCD179358       TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
TMI83959        TTACCAGTCT CAACCTGGCT GTAAGTTCCC GCCGTTCCTT CTGGCCCATA
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
TMI1647764      TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD537014       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD197403       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD185048       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD173225       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD678093       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD619382       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD409778       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TCD179358       TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG
TMI83959        TTGTGGAGAT AATCCTGCAG TTCCTTCTCA TTCATGGTTT TAATATCGGG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       160        170        180        190        200
TMI1647764      TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TCD537014       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TCD197403       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TCD185048       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTAGAGT
TCD173225       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TCD678093       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTAGAGT
TCD619382       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTAGAGT
TCD409778       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TCD179358       TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTATTGC
TMI83959        TAAATCACCC CCTGATTTGT TCAGATTTTC AACAGGGGGA CAGTTAGAGT

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       210        220        230        240        250
TMI1647764      CACGAGTCCG AGTCCAAGGG GCGCAAGCGC CCTGGTCGGC TGCCGCCTCC
TCD537014       CACGAGTCCG AGTCCGAGTC CAAGGGGCGC AAGCGCCCTG GTCGGCTGCC
TCD197403       CACGAGTCCG AGTCCGAGTC CAAGGGGCGC AAGCGCCCTG GTCGGCTGCC
TCD185048       CCAAGGGGCG CAAGCGCCCT GGTCGGCTGC CGCCTCCTGA ACGTCAACGG
TCD173225       CACGAGTCCG AGTCCGAGTC CAAGGGGCGC AAGCGCCCTG GTCGGCTGCC
TCD678093       CCAAGGGGCG CAAGCGCCCT GGTCGGCTGC CGCCTCCTGA ACGTCAACGG
TCD619382       CCAAGGGGCG CAAGCGCCCT GGTCGGCTGC CGCCTCCTGA ACGTCAACGG
TCD409778       CACGAGTCCG AGTCCAAGGG GCGCAAGCGC CCTGGTCGGC TGCCGCCTCC
TCD179358       CACGAGTCCG AGTCCAAGGG GCGCAAGCGC CCTGGTCGGC TGCCGCCTCC
TMI83959        CCAAGGGGCG CAAGCGCCCT GGTCGGCTGC CGCCTCCTGA ACGTCAACGG
```

```
                ....|....| ....|....| ....|....| ....|....| ....|....
                        260        270        280        290
TMI1647764      TGAACGTCAA CGGCTTTACG AACCTTTTTC CACTTCACCG CAA......
TCD537014       GCCTCCTGAA CGTCAACGGC TTTACGAACC TTTTTCACTT CACCGCA..
TCD197403       GCCTCCTGAA CGTCAACGGC TTTACGAACC TTTTTCCACT TCACCGCAA
TCD185048       CTTTACGAAC CTTTTTCCAC TTCACCGCA. .......... .........
TCD173225       GCCTCCTGAA CGTCAACGGC TTTACGAACC TTTTTCCACT TCACCGCA.
TCD678093       CTTTACGAAC CTTTTTCCAC TTCACCGCA. .......... .........
TCD619382       CTTTACGAAC CTTTTTCCAC TTCACCGCA. .......... .........
TCD409778       TGAACGTCAA CGGCTTTACG AACCTTTTTC CACTTCACCG CA.......
TCD179358       TGAACGTCAA CGGCTTTACG AACCTTTTTC CACTTCACCG CA.......
TMI83959        CTTTACGAAC CTTTTTCCAC TTCACCGCA. .......... .........
```

D.7 Sequence Alignment: VNTR locus Sal20

Multiple sequence alignment of the VNTR locus Sal20 sequences from ten *Salmonella* Typhi isolates. Sequences highlighted in Yellow and in Green indicate the forward and reverse primers respectively. Sequences highlighted in Purple indicate the tandem repeat regions.

```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        10         20         30         40         50
TCD619382       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TMI83959        TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD105056       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD111848       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD179358       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD316823       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD110945       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD185048       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD298424       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
TCD678093       TCAGCCGACA CAACTTAACG AAGTGCCGTG GAACGAGCAG ACGCCAGAAC
```

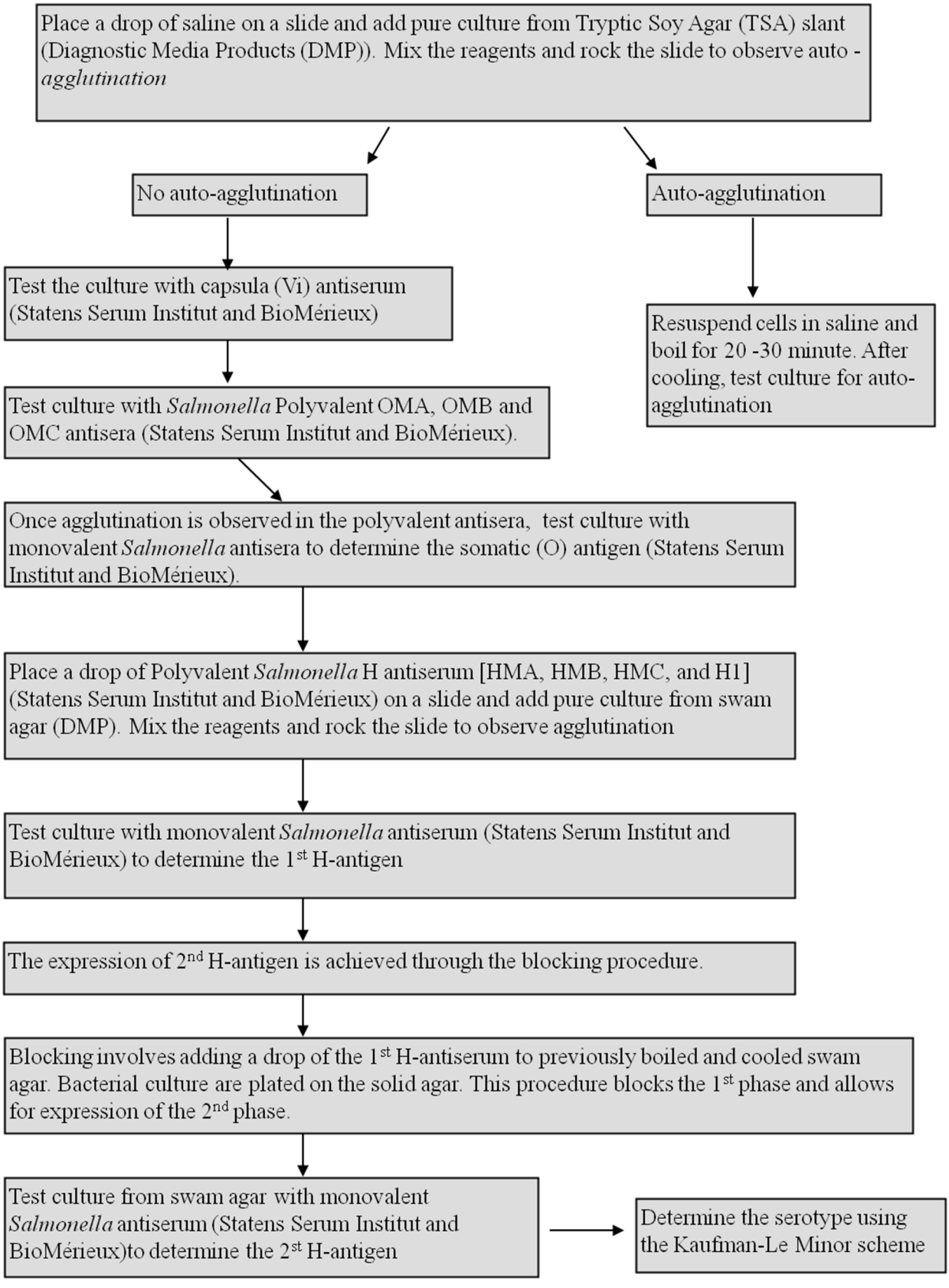
```
                ....|....| ....|....| ....|....| ....|....| ....|....|
                        60         70         80         90        100
TCD619382       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TMI83959        AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD105056       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD111848       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD179358       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD316823       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD110945       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD185048       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD298424       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG
TCD678093       AACGCCAGCA AACGCTACAG CGCCAGCGTC AGGCGCAGCA GCAGCAGCAG

                ....|....| ....|....| ....|....| ....|....| ....|....|
                       110        120        130        140        150
TCD619382       CAGCAGCAGC AGCAGCAGTG GGCTCAAACG CAGCCTGTTC AGCAACCACG
TMI83959        CAGCAGCAGC AGCAGCAGTG GGCTCAAACG CAGCCTGTTC AGCAACCACG
TCD105056       CAGCAGCAGT GGGCTCAAAC GCAGCCTGTT CAGCAACCAC GCACGCAGCC
TCD111848       CAGCAGCAGC AGCAGTGGGC TCAAACGCAG CCTGTTCAGC AACCACGCAC
TCD179358       CAGCAGCAGC AGCAGTGGGC TCAAACGCAG CCTGTTCAGC AACCACGCAC
TCD316823       CAGCAGCAGT GGGCTCAAAC GCAGCCTGTT CAGCAACCAC GCACGCAGCC
TCD110945       CAGCAGTGGG CTCAAACGCA GCCTGTTCAG CAACCACGCA CGCAGCCGCG
TCD185048       CAGCAGCAGC AGCAGCAGTG GGCTCAAACG CAGCCTGTTC AGCAACCACG
TCD298424       CAGCAGTGGG CTCAAACGCA GCCTGTTCAG CAACCACGCA CGCAGCCGCG
TCD678093       CAGCAGCAGC AGCAGCAGTG GGCTCAAACG CAGCCTGTTC AGCAACCACG


                ....|....| ....|....| ....|....| ....|....| ....|..
                       160        170        180        190
TCD619382       CACGCAGCCG CGGGTAAACG AACAGCCGCA AACGCGCACG GTACAGT
TMI83959        CACGCAGCCG CGGGTAAACG AACAGCCGCA AACGCGCACG GTACAGT
TCD105056       GCGGGTAAAC GAACAGCCGC AAACGCGCAC GGTACAGTA. .......
TCD111848       GCAGCCGCGG GTAAACGAAC AGCCGCAAAC GCGCACGGTA CAGT...
TCD179358       GCAGCCGCGG GTAAACGAAC AGCCGCAAAC GCGCACGGTA CAGT...
TCD316823       GCGGGTAAAC GAACAGCCGC AAACGCGCAC GGTACAGT.. .......
TCD110945       GGTAAACGAA CAGCCGCAAA CGCGCACGGT ACAGT..... .......
TCD185048       CACGCAGCCG CGGGTAAACG AACAGCCGCA AACGCGCACG GTACAGT
TCD298424       GGTAAACGAA CAGCCGCAAA CGCGCACGGT ACAGT..... .......
TCD678093       CACGCAGCCG CGGGTAAACG AACAGCCGCA AACGCGCACG GTACAGT
```

Appendix E

Salmonella serotyping flowchart



Appendix F

Agarose gel electrophoresis

F.1 10X TAE buffer (Tris: acetate: EDTA)

48 g Tris (Merck)

7.5 g EDTA (Sigma)

Dissolve in 500 ml distilled H2O

11 ml of glacial acetic acid (Merck)

Adjust volume to 1L using distilled H2O

F.2 1X TAE buffer

100 ml of 10X TAE buffer

900 ml of distilled dH20

F.3 Loading dye: 0.25% Bromophenol blue

0.25 g Bromophenol blue (Merck)

40 g Sucrose (Merck)

Dissolve in 100 ml distilled dH20

F.4 SeaKem LE Agarose gel (1.5%)

1.5 g SeaKem LE agarose

100 ml 1X TAE buffer

Dissolve agarose by boiling. Add 6 µl of ethidium bromide solution to the cooled agarose solution. Use 1 X TAE buffer as running buffer

Appendix G

MLVA profiles identified through MLVA typing of *Salmonella* Typhi from SSA

TR1	TR2	Sal02	Sal20	TR4699	MLVA profile*
176	303	201	169	187	STyMT-48
176	303	201	169	0	STyMT-49
176	303	201	0	187	STyMT-50
176	303	237	199	187	STyMT-51
253	383	303	178	223	STyMT-52
204	495	297	178	253	STyMT-54
183	335	201	184	259	STyMT-55
218	351	243	187	331	STyMT-56
183	383	237	184	223	STyMT-57
211	463	261	187	205	STyMT-58
197	399	315	190	241	STyMT-59
204	391	243	187	241	STyMT-60
190	455	243	187	247	STyMT-61
190	535	249	190	241	STyMT-62
197	567	231	190	229	STyMT-63
197	309	261	187	247	STyMT-64
197	655	255	190	259	STyMT-65
197	343	285	178	211	STyMT-66
197	367	273	190	241	STyMT-67
197	391	315	190	241	STyMT-68
197	399	219	187	241	STyMT-69
197	399	285	187	247	STyMT-70
197	479	303	190	241	STyMT-71
197	407	315	190	241	STyMT-72
197	415	315	190	235	STyMT-73
197	423	303	190	241	STyMT-74
197	423	315	190	235	STyMT-75
197	431	309	190	241	STyMT-76
197	447	315	190	247	STyMT-77
197	455	249	190	247	STyMT-78
197	471	219	187	241	STyMT-79
197	471	303	190	241	STyMT-80
197	471	315	190	235	STyMT-81
197	479	303	190	241	STyMT-82
197	487	213	193	247	STyMT-83
197	487	285	190	241	STyMT-84
197	487	309	196	229	STyMT-85

197	487	309	196	241	STyMT-86
197	495	255	190	247	STyMT-87
197	495	297	190	241	STyMT-88
197	503	225	187	241	STyMT-89
197	503	249	190	241	STyMT-90
197	503	255	190	247	STyMT-91
197	503	279	196	247	STyMT-92
197	503	285	190	241	STyMT-93
197	511	255	190	247	STyMT-94
197	511	255	190	259	STyMT-95
197	511	309	190	241	STyMT-96
197	519	249	190	241	STyMT-97
197	519	303	190	259	STyMT-98
197	527	209	190	247	STyMT-99
197	527	237	187	229	STyMT-100
197	527	243	187	235	STyMT-101
197	527	249	190	247	STyMT-102
197	527	255	190	247	STyMT-103
197	527	309	190	247	STyMT-104
197	527	315	190	247	STyMT-105
197	535	243	190	235	STyMT-106
197	543	249	187	253	STyMT-107
197	543	255	190	247	STyMT-108
197	551	249	190	247	STyMT-109
253	359	273	178	271	STyMT-110
197	551	255	190	247	STyMT-111
197	559	249	190	247	STyMT-112
197	567	243	190	253	STyMT-113
197	567	249	190	247	STyMT-114
197	567	249	190	253	STyMT-115
197	567	249	190	259	STyMT-116
197	567	255	190	247	STyMT-117
197	567	261	187	274	STyMT-118
197	567	315	190	235	STyMT-119
281	519	267	190	271	STyMT-120
197	575	237	190	229	STyMT-121
197	575	249	187	253	STyMT-122
197	575	249	190	259	STyMT-123
197	575	455	190	253	STyMT-124
197	583	237	190	229	STyMT-125
197	583	249	187	259	STyMT-126
197	583	249	190	241	STyMT-127
197	583	249	190	247	STyMT-128
197	583	249	190	253	STyMT-129

197	583	255	190	247	STyMT-130
197	583	367	190	241	STyMT-131
197	591	255	190	247	STyMT-132
197	591	255	190	253	STyMT-133
197	591	261	190	247	STyMT-134
197	599	249	190	241	STyMT-135
197	599	255	190	247	STyMT-136
197	599	261	190	247	STyMT-137
197	615	255	190	247	STyMT-139
197	650	321	190	241	STyMT-140
204	335	267	187	337	STyMT-141
204	391	315	178	301	STyMT-142
204	399	219	190	253	STyMT-143
204	431	315	190	253	STyMT-144
204	471	249	187	253	STyMT-145
204	583	249	190	253	STyMT-146
211	351	255	187	295	STyMT-147
211	359	255	187	289	STyMT-148
211	367	138	199	241	STyMT-149
211	375	219	196	241	STyMT-150
211	447	249	175	211	STyMT-151
211	463	261	181	211	STyMT-152
211	503	255	181	205	STyMT-153
211	511	249	187	211	STyMT-154
211	527	255	181	205	STyMT-155
218	343	237	187	211	STyMT-156
218	343	249	187	319	STyMT-157
295	311	249	187	331	STyMT-158
218	351	249	187	331	STyMT-159
218	359	249	187	319	STyMT-160
218	375	243	187	247	STyMT-161
218	455	237	187	211	STyMT-162
218	463	255	187	247	STyMT-163
218	479	255	184	241	STyMT-164
218	479	255	187	241	STyMT-165
218	511	273	187	307	STyMT-166
218	511	309	178	267	STyMT-167
225	315	249	184	289	STyMT-168
225	359	273	187	247	STyMT-169
225	375	0	178	223	STyMT-170
225	399	175	178	223	STyMT-171
225	423	267	184	229	STyMT-172
225	463	225	184	241	STyMT-173
225	463	261	199	271	STyMT-174

225	503	273	178	295	STyMT-175
225	543	231	199	223	STyMT-176
225	543	249	187	223	STyMT-177
232	335	261	184	259	STyMT-178
232	335	261	187	232	STyMT-179
232	335	261	187	259	STyMT-180
232	335	263	187	187	STyMT-181
232	351	303	178	229	STyMT-182
232	359	219	181	313	STyMT-183
232	375	163	184	241	STyMT-184
232	399	583	193	205	STyMT-185
232	407	273	184	295	STyMT-186
232	415	151	184	229	STyMT-187
232	423	231	187	223	STyMT-188
232	423	261	184	223	STyMT-189
232	455	219	178	247	STyMT-190
232	455	261	184	247	STyMT-191
232	487	261	178	229	STyMT-192
232	503	201	184	259	STyMT-193
232	503	261	187	253	STyMT-194
232	543	255	184	247	STyMT-195
232	551	0	184	247	STyMT-196
232	551	267	184	241	STyMT-197
232	559	261	184	241	STyMT-198
232	559	267	184	241	STyMT-199
232	575	261	187	223	STyMT-200
232	650	261	184	247	STyMT-201
239	359	285	181	265	STyMT-202
239	367	279	181	265	STyMT-203
239	375	285	187	289	STyMT-204
239	383	0	184	247	STyMT-205
239	383	168	184	253	STyMT-206
239	383	267	181	253	STyMT-207
239	391	163	184	253	STyMT-208
239	391	267	181	253	STyMT-209
239	399	213	187	247	STyMT-210
239	399	279	178	283	STyMT-211
239	431	285	178	229	STyMT-213
239	439	267	178	235	STyMT-214
239	455	0	184	259	STyMT-215
239	455	237	187	271	STyMT-216
239	455	291	178	211	STyMT-217
239	463	291	190	235	STyMT-218
239	503	285	178	229	STyMT-219

239	519	267	184	235	STyMT-220
239	519	273	178	253	STyMT-221
239	567	261	184	229	STyMT-222
239	575	261	184	253	STyMT-223
246	415	249	178	241	STyMT-224
239	583	273	187	247	STyMT-225
246	359	303	178	277	STyMT-227
246	359	309	178	271	STyMT-228
246	367	273	178	229	STyMT-229
246	367	303	178	271	STyMT-230
246	367	309	178	265	STyMT-231
246	367	315	178	271	STyMT-232
246	375	303	178	295	STyMT-233
246	383	273	178	295	STyMT-234
246	383	527	190	217	STyMT-235
246	447	261	184	241	STyMT-236
246	471	291	178	241	STyMT-237
246	535	0	184	223	STyMT-238
253	311	267	187	337	STyMT-239
253	383	309	178	271	STyMT-240
253	463	297	178	241	STyMT-241
253	503	255	184	241	STyMT-242
255	479	249	187	223	STyMT-243
260	327	267	175	223	STyMT-244
260	375	273	178	217	STyMT-245
260	415	237	178	253	STyMT-246
260	423	273	178	255	STyMT-247
267	327	186	175	259	STyMT-248
267	343	279	187	349	STyMT-249
267	375	219	178	199	STyMT-250
271	479	297	196	235	STyMT-251
281	431	219	178	223	STyMT-252
281	463	249	178	271	STyMT-253
197	471	315	196	241	STyMT-254
197	487	303	190	241	STyMT-255
197	487	309	193	241	STyMT-256
197	503	321	187	229	STyMT-257
253	311	267	184	253	STyMT-258
197	543	351	187	289	STyMT-259
197	567	249	190	241	STyMT-260
197	583	249	187	241	STyMT-261
218	343	249	187	295	STyMT-262
225	319	249	187	225	STyMT-263
225	319	249	187	289	STyMT-264

225	511	249	184	235	STyMT-265
232	351	303	178	289	STyMT-266
246	359	309	178	277	STyMT-267
246	359	321	178	211	STyMT-268
246	383	309	178	271	STyMT-269
197	261	0	190	253	STyMT-270
197	309	303	190	235	STyMT-271
197	319	249	187	253	STyMT-272
197	343	243	190	259	STyMT-273
176	303	201	169	253	STyMT-274
239	519	0	184	253	STyMT-275
239	543	0	184	253	STyMT-276
197	591	249	190	247	STyMT-277

* STyMT - *Salmonella* Typhi MLVA type

Appendix H

Ethics clearance certificate

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG
Division of the Deputy Registrar (Research)

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
R14/49 Ms Nomsa P Tau

CLEARANCE CERTIFICATE	**M120951**
PROJECT	Development and Evaluation of New Molecular Epidemiological Methods for Analysis of Salmonella Typhi
INVESTIGATORS	Ms Nomsa P Tau.
DEPARTMENT	Centre for Enteric Disease
DATE CONSIDERED	28/09/2012
DECISION OF THE COMMITTEE*	Approved unconditionally

Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.

DATE 28/09/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Dr Karen Keddy

DECLARATION OF INVESTIGATOR(S)
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

Appendix I

Plagiarism declaration with 'Turnitin' report

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

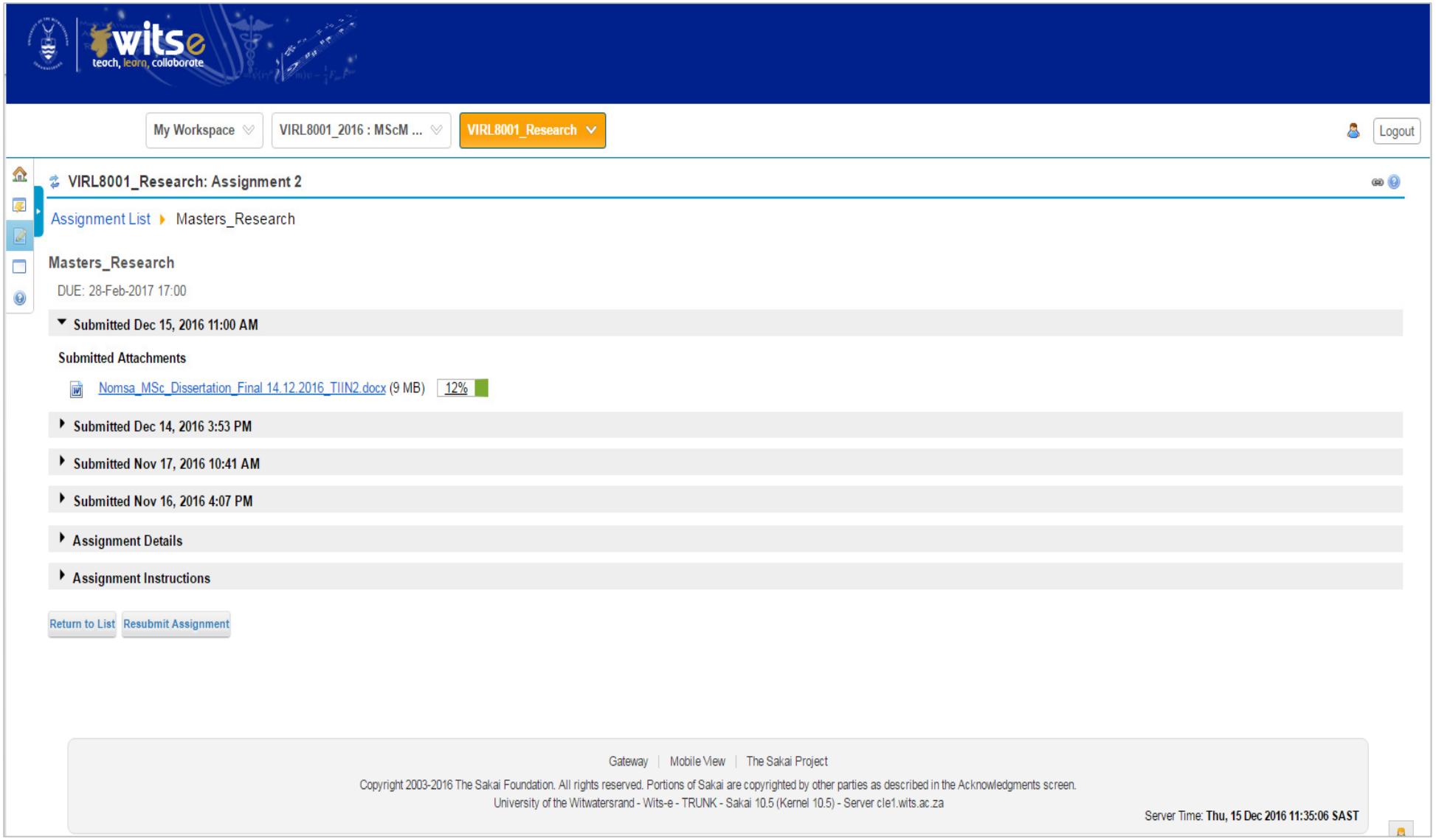
SENATE PLAGIARISM POLICY: APPENDIX ONE

I Nomsa Tau (Student number: 693689) am a student registered for the degree of Master of Science in the academic year 2012.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: Tau Date: 15-Dec-2016


witse
teach, learn, collaborate
My Workspace
VIRL8001_2016 : MScM ...
VIRL8001_Research
Logout
VIRL8001_Research: Assignment 2
Assignment List
Masters_Research
Masters_Research
DUE: 28-Feb-2017 17:00
Submitted Dec 15, 2016 11:00 AM
Submitted Attachments
Nomsa_MSc_Dissertation_Final 14.12.2016_TIIN2.docx (9 MB)
12%
Submitted Dec 14, 2016 3:53 PM
Submitted Nov 17, 2016 10:41 AM
Submitted Nov 16, 2016 4:07 PM
Assignment Details
Assignment Instructions
Return to List
Resubmit Assignment
Gateway
Mobile View
The Sakai Project
Copyright 2003-2016 The Sakai Foundation. All rights reserved. Portions of Sakai are copyrighted by other parties as described in the Acknowledgments screen.
University of the Witwatersrand - Wits-e - TRUNK - Sakai 10.5 (Kernel 10.5) - Server cle1.wits.ac.za
Server Time: Thu, 15 Dec 2016 11:35:06 SAST