
UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG



**Snapshot of South African *Salmonella*:
Development of a sensitive tool to investigate
the distribution of *Salmonella* Enteritidis
clades in 2012 and 2013**

Sarah Gallichan

(Student number: 1726349)

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, in fulfillment of the requirements for the degree of Master of Science in Medicine

Johannesburg, 2020

Advisory Committee

Supervisor 1

Doctor Anthony Marius Smith

BSc(Hons), PhD

Supervisor 2

Doctor Juno Thomas

MBBCh, DTMH, DipHIVMan, FCPATH

Department

Centre for Enteric Diseases at the National Institute for Communicable Diseases (NICD),
Johannesburg, South Africa AND Department of Clinical Microbiology and Infectious
Diseases, School of Pathology, Faculty of Health Sciences, University of the
Witwatersrand, Johannesburg, South Africa

Supervisor 3

Professor Nicholas Feasey

BSc, MSc, PhD, MBBS, MRCP, FRCPath, DTMH

Department

Liverpool School of Tropical Medicine, Pembroke Place, Liverpool, UK AND Malawi
Liverpool Wellcome Clinical Research Programme, Blantyre, Malawi

Declaration

The experimental work described in this dissertation was conducted under the supervision of **Dr Anthony M. Smith** and **Dr Juno Thomas** at the Centre for Enteric Diseases (CED), National Institute for Communicable Diseases (NICD), National Health Laboratory Service (NHLS).

I **Sarah Gallichan**, student number **1726349** declare that this dissertation is my own, unaided work. It is being submitted for the fulfilment for the degree of Master of Science in Medicine to the Faculty of Health Sciences, University of Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.



Sarah Gallichan

08/06/2020

Date

Dedication

To my parents who have endured my endless scientific monologues, for your love and dedication to my success I cannot thank you enough.

To Ross Gray, my other half, for all your support and kindness through this stressful process. You always bring a smile to my face, hopefully this will bring a smile to yours.

Presentations

Gallichan, S. Development of a real-time PCR assay to characterise clades of *Salmonella* Enteritidis. Presented at the Federation of Infectious Diseases Societies of Southern Africa (FIDSSA), Indaba Hotel & Conference Centre, Fourways, Johannesburg, South Africa, 8 November 2019. (Oral presentation)

Gallichan, S. Development of a real-time PCR assay to characterise clades of *Salmonella* Enteritidis. Presented at the Scientific Research Forum, National Institute for Communicable Diseases, Johannesburg, South Africa, 27 November 2019. (Oral presentation)

Gallichan, S. Development of a real-time PCR assay to characterise clades of *Salmonella* Enteritidis. Presented at the Molecular Biosciences Research Thrust (MBRT) Research Day, Wits Adler Museum, University of the Witwatersrand Medical Campus, Johannesburg, South Africa 28 November 2019. (Poster)

Papers in preparation

S Gallichan, N Feasey, N Ramalwa, J Thomas, AM Smith Investigation of the prevalence and impact of *Salmonella* Enteritidis clades in South Africa. Viewpoint paper to be published in PLOS Neglected Tropical Diseases.

S Gallichan, N Feasey, J Thomas, AM Smith Development of a real-time PCR assay to characterize clades of *Salmonella* Enteritidis. Full research paper to be published in the Journal of Clinical Microbiology.

Abstract

Salmonella Enteritidis is the most commonly reported non-typhoidal *Salmonella* serovar causing human disease, and is implicated in both gastroenteritis and invasive non-typhoidal *Salmonella* disease worldwide. Whole-genome sequence (WGS) comparison of *Salmonella* Enteritidis isolates from across the world identified four definitive clades: Outlier, Global, East African and West African. Three of these clades were implicated in epidemics: the Global Clade was linked to poultry-associated gastroenteritis, while the two African clades were related to invasive disease. Despite the recognition of different *Salmonella* Enteritidis clades, the distribution and epidemiology of these clades across Africa is poorly understood. To investigate this, we developed a sensitive real-time PCR assay that was capable of differentiating between the *Salmonella* Enteritidis clades. This assay was then used to investigate 618 *Salmonella* Enteritidis isolates originating from four South African provinces over a period of two years (2012 – 2013). Associations between particular clades and province, antimicrobial resistance and HIV status were explored using statistical analysis. Majority of the South African isolates were classified within the Outlier clade (61.00%), with fewer classified within the Global clade (38.83%) and only one isolate classified within the West African Clade (0.16%). Of note, the isolates within the Global clade were strongly associated with decreased ciprofloxacin susceptibility (37/240, 15.42%, OR: 7.45, CI: 3.526 – 15.751) and invasive disease (140/240, 58.33%, OR: 1.57, CI: 1.13 – 2.17). Thus, we have highlighted the importance of monitoring *Salmonella* Enteritidis clades in sub-Saharan Africa in order to gain a better understanding of the distribution and epidemiology. The development of a sensitive real-time PCR assay provides a cost-effective alternative to WGS for use in resource-limited settings.

Acknowledgements

Doctor Anthony Smith and Doctor Juno Thomas, thank you for your skilled guidance during these past two years. I greatly appreciate you both for always being available for questions and discussions about my project at all times.

Professor Nicholas Feasey, I cannot thank you enough for the invaluable assistance you have provided me. Your knowledge and advice was instrumental to the successful completion of this dissertation. Through your willingness to share and advise me, I have learnt so much about the fields of public health and epidemiology.

Ntsieni Ramalwa, for keeping me motivated throughout the writing process and for your assistance with all the statistics and epidemiology required by this study. Your smiles and kind words have helped me immensely during this process, thank you for each and every one of them.

Doctor Blanca Perez-Sepulveda and Professor Jay Hinton, for sending me all the validation panel isolates I required as well as for tirelessly answering all the questions I had. Thank you for making me feel very welcome during my time at the University of Liverpool, I felt privileged to share in your immense knowledge.

To all the staff members at the National Institute of Communicable Diseases who assisted me, I greatly appreciate all your help. A special thanks to Masindi Ramudzulu and Dominic Ndlovu for assisting in the tedious task of finding all the necessary isolates and culturing them. Sue Candy, for assisting in the acquisition of additional data from the Central Data Warehouse. Nevashan Govender for the helpful epidemiology discussions.

Lastly, but most certainly not least, I would like to thank my family. To my parents, Anna and Grant Gallichan, there aren't enough words to completely express how grateful I am for you both. It has been a long, stressful road and you both have patiently supported me throughout. Thank you for helping make my dreams of being a scientist possible. To my brother, Bryce Gallichan, for always taking interest in my project and adding helpful thoughts to each discussion. To my boyfriend, Ross Gray, for being with me through every part of this process. I would not have managed without your loving support and encouragement.

This study was funded by the German Federal Ministry of Education and Research
(Grant Number 81203616)

Table of Contents

Advisory Committee	i
Declaration.....	ii
Dedication.....	iii
Presentations.....	iv
Abstract	v
Acknowledgements	vi
Table of Contents.....	vii
Nomenclature	x
List of Figures.....	xii
List of Tables	xvi
CHAPTER ONE.....	1
Literature Review	1
1.1. <i>Salmonella</i> through the years	1
1.1.1. Historical outbreaks.....	1
1.1.2. Classifying <i>Salmonella</i>	2
1.1.3. Current taxonomy of <i>Salmonella</i>	3
1.2. Non-typhoidal <i>Salmonella</i>	4
1.2.1. Characteristics of non-typhoidal <i>Salmonella</i>	4
1.2.2. Method of host cell invasion.....	5
1.2.3. Symptoms of non-typhoidal <i>Salmonella</i> infection.....	5
1.2.4. Transmission of non-typhoidal <i>Salmonella</i>	6
1.2.5. Risk factors associated with non-typhoidal <i>Salmonella</i>	7
1.3. Typing <i>Salmonella</i>	8
1.3.1. Phenotypic laboratory identification of non-typhoidal <i>Salmonella</i>	8
1.3.2. Molecular subtyping and characterisation of non-typhoidal <i>Salmonella</i>	10
1.4. <i>Salmonella enterica</i> serovar Enteritidis.....	14
1.4.1. Disease surveillance	14
1.4.2. Global <i>Salmonella</i> Enteritidis burden.....	15
1.4.3. The burden of <i>Salmonella</i> Enteritidis in Africa	16
1.4.4. Preventive measures for <i>Salmonella</i> Enteritidis	17

1.5. Genomic composition of <i>Salmonella</i> Enteritidis	19
1.5.1. Genome characteristics	19
1.5.2. Evolution of a genome	19
1.5.3. Population structure	20
1.6. Study Objectives.....	23
CHAPTER TWO.....	24
Development of a clade-classifying real-time PCR assay for <i>Salmonella</i> Enteritidis.....	24
2.1. Introduction	24
2.2. Materials and methods	26
2.2.1. Workflow for the development of a clade-classifying real-time PCR assay.....	26
2.2.2. Research facility	27
2.2.3. Real-time PCR assay validation panel isolates	27
2.2.4. Additional samples for validation panel	28
2.2.5. Whole-genome sequencing	28
2.2.6. Selection of category- and clade-specific genes	29
2.2.7. Gene selection from whole-genome sequencing data	29
2.2.8. Primer and probe design for real-time PCR assay.....	30
2.2.9. <i>In silico</i> PCR assay	31
2.2.10. Preparation of primers and probes	31
2.2.11. Crude genomic DNA extraction of validation panel isolates	31
2.2.12. Validation of real-time PCR assay.....	32
2.2.13. Optimisation of primer concentrations for singleplex real-time PCR assay.....	34
2.2.14. Optimisation of probe concentrations for singleplex real-time PCR assay.....	36
2.2.15. Master mixes for multiplex real-time PCR assays.....	37
2.2.16. Pure whole-genome DNA extraction from isolates using kits	38
2.2.17. Analytical verification of the multiplex real-time PCR assay	38
2.3. Results.....	40
2.3.1. Whole-Genome Sequence analysis of the validation panel isolates	40
2.3.2. Selection of category- and clade-specific genes	40
2.3.3. <i>In silico</i> primer and probe validation	41
2.3.4. Sensitivity and specificity of the real-time PCR assay	41
2.3.5. Primer and probe concentration optimisation	42
2.3.6. Analytical verification of primers and probe sets within multiplex assays.....	42

2.4. Discussion	46
CHAPTER THREE.....	48
Epidemiological investigation of <i>Salmonella</i> Enteritidis clades in South Africa.....	48
3.1. Introduction	48
3.2. Materials and methods	50
3.2.1. Workflow for the epidemiological investigation of <i>Salmonella</i> Enteritidis clades in South Africa	50
3.2.2. GERMS-SA laboratory surveillance program	51
3.2.3. GERMS-SA non-typhoidal <i>Salmonella</i> isolates repository.....	51
3.2.4. Isolate referral bias	52
3.2.5. Selection of South African <i>Salmonella</i> Enteritidis isolates	53
3.2.6. Crude genomic DNA extraction	54
3.2.7. Classification of isolates into clades using clade-classifying real-time PCR assay...54	
3.2.8. Line list of epidemiological data variables for <i>Salmonella</i> Enteritidis isolates	54
3.2.9. Bivariate statistical analysis of epidemiological data.....	55
3.2.10. Bivariate statistical analysis of HIV data subset	55
3.2.11. Ethics statement.....	56
3.3. Results.....	57
3.3.1. Description of the study cohort	57
3.3.2. Clade classification of <i>Salmonella</i> Enteritidis isolates across South Africa.....	58
3.3.3. <i>Salmonella</i> Enteritidis isolates associated with invasive non-typhoidal <i>Salmonella</i> disease in patients 5 years and older	59
3.3.4. Antimicrobial resistance in <i>Salmonella</i> Enteritidis isolates.....	61
3.4. Discussion	63
CHAPTER FOUR.....	68
Conclusions.....	68
4.1. Summary of work.....	68
4.2. Limitations.....	68
4.3. Implication of this work	70
CHAPTER FIVE.....	71
References.....	71
SUPPLEMENTARY MATERIAL	71
Plagiarism Report	97

Nomenclature

- **μL**- Microliter
- **μM**- Micro molar
- **BAPS**- Bayesian Analysis of Population Structure
- **US CDC**- United States Centers for Disease Control and Prevention
- **CDW**- Central Data Warehouse
- **CED**- Centre for Enteric Diseases
- **CI**- Confidence Interval
- **Ct**- Cycle threshold
- **DALY**- Disability-adjusted life years
- **DCS** – Decreased ciprofloxacin susceptibility
- **DNA**- Deoxyribonucleic acid
- **EFSA**- European Food Safety Authority
- **ERIC**- Enterobacterial repetitive intergenic consensus
- **EU**- European Union
- **CLSI**- Clinical and Laboratory Standards Institute
- **FQ**- Fluoroquinolone
- **GERMS-SA**- Group for Enteric, Respiratory and Meningitides Surveillance for South Africa
- **HIV**- Human immunodeficiency virus
- **iNTS**- Invasive non-typhoidal *Salmonella*
- **kb**- Kilobase pairs
- **LMIC**- low- and middle-income countries
- **MIC**- Minimum inhibitory concentration
- **MDR**- Multidrug-resistant
- **MLST**- Multi-locus sequence typing
- **MLVA**- Multi-locus variable number tandem repeat analysis
- **N**- total number of bacterial isolates in validation panel
- **n**- number of bacterial isolates
- **ng**- Nano gram
- **NICD**- National Institute for Communicable Diseases
- **nm**- Nano meter
- **NTS**- Non-typhoidal *Salmonella*

- **OR**- Odds ratio
- **PCR**- Polymerase chain reaction
- **PFGE**- Pulsed-field gel electrophoresis
- **R**- Determination coefficient
- **RAPD**- Random amplified polymorphic DNA
- **RNA**- Ribonucleic acid
- **rRNA**- Ribosomal ribonucleic acid
- **SPI**- *Salmonella* pathogenicity island
- **sSA**- Sub-Saharan Africa
- **US**- United States
- **VNTR**- Variable number tandem repeat
- **WGS**- Whole-genome sequencing

List of Figures

Note on Figure and Table labelling: Within this dissertation, Figures and Tables are labelled using three numbers. The first number indicates the chapter and the second number indicates the section wherein the Figure or Table is located. The third number simply indicates the number of the Table or Figure within the given chapter and section.

CHAPTER ONE:

- Figure 1.1.1** An unrooted phylogenetic tree representing the relationship between the five Proteobacteria classes (α , β , δ , ϵ and γ) and the bacterial genera of medical importance within each class. 3
- Figure 1.2.1** Electron micrograph of *Salmonella enterica* serovar Enteritidis. Flagella, extruding from the bacterial cell wall, assist with locomotion and host cell attachment.....4
- Figure 1.2.2** Illustration of gut epithelial cells situated closely together in the small intestine. The brush border and junctional complex are cell characteristics that make the intracellular invasion of *Salmonella* bacteria possible. 5
- Figure 1.2.3** A bar graph representing the percentage of laboratory-confirmed *Salmonella* Enteritidis isolates taken from non-human sources from 1968 – 2011 in the United States. Chicken is significantly higher than any other source.....6
- Figure 1.3.1** Visible agglutination (left) is observed as white dots in the solution as antigen-antibody complexes form. 8
- Figure 1.3.2** Diagram representing the phage typing of a *Salmonella enterica* serovar. The formation of plaques (dark circles) indicate the bacterial susceptibility to a particular bacteriophage.....9
- Figure 1.3.3** A Pulsed-Field Gel Electrophoresis image of *Salmonella enterica* serotypes. The patterns produced by the fragmented DNA make isolates comparable which is beneficial for outbreak detection. 10

Figure 1.3.4 An electropherogram produced by the MLVA procedure. The red peaks indicate DNA standards of a known size and the green, blue and black peaks represent the sizes of PCR assay products from a bacterial isolate. A software program converts this data into a serotype.....	11
Figure 1.4.1 Disability-adjusted Life Years (DALYs) due to invasive non-typhoidal <i>Salmonella</i> based on estimates calculated by the World Health Organization.....	16
Figure 1.4.2 Number of confirmed non-typhoidal <i>Salmonella</i> isolates (invasive and non-invasive) of different serovars submitted to the South African GERMS program over a 12 year period. Adapted from GERMS-SA database by the author, Sarah Gallichan.....	17
Figure 1.4.3 Number of confirmed non-typhoidal <i>Salmonella</i> isolates of different serovars in the United States over a 15 year period.....	18
Figure 1.5.1 A cladogram representing the evolutionary relationship between the four <i>Salmonella</i> Enteritidis clades. Diagram created by the author, Sarah Gallichan.....	20
Figure 1.5.2 Diagrammatic representation of the differences between the accessory genomes of <i>Salmonella</i> Enteritidis clades.....	21

CHAPTER TWO:

Figure 2.2.1 Flow chart showing the workflow for the identification of <i>Salmonella</i> Enteritidis genes for use as clade molecular markers and the development and validation of a clade-classifying real-time PCR assay.....	26
Figure 2.2.2 96-well plate layout for optimal primer concentrations to amplify clade-defining genes in <i>Salmonella</i> Enteritidis isolates in a real-time PCR assay	34
Figure 2.2.3 96-well plate layout for optimal probe concentrations to amplify clade-defining genes in <i>Salmonella</i> Enteritidis isolates in a real-time PCR assay.....	36
Figure 2.2.4 Serial dilutions of pure DNA used to test the efficiency of the primers and probes in the clade-classifying PCR assay.....	38
Figure 2.3.1 Workflow depicting the clade classification of a <i>Salmonella</i> Enteritidis isolate based on the presence (green tick) or absence (red cross) of genes targeted by the clade-classifying real-time PCR assay. Diagram created by the author, Sarah Gallichan	39

Figure 2.3.2 Logarithmic efficiency plot depicting the change in cycle threshold value with the change in DNA concentration for the African and Global/ Outlier category primer and probe sets within the category-specific real-time PCR assay.....42

Figure 2.3.3 Logarithmic efficiency plot depicting the change in cycle threshold value with the change in DNA concentration for the East African and Global clade primer and probe sets within the clade-specific real-time PCR assay.....42

CHAPTER THREE:

Figure 3.2.1 Flow chart showing the workflow for the statistical analysis of clade-classifying assay results and corresponding patient and isolate data.....49

Figure 3.2.2 A diagram representing the proportion of potential pathogens present at different stages. Different types of surveillance programs sample pathogens at different stages, therefore have varying sensitivities and specificities. The blue area of the pyramid represents the proportion of isolates reported to surveillance organizations as with GERMS-SA. Adapted from Crump *et al.* 2003.....51

Figure 3.3.1 Frequency (grey bars) of the total *Salmonella* Enteritidis isolates, invasive isolates (red line) and non-invasive isolates (blue line) collected from patients in each age group.....56

Figure 3.3.2 Graphical representation of the proportions of *Salmonella* Enteritidis isolates classified within the Outlier, Global and Western African Clades in four South African provinces (Gauteng, Mpumalanga, KwaZulu-Natal and Western Cape).....57

Figure 3.3.3 The number of invasive disease cases with the corresponding number of HIV positive cases across all age groups 5 years and above. The asterisks (*) indicate the age groups where there was evidence for an association between invasive disease and HIV status.....58

Figure 3.3.4 Proportion of invasive **(A)** and non-invasive **(B)** *Salmonella* Enteritidis isolates classified within each of the two clades (Outlier and Global Clade).....59

Figure 3.3.5 The proportion of *Salmonella* Enteritidis isolates that are resistant and susceptible to various antimicrobial compounds. (SUL= sulphamethoxazole; CIP= ciprofloxacin; NAL= nalidixic acid; TET= tetracycline; TRI= trimethoprim; AMP= ampicillin; CHL= chloramphenicol; COT= cotrimoxazole; AUG= amoxicillin-clavulanate; CAS= ceftazadime; CRO= ceftiaxone; STR= streptomycin; KAN= kanamycin; IMI= imipenem).....60

Figure 3.4.1 Proportion of invasive *Salmonella* isolates, by age group, from high-income and low-to middle-income settings in 2017. 62

Figure 3.4.2 Consumption per capita of four different meat types (kg) in South Africa from 2000 to 2013. The two years (2012 and 2013) covered in this study are highlighted in the red box.....63

List of Tables

CHAPTER TWO:

Table 2.2.1 Validation panel isolates sent from the University of Liverpool.....	27
Table 2.2.2 Additional validation panel isolates from GERMS-SA surveillance repository.....	28
Table 2.2.3 Classification of isolates into clades according to real-time PCR assay results.....	29
Table 2.2.4 Optimal criteria for primer and probe characteristics.....	30
Table 2.2.5 Selected Assay Sets for primers and probes (including the fluorophores and quenchers) of target genes.....	30
Table 2.2.6 Set up of validation of primers and probes with validation panel isolates in a 96-well plate.....	32
Table 2.2.7 7500 Real-time PCR System run conditions.....	33
Table 2.2.8 Volumes of 20 µM primer Stock Solution and nuclease-free water required for particular primer concentrations to make up different concentrations of primer	34
Table 2.2.9 Volumes of 20 µM probe Stock Solution and nuclease-free water required to make up different concentrations of probe for particular probe concentrations	35
Table 2.2.10 Constituents of the four Master Mixes used in the category-classifying (Master Mix One and Two) clade-classifying (Master Mix Three and Four) real-time PCR.....	36
Table 2.2.11 Volume of water and DNA required for each validation panel isolate to obtain a final DNA concentration of 100 µM in a volume of 50 µl.....	37
Table 2.3.1 Average Ct values from the three technical replicate real-time PCR assays performed using the validation panel isolates. Target gene Ct values shaded in green indicate a positive result (Ct<30) and values in red indicate a negative result.....	40
Table 2.3.2 Ct values from the three technical replicate real-time PCR assays performed with varying DNA concentrations. Values shaded in green indicate a positive result (Ct<30) and values in red indicate a negative result (Ct≥30).....	44

CHAPTER THREE:

Table 3.2.1 Number of invasive and non-invasive *Salmonella* Enteritidis isolates selected from the 2012 and 2013 GERMS-SA repository to represent four South African provinces (Gauteng, Western Cape, KwaZulu-Natal, Mpumalanga).....52

Table 3.3.1 Association between invasive *Salmonella* Enteritidis isolates and South African provinces from which the *Salmonella* Enteritidis isolates were collected.....59

Table 3.3.2 Association between antimicrobial resistance of *Salmonella* Enteritidis isolates and the South African province from which the isolates were collected.....61

Table 3.3.3 Association between antimicrobial resistance and Global Clade isolates.....61

Glossary

- **Actin-** a protein component of the muscle responsible for the contraction and relaxation
- **Anthropo-notic disease-** an infectious disease which is carried and spread by humans
- **Antibody-** a molecule, produced as a result of an immune response, that specifically combines with an antigen in order to neutralise it
- **Antigen-** a substance capable of eliciting an immune response
- **Bacteraemia-** the presence of bacteria in the blood, which is usually a sterile environment
- **Bacteriophage-** a virus which infects and replicates within bacteria
- **Bar-coding-** (in relation to Whole-genome sequencing) the addition of unique sequences to DNA fragments for the sorting and identification of sequences
- **Base pairs-** a complementary pair of nucleic acids, which make up the double stranded structure of deoxyribonucleic acid. The base pair is often used as a metric for the size of a genome or genetic material
- **Biochemical-** chemical reactions with relation to bacteria
- **CD18-** (as in CD18-expressing) a protein which functions in cell surface signalling and immune responses
- **Decreased susceptibility-** bacteria with an MIC that is higher than expected for the susceptible population and has one or more resistance mechanisms
- **Denuding-** the removal of an overlying material
- **Disability-Adjusted Life Years-** the number of potential years of a person's life lost due to the lack of productivity or mortality resulting from a disease, as defined by the World Health Organization
- **Endonuclease-** a restriction enzyme that cleaves the phosphodiester bonds within a DNA molecule
- **Epidemiology-** the study of factors which determine or influence the occurrence and distribution of health-related events
- **Epithelium-** a cellular layer covering internal and external surfaces of an multi-cellular organism
- **Fimbriae-** hair-like extensions branching from the cell surface
- **Genomic degradation-** the reduction in genetic material due to deleterious mutations,
- **Homogenous-** being alike or having the same qualities

- **Homoplasmy-** a characteristic or genetic element shared by two or more bacteria as a result of a common ancestor
- **Immunogenic-** relating to the elicitation of an immune response
- **Immunosuppressed-** a state of the immune system whereby the function is greatly decreased as a result of a disease or medication
- **Incident-** number of times an event has occurred
- **Invasive-** (as in invasive salmonellosis) the ability of a *Salmonella* bacterium to invade the bloodstream of the human host and subsequently cause systemic disease
- **Junctional complexes-** a specialized site of attachment between cells or to the extracellular matrix
- **Lineage-** a line of descendants from a common ancestor
- **Low-level resistance-** bacteria which has an MIC that is higher than expected for the susceptible population but lacks any resistance mechanisms
- **Lysis-** the disruption of cellular structure leading to the destruction of a bacterial cell or tissue
- **Meningitis-** the inflammation of the membranes surrounding the brain and spinal cord (meninges) as a result of an infection
- **Metagenomics-** the study of a range of nucleic acids extracted directly from the environmental sample
- **Molecular/ Nucleic-acid based methodology-** the use of genetic material in various procedures to classify and investigate bacteria
- **Mortality rate-** the number of deaths in a defined population occurring at a given time period
- **Multidrug-resistant-** resistance to three or more antibiotics
- **Niche-** a particular place which is defined by specific biotic and abiotic factors
- **Nomenclature-** the designation of a name
- **Operon-** closely linked genes which are grouped together
- **Osteomyelitis-** an infection of bone tissue
- **Outbreak-** a sudden increase in the number of cases of a disease above what is normally expected in that population in that area
- **Paraphyletic-** a taxonomic group that does not include all descendants of a common ancestor
- **Pathogenicity-** bacteria's ability to cause disease

- **Permease**- a category of proteins functioning in the transport of various molecules across biological membranes
- **Phagocyte**- a cell that ingests and destroys foreign particles, bacteria or cell debris
- **Phenotypic**- the observable appearance of bacteria resulting from the interactions of gene expression and the environment
- **Phylogeny**- the evolutionary history of a group of bacteria
- **Prophage region**- an area in the bacterial genome that consists of genes transferred by bacteriophages
- **Resistant**- bacteria that have a very high MIC and has multiple resistance mechanisms
- **Robust**- (in relation to real-time PCR) an assay's ability to produce the same results under different conditions
- **Sensitive**- (in relation to real-time PCR) an assay's ability to correctly identify positive isolates as positive
- **Septic Arthritis**- infection of the joint
- **Serology**- the study of antibody and antigen reactions using serum
- **Serovar**- a subdivision of a species of bacteria distinguished on the basis of serological variation
- **Shearing**- (in relation to Whole-genome sequencing) the cutting of DNA into fragments of a defined size using enzymes or mechanical disruptions
- **Specific**- (in relation to real-time PCR) an assay's ability to correctly identify negative results as negative
- **Subtyping**- process whereby bacteria are precisely classified using antisera
- **Zoonotic**- an infectious disease which is capable of being transmitted from animals to humans

CHAPTER ONE

Literature Review

1.1. *Salmonella* through the years

1.1.1. Historical outbreaks

The earliest documented outbreak of *Salmonella* dates back to 430 B.C., Athens. According to the detailed accounts of the Greek historian Thucydides, nearly 100 000 people were suffering from diarrheal illness, fever and rash with a 25% mortality rate (1). Metagenomic sequencing of human remains from a mass grave, in the ancient Kerameikos cemetery, confirmed the presence of bacteria closely related to modern-day *Salmonella enterica* serovar Typhi (*Salmonella* Typhi) (2).

Subsequent historical records show typhoid fever outbreaks of significant proportions. Between 1607 and 1624, 6 000 English settlers in Jamestown, Virginia died after contracting *Salmonella* Typhi from drinking putrid river water (3,4). Similarly unsanitary conditions during the American Civil War in the 1860s resulted in more deaths from typhoid fever than from the war itself (5). The mortality rate due to typhoid fever remained high throughout the 19th century, with 174 deaths recorded per 100 000 people in Chicago alone (6). Only in 1896, with the development of a vaccine by Almroth Wright and experimental inoculation of British troops, did the mortality rate of typhoid fever drastically decrease (7).

Large outbreaks caused by non-typhoidal *Salmonella* outbreaks have also been recorded. In 1974, nearly 300 people contracted *Salmonella enterica* serovar Typhimurium (*Salmonella* Typhimurium) from drinking contaminated cider in New Jersey and became ill (8). However, the largest documented outbreak of *Salmonella* Typhimurium occurred in 1985 when more than 16 000 people became ill after consuming tainted milk, which was distributed across several states in the United States (US) (9). *Salmonella enterica* serovar Enteritidis (*Salmonella* Enteritidis), a close relative of *Salmonella* Typhimurium, was also implicated in a multistate outbreak in 1986 involving raw pasta (10). *Salmonella* Typhimurium and *Salmonella* Enteritidis have emerged as pathogens of global public health concern over the past 50 years (11).

1.1.2. Classifying *Salmonella*

Karl Joseph Eberth is largely credited as the first person to have observed and described *Salmonella* bacteria (12). In 1880 Eberth published a report describing bacilliform bacteria on tissue he collected from a patient that had died from typhoid fever (13). The bacteria detailed in this paper was later referred to as “Eberth’s Typhoid Bacillus” (14) in European journals describing similar bacteria present in diseased tissue. Thus, a defining point in the history of *Salmonella* was the acceptance of a group name, which spurred much of the taxonomical classifications used today.

At the same time, Theobald Smith, a research assistant of chief veterinary pathologist Daniel Elmer Salmon, isolated and described the pathogenicity and the microscopic structure of the bacilliform bacteria collected from swine afflicted with hog cholera (15). In 1900, Smith suggested that these bacteria be named *Salmonella* for the chief pathologist (16). Salmon wrongly claimed the discovery for himself, not acknowledging the work of Smith (17). Nevertheless, Smith’s further work pioneered the current fields of epidemiology, pathology and bacteriology (18).

Following the wide acceptance of the *Salmonella* group name in the scientific community, numerous reports were published describing bacteria belonging to this group. Each newly discovered *Salmonella* bacterium was classified according to their antigenic composition, a method known as the Kauffman-White scheme (19). Clinically significant *Salmonella* strains were commonly assigned a name according to the syndrome caused (eg. *Salmonella* Typhi) or host specificity (eg. *Salmonella* Abortusovis) (20,21). However, this nomenclature led to confusion in the scientific community due to the non-specific nature of some *Salmonella* strains. Hereafter, each strain was named according to the geographical location of isolation, eg. *Salmonella* London (21). In 1966, it was suggested that every description of *Salmonella* bacteria with a unique antigenic composition be considered an individual species (22). Thus, the antigenic formula was designated as the primary nomenclature system of *Salmonella*.

1.1.3. Current taxonomy of *Salmonella*

The *Salmonella* genus is classified under the broad Gammaproteobacteria class, the largest and most diverse class of Proteobacteria (**Figure 1.1.1**) (23,24). Within this class, Enterobacteriales is one of the most medically important orders (25). This order encompasses the *Salmonella*, *Escherichia*, *Shigella* and *Vibrio* genera, all of which have a significant medical impact on populations worldwide (23,24). These bacteria inhabit intestinal tracts of animals therefore are commonly known as enterics.

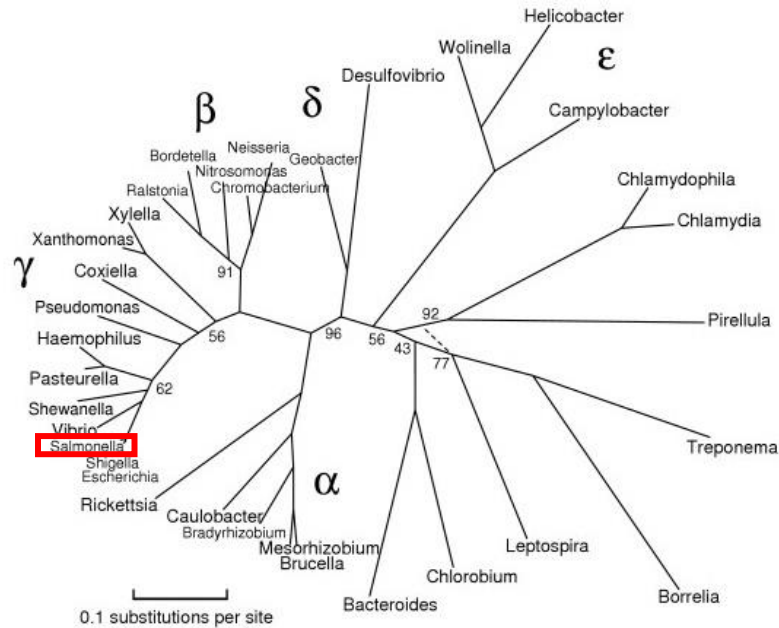


Figure 1.1.1 An unrooted phylogenetic tree representing the relationship between the five Proteobacteria classes (α , β , δ , ϵ and γ) and the bacterial genera of medical importance within each class. (26)

Molecular methods divided *Salmonella* into two distinct species, namely *bongori* and *enterica*, based on genetic relatedness (27,28). The latter, *Salmonella enterica*, accounts for the majority of disease in humans (25). Biochemical variations further differentiate this species into six subspecies: *enterica*, *salamae*, *arizonae*, *diarizonae*, *houtenae* and *indica* (29). The subspecies *enterica* contains the greatest number of serovars pathogenic to humans (30).

More than 2 600 serovars of *Salmonella enterica* have been described (31). These are loosely classified according to the clinical syndrome and antigenic formulae, as either, typhoidal or non-typhoidal. The typhoidal serovars are non-zoonotic in transmission and only pathogenic in humans (32). Conversely, non-typhoidal serovars are extremely diverse and have a multitude of hosts (32), ranging from cattle and swine (33–36) to plant tissue (37).

1.2. Non-typhoidal *Salmonella*

1.2.1. Characteristics of non-typhoidal *Salmonella*

As previously described by Eberth and Salmon, the *Salmonella* genus consists of bacilliform (rod-shaped) bacteria (13,17). The characteristic shape is due, in part, to the narrow layer of peptidoglycan in the cell wall (38). Gram stain tests of *Salmonella* bacteria result in gram-negative cells, since the thin cell walls are unable to retain the crystal violet-iodine complex (39,40). Multiple flagella embedded in the apical end of cells further distinguish *Salmonella* isolates (**Figure 1.2.1**). These whip-like structures are important for bacterial locomotion, attachment and invasion of host cells (41–43).

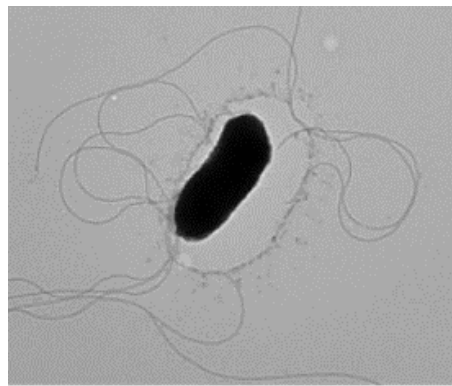


Figure 1.2.1 Electron micrograph of *Salmonella enterica* serovar Enteritidis. Flagella, extruding from the bacterial cell wall, assist with locomotion and host cell attachment. (44)

As facultative anaerobes, *Salmonella* are able to perform metabolic processes in the presence as well as the absence of oxygen (23,38). Respiration is independent of oxygen as it is not required as a terminal electron receptor. However, in the absence of oxygen, *Salmonella* makes use of another electron receptor, tetrathionate, which is produced when the bacteria elicit mammalian gut inflammation (45,46). When oxygen is present, catalase enzymes prevent the accumulation of toxic oxygen by-products within bacterial cells (47,48). The adaptability of these bacteria to live in both oxygen-rich and poor environments facilitates both free-living and intracellular lifestyles (45–48).

The adaptability of NTS to diverse niches is further supported by the ability of the bacteria to obtain carbon and nitrogen from a variety of sources (49). *Salmonella* metabolize a wide range of sugars through the use of permease systems capable of catabolic activity. Metabolism of glucose produces succinate, lactate and acetate, classifying *Salmonella* as mixed acid fermenters (23,50). Furthermore, the bacteria are capable of using both nitrite and nitrate as electron receptors facilitated by two types of membrane-bound reductase enzymes (51).

1.2.2. Method of host cell invasion

Once *Salmonella* bacteria enter the mammalian digestive tract, the bacteria penetrate the brush border cells (**Figure 1.2.2**) of the gut epithelium (50,52). Type I fimbriae, surrounding the bacterial cell, and flagella facilitate the binding of bacteria to the brush border cells (23,41,42). Modification of the host cell's filamentous actin cytoskeleton induces cell ruffling and enables bacterial entry (53–55). These modifications often disrupt the brush border cells, causing denuding and elongation of the villi (52). As the bacteria replicate and degrade the host cell, adjacent host cells are invaded through the junctional complexes (52,56).

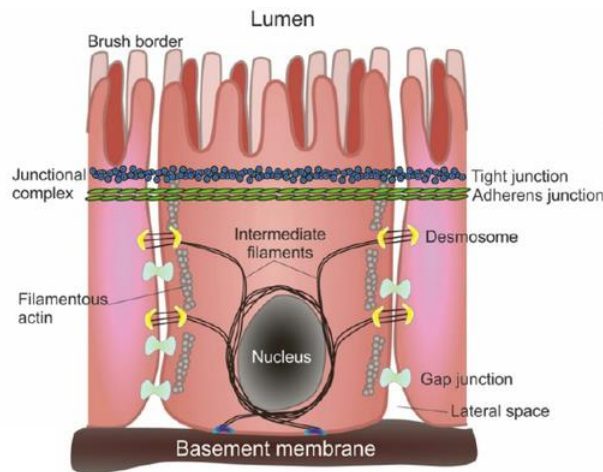


Figure 1.2.2 Illustration of gut epithelial cells situated closely together in the small intestine. The brush border and junctional complex are cell characteristics that make the intracellular invasion of *Salmonella* bacteria possible. (57)

Alternatively, the invasion of the intestinal epithelium may take place through uptake via CD18-expressing phagocytes (58). However, *Salmonella* preferentially invade the microfold (M) cells of the Peyer's patches, small masses of lymphatic tissue within the small intestine, and pass into the host lymphatic or cardiovascular systems of vulnerable hosts (23,59). The numerous methods of cell invasion employed by *Salmonella* emphasise the importance of intracellular environments for the successful replication of *Salmonella* (60,61). Thus, the *Salmonella* bacteria require an intracellular environment for at least part of the pathogenic lifecycle.

1.2.3. Symptoms of non-typhoidal *Salmonella* infection

Salmonella produces enterotoxins within the host intestinal tract (23). These specialised proteins bind to the host cell receptors and cause the rapid secretion of fluids into the intestinal lumen (62). Excessive fluid secretion leads to symptomatic diarrhoea and vomiting (23,63). The increased loss of fluids puts the host in danger of severe dehydration (23). Thus, the enterotoxins are highly immunogenic and result in the stimulation of neutralising antibody production as well as intestinal inflammation (23, 62, 63).

Symptoms develop in the host only once the bacteria exploit host nutrients and replicate to high tissue loads. Reportedly, the incubation period of NTS *Salmonella* is 6 to 72 hours from the time of contaminated food consumption (64,65). However, cases of incubation periods of up to 14 days after exposure have been recorded (66).

Salmonella most commonly manifests as gastroenteritis and does not usually cause systemic disease in humans (63). However, in 6% of all confirmed NTS cases, an invasive *Salmonella* infection occurs (67,68). Diarrhoea, vomiting and dehydration are uncommon in systemic infections (69). Instead, invasive NTS (iNTS) disease closely resembles the enteric fever syndrome of patients suffering from typhoid fever (70). Patients may also develop bacteraemia, or other extraintestinal sites of infection (including meningitis, osteomyelitis, and septic arthritis) (71,72). The mortality rate of these systemic infections remains high, even with the administration of antibiotics (71,73–75).

1.2.4. Transmission of non-typhoidal *Salmonella*

All non-typhoidal *Salmonella* are considered potential zoonotic pathogens, of which a small number are responsible for the majority of human infections(23,31). NTS infects a broad range of both warm- and cold-blooded animals (31). Poultry, in particular contaminated chicken products (**Figure 1.2.3**), are considered the primary source of *Salmonella* Enteritidis (12,76–80). Infection of human hosts generally occurs through faeco-oral transmission, either through ingestion of faecally-contaminated food or water, or from direct contact with animals (61). Globally, poultry is the second most widely consumed meat (81,82) and is currently the most rapidly growing agricultural industry (81). Thus, NTS poses a huge threat to the economy, food security, and healthcare systems worldwide (31,76,79,83).

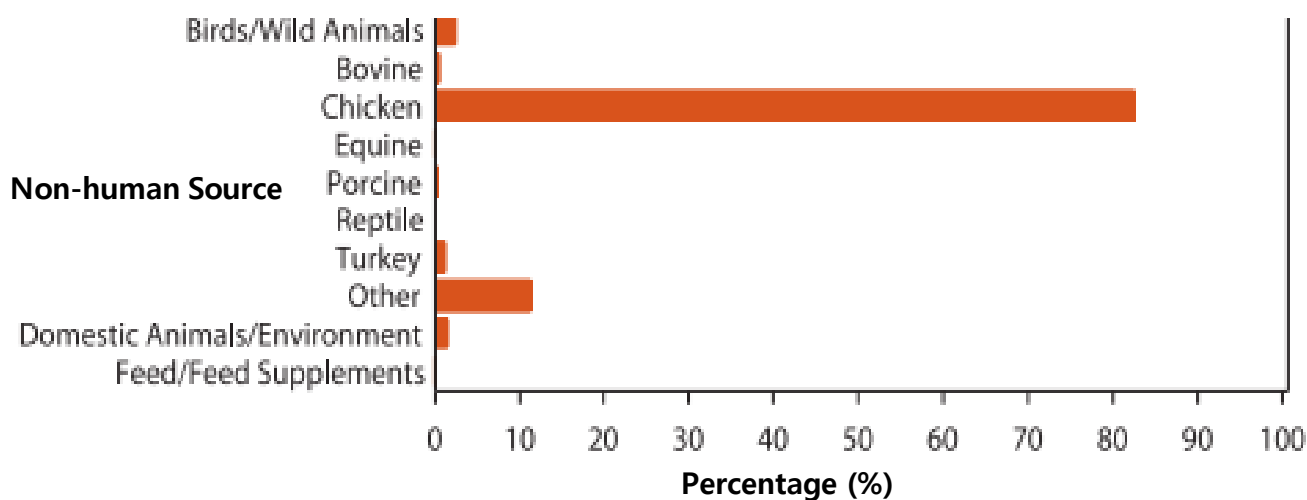


Figure 1.2.3 A bar graph representing the percentage of laboratory-confirmed *Salmonella* Enteritidis isolates taken from non-human sources from 1968 – 2011 in the United States. Chicken is significantly higher than any other source. Adapted from the United States National Veterinary Services Laboratories. (12)

The exact mode of transmission of iNTS disease remains largely unknown (84). A large proportion of *Salmonella* serovars isolated from production animals are not associated with invasive human infection, suggesting alternate transmission routes of invasive *Salmonella* (85–87). Similarly, water-derived samples are distinct from the human serovars causing iNTS disease (84). Some studies have proposed that the major route of transmission for serovars involved in the recently evolved African iNTS disease is anthroponotic (88,89).

1.2.5. Risk factors associated with non-typhoidal *Salmonella*

Non-typhoidal *Salmonella* is an opportunistic pathogen, therefore individuals with weakened immune systems are most susceptible to iNTS (71). Age has long been recognised as an independent risk factor for iNTS. Infants and children younger than two years have immature immune systems and are at high risk for disease (72,90,91). As the immune system matures, fewer NTS infections are observed, with the lowest iNTS incident observed in persons aged 10 to 44 years of age (90). However, the risk of infection as well as the mortality rate increases again in individuals over 65 years, due to immunosenescence and additional comorbid diseases (91).

Immunosuppressive conditions predispose individuals to iNTS disease. HIV infection, in particular, is associated with a high risk of developing iNTS disease (71,92–94). Other well-described risk factors include malnutrition, malaria, sickle-cell disease and anaemia (73,95). In absence of an efficient immune response, *Salmonella* bacteria are rapidly internalised, which also reduces the effectiveness of antibiotic therapy (94).

Ongoing global surveillance of NTS serovars has revealed that *Salmonella* Enteritidis is the only serovar that displays a significant seasonality. In both northern and southern hemispheres, the majority of infections occur in the respective summer/ autumn months (90,96,97). iNTS disease has also been shown to be seasonal, with a higher incident rate during the rainy season (98). It is unclear whether this is due to enhanced waterborne transmission, seasonal malaria, an association with food scarcity and malnutrition, or a combination of risk factors (98).

1.3. Typing *Salmonella*

Typing of bacterial isolates provides invaluable aid in the investigation of outbreaks. Traditional serotyping relies on the phenotypic characterisation of bacteria. These methods are commonly time consuming and require skilled personnel to perform, and even then are not consistently reproducible (99). Nucleic acid-based methodologies have improved on the performance of traditional phenotypic methods with high subtyping sensitivity and use of commonly available laboratory agents (100). As technologies develop and become more affordable, molecular methods will replace phenotypic characterisation as gold standards for serotyping (101–103).

1.3.1. Phenotypic laboratory identification of non-typhoidal *Salmonella*

a) Slide agglutination serotyping

Serotyping by slide agglutination, which defines serovars within a subspecies, remains the standard approach to *Salmonella* subtyping in well-resourced routine diagnostic laboratories and is a key aspect of typing in reference laboratories (101). Serovars are designated according to the White-Kauffmann-Le Minor scheme of antigenic formulae (28). Highly variable O and H antigens are detected using antisera, which form specific antigen-antibody complexes that are visible as agglutination (**Figure 1.3.1**). Somatic (O) antigens are the saccharide components of the exposed lipopolysaccharides on the bacteria surface. At least 64 O antigen variations have been identified (29). Flagellar (H) antigens have nearly double the number of variants, with at least 114 H antigen variations described (29). The increased variation of H antigens are due to a single *Salmonella* isolate containing more than one type of H antigen.



Figure 1.3.1 Visible agglutination (left) is observed as white dots in the solution as antigen-antibody complexes form. Photographed by Statens Serum Institut.(104)

Most *Salmonella* serovars possess two distinct H antigens, termed H1 (phase 1) and H2 (phase 2) (105). Serovars may express only one flagellar antigen type at a time (monophasic) or both antigen types at the same time (diphasic). Therefore, to determine whether the isolate has an H antigen other than that identified by initial slide agglutination, phase inversion is used. The isolate is passed through a special media containing antibodies that inhibit the dominant antigen to select for cells expressing the other H antigen (106). Once the type of H and O antigens of a particular isolate have been resolved, the Kauffman-White-Le Minor scheme is used to classify the isolate as a particular serovar (29).

Serotyping by slide agglutination is fairly quick for common isolates, with the average hands-on time being under one hour. Any microbiologist with the full array of antisera will be able to identify any *Salmonella* isolate. However, the method is very labour intensive and requires personnel to be carefully trained (101). False positives may also occur due to weak, non-specific agglutination (107) which limits the utilisation of traditional serotyping techniques, such as this, in detection of outbreak strains (101). In addition, the cost of maintaining a complete array of antisera has increased considerably, therefore is typically only available at reference laboratories.

b) Phage typing

Phage typing is used to discriminate between *Salmonella* isolates within the same subspecies. The phage type is assigned according to the ability of a specific bacteriophage to lyse the isolate under investigation (**Figure 1.3.2**) (108). The molecular characteristics of the bacteriophage and phage receptor of the bacteria are specific, in that the bacteriophage only infects particular *Salmonella* serovars (108). Initial use of this method was to discriminate between *Salmonella* serovars Typhi, Paratyphi A and Typhimurium, which are known to be highly monophyletic. Later, the phage typing methodology was adapted to include *Salmonella* Enteritidis (109).

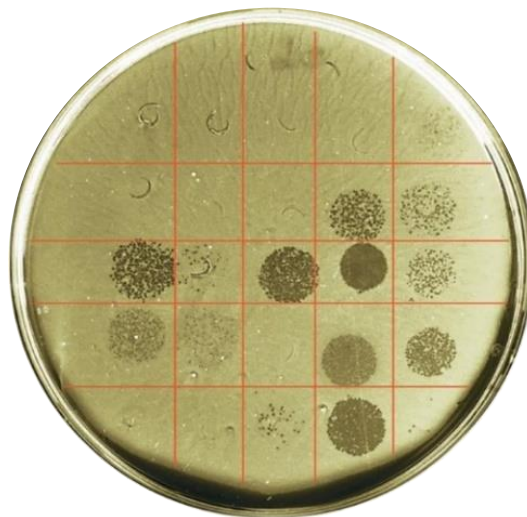


Figure 1.3.2 Diagram representing the phage typing of a *Salmonella enterica* serovar. The formation of plaques (dark circles) indicate the bacterial susceptibility to a particular bacteriophage. Adapted from Pearson Education Limited, 2015. (50)

Serotyping *Salmonella* with the use of specific bacteriophages is very simple to implement and only requires basic laboratory equipment (110). The simplicity of the test also enables results to be obtained in less than 24 hours (109). Careful coordination between reference laboratories is required to ensure tests are standardised, nevertheless results are not easily comparable (110). Bacteriophage lysis is also commonly ambiguous, which further affects the reliability of results. The limited number of available bacteriophages confines the range of serotyping by this method (101). Phage lysis may also be affected by genetic events such as the gain or loss of a genetic element resulting in the conversion of the phage (111,112).

1.3.2. Molecular subtyping and characterisation of non-typhoidal *Salmonella*

c) Pulsed-Field Gel Electrophoresis analysis

Pulsed-field gel electrophoresis (PFGE) analysis was considered the gold standard for *Salmonella* typing (101), as it has the capacity to identify strains through sensitive fingerprinting, which has important applications in outbreak investigation (113–116). It has recently been replaced by whole-genome sequencing (WGS) as a gold standard (117). The PFGE method makes use of patterns (**Figure 1.3.3**) created by large genomic DNA fragments passed through an agarose gel by alternating, multi-directional electric current (118).

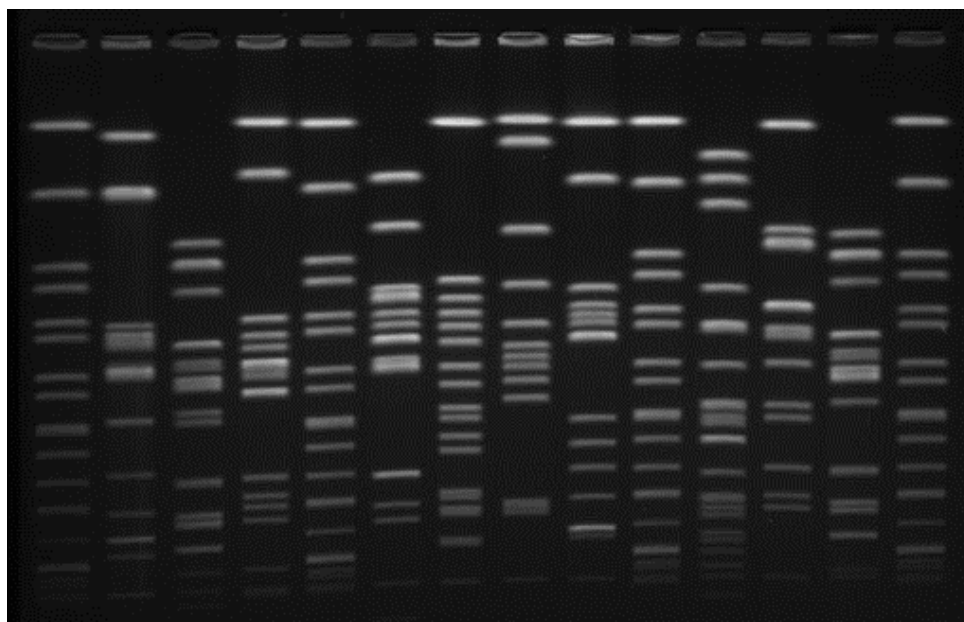


Figure 1.3.3 A Pulsed-Field Gel Electrophoresis image of *Salmonella enterica* serotypes. The patterns produced by the fragmented DNA make isolates comparable which is beneficial for outbreak detection. (119)

Standardisation of this test by the United States Centers for Disease Control and Prevention (US CDC) and the PulseNet network has allowed test results to be comparable through a set of databases shared worldwide (120). Thus, along with the reasonably discriminatory nature of the test, PFGE has enabled the successful tracking of numerous sources of *Salmonella* infections (113–116). However, as the database of isolate fingerprints increases, the reliability of clustering isolates becomes more difficult due to many large clusters differentiated by only minor differences (100,101). Sensitivity is also not displayed equally between different serovars (121). PFGE is a time-consuming process and labour-intensive, which limits the number of isolates that can be processed at one time.

d) Polymerase Chain Reaction based approaches

Polymerase chain reaction (PCR) is a method of target gene amplification using specifically designed primers and specialised thermal cyclers. The fluctuation of temperature facilitates the denaturation of DNA, annealing of primers to target sequences and the extension of the amplified strand (122). An efficiently designed PCR duplicates the target gene sequences with every cycle (123). The resulting amplification can be qualitatively measured using gel electrophoresis, as with conventional PCR (122). Alternatively, the gene amplification can be quantified in real-time by measuring the intensity of a fluorescent signal emitted by gene specific probes, as with real-time PCR (124).

Multi-locus variable number tandem repeat analysis (MLVA) makes use of the unique variable number tandem repeat (VNTR) regions within related bacterial genomes (125). Analysis of the VNTR fragment sizes produced by digestion of five gene loci results in distinct bacterial profiles (**Figure 1.3.4**). Numerous methods for serovars Typhimurium (125–127) and Enteritidis (128–131) have been developed. Each of these methods describe different target loci, primer designs and resulting fragment lengths. Thus, standardisation of the MLVA methodology is necessary for inter-laboratory comparison (132).

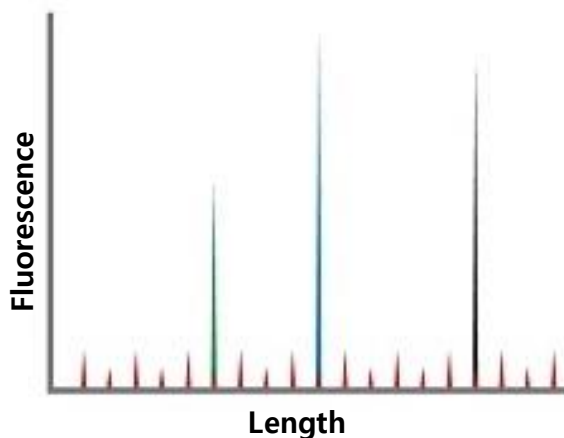


Figure 1.3.4 An electropherogram produced by the MLVA procedure. The red peaks indicate DNA standards of a known size and the green, blue and black peaks represent the sizes of PCR assay products from a bacterial isolate. A software program converts this data into a serotype. (133)

MLVA is a highly discriminatory test, capable of differentiating between genetically similar bacterial isolates (131,134). Presentation of results as a string of five numbers allows ease of international sharing and database setup (135). This requires collaboration for standardisation of the test between laboratories (135,136). However, MLVA is commonly restricted to specific serovars and is not useful for global phylogenetic studies or for assigning serovars (101). Furthermore, genetic homoplasmy among bacterial strains leads to apparent relatedness and misleading isolate groupings (137). Isolates frequently typed by MLVA require an additional complementary subtyping method during outbreak investigations to distinguish the unique serovars (138).

Random Amplified Polymorphic DNA (RAPD) analysis has proven an effective method of subtyping *Salmonella* species (139). The technique makes use of multiple short primers that amplify random polymorphisms within the bacterial genome (140,141). Replication of particular gene fragments creates a unique fingerprint when analysed using gel electrophoresis (140,141). The length of each gene fragment differs between *Salmonella* serovars, which can thus be used for accurate subtyping (139). RAPD analysis has a high discriminatory power, as observed with *Salmonella* Enteritidis subtyping (142). However, the arbitrary amplification of gene fragments is not easily reproducible between laboratories and is not useful for comparative studies (143). The discriminatory power of this method, when applied to highly homologous *Salmonella* serovars, is not as powerful as other PCR techniques (144).

Enterobacterial repetitive intergenic consensus (ERIC) sequences are imperfect palindromes of 127 base pairs (bp) located between coding regions (145). The presence or absence of ERIC sequences between particular genes in different bacterial species results in copy number variance (146). Thus, PCR amplification of the regions between ERIC sequences results in varied fragment lengths (145). This method provides a quick and cost effective way of differentiating between bacterial isolates (147). The use of a single primer set in this methodology also ensures a greater reproducibility than RAPD. However, the reproducibility of this technique as well as the discriminatory power remains lower than that of other molecular techniques, such as PFGE (147).

e) Ribotyping

Ribotyping compares the ribosomal operons (genes encoding the ribosomal RNA and associated sequences) of bacterial isolates (148,149). Ribosomal sequences 16S and 23S are highly conserved regions within the bacterial genome. Therefore, these regions are effective and stable for use in differentiating *Salmonella* serovars (150). The method makes use of rRNA probes which are hybridised to endonuclease digested, electrophoresed genomic DNA (149). The extent of hybridisation is then analysed by autoradiography.

The multi-step process of this methodology is labour-intensive, time-consuming and expensive (151). Furthermore, the method has a low reproducibility in different laboratories and has a limited discriminatory power between isolates digested with a single endonuclease. However, the use of two-enzymes in the digestion of DNA increases the discrimination between isolates (152,153). This method is suitable for distinguishing between bacterial isolate clusters, but the throughput is not ideal for use in outbreak investigations (152).

f) Multi-locus Sequence Typing

Multi-locus sequence typing (MLST) methodology is a gene-by-gene comparison approach based on the sequencing of seven metabolic housekeeping genes (101,154). These genes were specifically selected based on their scattered positions in the genome, the presence of genes with known function in the surrounding and the lack of any diversifying selection pressure (155). Establishment of a global MLST *Salmonella enterica* database (156) has made the method popular worldwide.

MLST is capable of discriminating between closely related bacterial isolates. The disadvantage is that the use of only seven loci limits the ability of the method to detect differentiating mutations (101). The method is also costly and requires a substantial amount of laboratory work in order to amplify, sequence and proof-read the nucleotide sequences (101,157). Thus, MLST is not ideal for routine laboratory testing. Recently, the MLST technique has been used to mine genes from WGS databases (eg. Bacterial Isolate Genome Sequence database) (154,158). This approach to MLST has increased in popularity due to the elimination of laboratory-intensive PCR analysis (154).

g) Whole-Genome Sequencing analysis

WGS is one of the few molecular methodologies that makes use of all genetic material within a bacterial cell. The entire genome undergoes multiple preparation steps, including shearing and bar-coding of the extracted DNA, before being sequenced by specialised machinery (159). Analysis of the resulting sequence reads using computer programs provides information on bacterial relatedness and the likeliness of linkage in an outbreak.

WGS provides detailed and precise data based on the comparison of millions of base pairs (103). The method is also rapid (160) and capable of high throughput on current sequencing platforms (161). Thus, WGS is ideal for investigating outbreaks (92). Nevertheless, preparation of DNA and sequence analysis requires skilled personnel. Machinery and software required for the testing and analysis of sequences is expensive, limiting the number of laboratories able to afford the technology. However, the continual improvement of WGS techniques has resulted in a gradual decrease in the cost of equipment and reagents (162,163). As this technique becomes more affordable, more equipment will become available

to laboratories worldwide and WGS will slowly replace other subtyping methods, as is already the case in several countries (including Europe and the US) (164).

h) Development of more sensitive subtyping methods

Clonal bacterial groupings, such as *Salmonella* Enteritidis, require highly discriminatory typing methods together with good epidemiological data to identify the source and the route of transmission of infections. A rapid and efficient molecular approach to outbreak detection is required for investigations, especially in South Africa where foodborne outbreaks frequently occur (165). However, subtyping tests are limited for NTS owing to high running costs and low sensitivity and specificity (31). Most work is focused on the phenotypic characteristics of iNTS disease, such as serotyping and multidrug-resistance (MDR) (73,74), and less on the phylogenetic relationships of the serovars involved.

Development of rapid, robust, portable, sensitive and specific tests which produce objective results are required to differentiate between epidemiologically unrelated strains (166). These tests should ideally group all isolates, originating from the same source, whilst conforming to the current classification of *Salmonella* subspecies and serovars (29). A balance between budget constraints and improved technologies which can be implemented on a mass scale in routine laboratories needs to be maintained with the ultimate goal of producing reliable results of benefit not only for optimal care of the individual patient, but also for broader public health purposes (101). Methods for the timely detection of *Salmonella* species contamination across the food chain continuum are essential to detect and prevent *Salmonella* infections.

1.4. *Salmonella enterica* serovar Enteritidis

1.4.1. Disease surveillance

Many countries (mostly high-income) have established surveillance programmes that monitor *Salmonella* Enteritidis incidence and patterns of disease, which facilitates outbreak detection. Regulations in Europe, the US and Australia have specified NTS infections, including *Salmonella* Enteritidis, as notifiable to health authorities (167–169). Networks such as the European Food Safety Authority (EFSA) in the European Union (EU) (96), the *Salmonella* Enteritidis Outbreak Reporting System in the US (170) and OzFoodNet in Australia (171) collect and analyse these reported disease data. Regular surveillance of *Salmonella* serovars provides valuable insight into the epidemiology, and facilitates informed disease prevention strategies.

However, human *Salmonella* Enteritidis surveillance in Africa is poor or non-existent. The major challenge is the scarcity of laboratories in low-resource settings (172). Limited funding resulted in general surveillance for *Salmonella* in South Africa being stopped in 1992 (173). Thereafter, most available information relates to *Salmonella* Enteritidis disease notification in animals. Animal Health acts and regulations established by the Namibian Ministry of Agriculture, Water and Forestry (Act 180 of 2011) (174) and the South African Department of Agriculture, Forestry and Fisheries (Act 35 of 1984 and regulations 2026 of 1986) (175) state the only official control and notification measures necessary for *Salmonella* Enteritidis infections in animals.

In response to the lack of data on human NTS disease in Africa, the South African GERMS program began active surveillance of NTS in 2003. Clinical laboratories, both public and private health sectors, regularly submitted NTS isolates to the National Institute of Communicable Diseases. The valuable information gained from this surveillance enabled detection and investigation of *Salmonella* outbreaks nationwide (176,177). Furthermore, the GERMS database allows for the constant monitoring of NTS disease trends for rapid outbreak detection and antimicrobial resistance monitoring.

1.4.2. Global *Salmonella* Enteritidis burden

Worldwide, NTS is responsible for 93.8 million enteric illnesses and 155 000 deaths annually (67). One of the five most commonly reported serovars causing human disease is *Salmonella* Enteritidis (96,178,179). Approximately 67% of all NTS cases reported across Europe are attributed to *Salmonella* Enteritidis (180), with an estimated 67 000 cases reported annually (96). Similarly, the majority of NTS cases reported in the US are attributed to *Salmonella* Enteritidis (170). However, the average annual rate of infection in the US was significantly less than the EU, with 3 cases per 100 000 people between 1996 and 2011 (90). In 2012 only 1 239 confirmed cases of this particular serovar were recorded (170).

An estimated 534 600 infections and 77 500 deaths were attributed to iNTS disease in 2017 (181), a problem which may worsen due to the increasing spread of antimicrobial resistant strains (182). Antimicrobial resistant strains originate due to the selection pressure of antibiotic use along the food chain. The use of antibiotics in food animals and the human population is the main driver of antimicrobial resistant strains causing disease in humans (179). Increased use of first-line drugs (ampicillin, chloramphenicol, sulphonamides and tetracycline) to treat severe and invasive NTS cases has led to rapid resistance (183,184). The high prevalence of MDR serovars causing iNTS disease necessitates the use of second-line antibiotics, including fluoroquinolones (ciprofloxacin) and third-generation cephalosporins (ceftriaxone) (71,185). Inevitably, this too has resulted in the emergence of fluoroquinolone (FQ) resistant *Salmonella* serovars (186). The global infection and mortality rate of iNTS disease will continue to increase if the antimicrobial resistance of individual *Salmonella* serovars remains unaddressed.

1.4.3. The burden of *Salmonella* Enteritidis in Africa

Sub-Saharan Africa (sSA) has experienced epidemic increases of *Salmonella* Enteritidis and *Salmonella* Typhimurium as MDR strains have emerged (31,98). Such MDR strains are typically resistant to all first-line antibiotics, making treatment in resource-limited areas challenging (187). Emerging FQ resistance is reported more commonly in *Salmonella* Enteritidis than in other serovars (84). This serovar alone accounts for 26% of all laboratory confirmed NTS cases (188). However, due to the lack of microbiological and epidemiological data on *Salmonella* serovars in African countries, accurate estimates of NTS infections and serovar prevalence are largely undetermined (69,189,190).

NTS is one of the most common causes of community-acquired bacteraemia across sSA (98). Accumulatively, iNTS disease accounts for 421 600 illnesses every year in Africa alone which is more than half of all reported cases globally (178,181). Mortality rates associated with iNTS disease remain high, with an estimated 49 440 iNTS related deaths in sSA for 2017 (181). The most common serovars implicated in iNTS disease are Typhimurium and Enteritidis (98,191–193).

iNTS disease is associated with high mortality rates in low- and middle-income countries (LMIC), due to weak healthcare systems, persistent, poor sanitary conditions and lack of safe water supplies (187). Furthermore, the incidence and deaths associated with iNTS disease in Africa constitute the world's highest as measured by Disability-adjusted Life Years (DALYs) (**Figure 1.4.1**) (194). This has a detrimental effect on the socioeconomic environment and healthcare systems of these countries, the majority of which classify as LMIC (195).

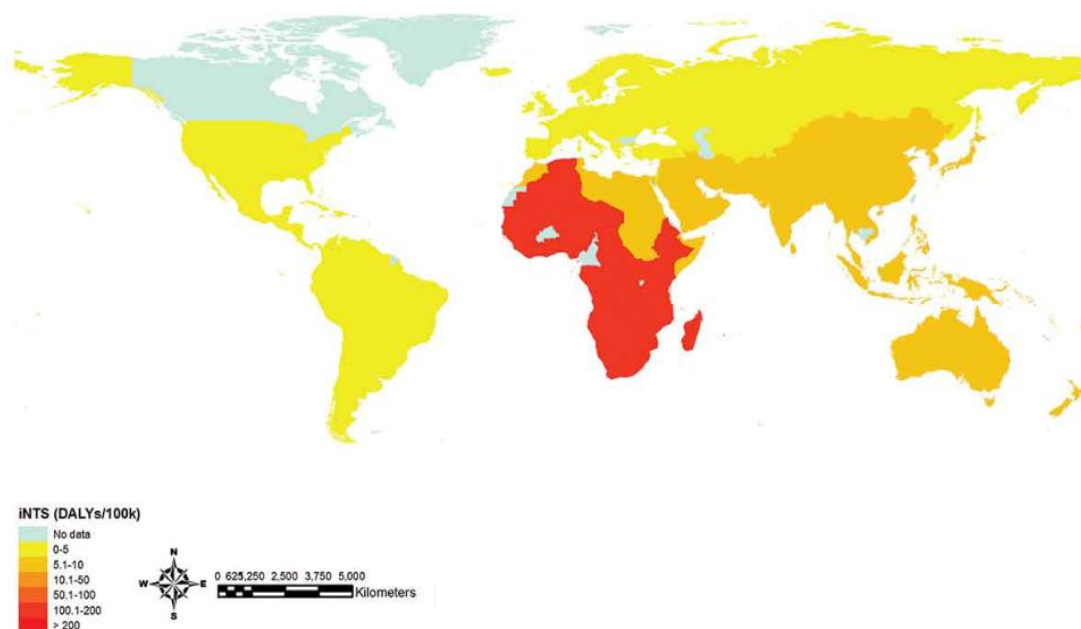


Figure 1.4.1 Disability-adjusted Life Years (DALYs) due to invasive non-typhoidal *Salmonella* based on estimates calculated by the World Health Organization. (194)

In South Africa, *Salmonella* Enteritidis surpassed *Salmonella* Typhimurium as the most commonly isolated serovar nationwide in 2011 and has remained the most common serovar since (**Figure 1.4.2**) (196). GERMS-SA surveillance reported the highest number of all laboratory-confirmed *Salmonella* Enteritidis cases in the years 2011 to 2014, with more than 3 500 isolates confirmed (196–198). This increase in cases corresponded with foodborne disease outbreaks in multiple provinces (199). South Africa has the lowest incident rate of iNTS disease (approximately 1.4 confirmed cases for every 100 000 people) of all the African countries (200).

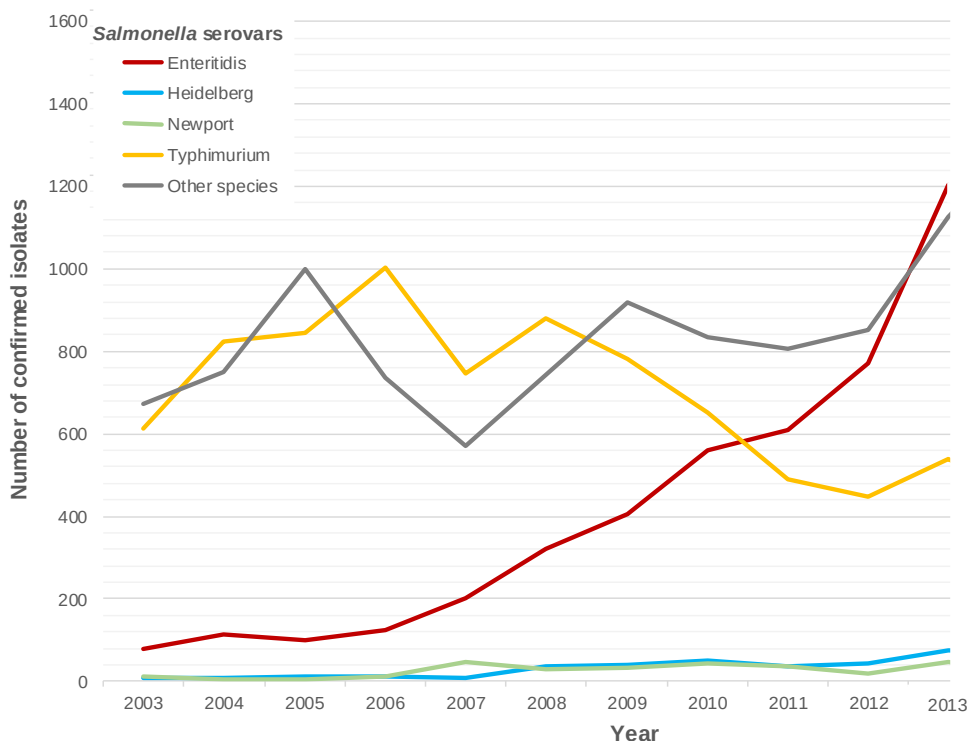


Figure 1.4.2 Number of confirmed non-typhoidal *Salmonella* isolates (invasive and non-invasive) of different serovars submitted to the South African GERMS program over a 12 year period. Data adapted from GERMS-SA database at the National Institute of Communicable Diseases by the author, Sarah Gallichan.

1.4.4. Preventive measures for *Salmonella* Enteritidis

Effective prevention of NTS infections requires a concerted effort along the farm-to-fork continuum (31). FQ resistance is especially important at the human-animal interface since both have the potential to be hosts for resistant *Salmonella* serovars (186). Regular use of FQ in agricultural practices, especially use in poultry, has been directly linked to resistance in human *Salmonella* infections (201). Complete banning of FQ use in medical and agricultural practice has proven effective in decreasing resistant bacteria in food animals and humans in Australia (202,203).

Recent reports have shown that the number of NTS cases in Europe and the US have remained constant over the years (**Figure 1.4.3**) despite intensive control efforts in those regions (170). A better understanding of the transmission dynamics of individual serovars may be necessary in order to implement more effective, targeted preventive measures (90). This can be achieved through the consistent monitoring of the evolution of *Salmonella* serovars and their resistance patterns (186). There is a significant need to strengthen global detection and integrated surveillance at various levels of the food cycle, especially in African countries, through rapid, reliable and cost-effective *Salmonella* serovar detection (31).

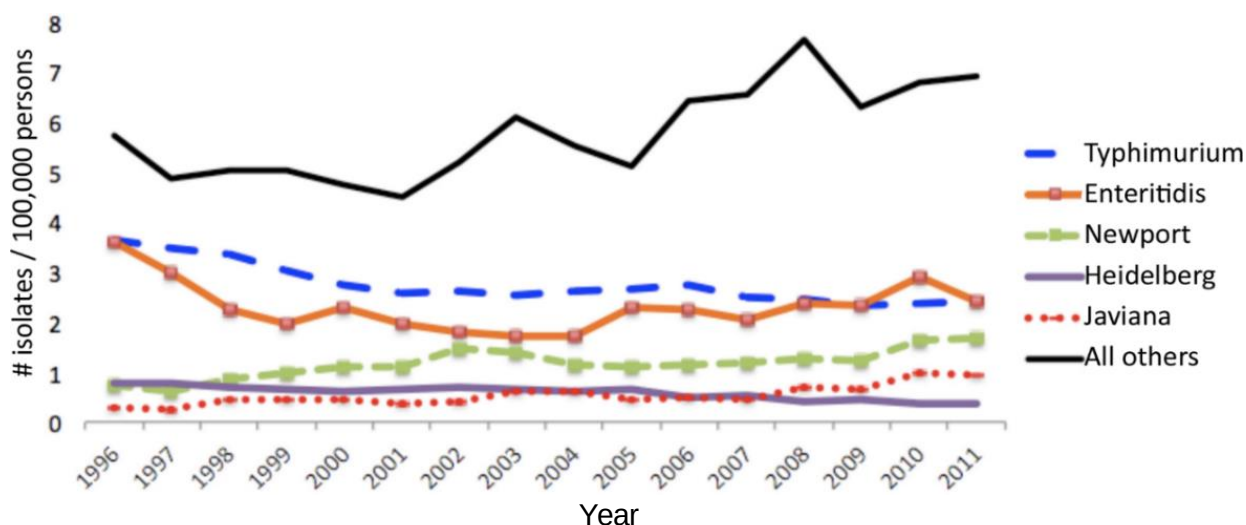


Figure 1.4.3 Number of confirmed non-typhoidal *Salmonella* isolates of different serovars in the United States over a 15 year period. (204)

Since the introduction of antiretroviral therapy (for the treatment of HIV) in South Africa in April 2004, there has been a steady decline in the number of reported invasive *Salmonella* Typhimurium cases (94,205). However, invasive *Salmonella* Enteritidis has not followed this trend (**Figure 1.4.2**). This suggests that, unlike *Salmonella* Typhimurium, HIV is not the only or most important risk factor predisposing a human host to disease (94,206). The burden of infection in poultry coupled with poor food safety control measures may have contributed to the increased prevalence (94,206). If prevention and control of *Salmonella* Enteritidis in the food chain is improved, the number of reported cases should decrease, as has been the experience in many other countries worldwide (94).

1.5. Genomic composition of *Salmonella* Enteritidis

1.5.1. Genome characteristics

The *Salmonella enterica* genome ranges in size from 4.6 to 4.9 million bp (207). Large regions of the chromosome are composed of clustered virulence genes necessary for invasion and survival within host cells (208,209). Several of these *Salmonella* pathogenicity islands (SPIs) are present in all *Salmonella enterica* serovars (31). SPI-1 and SPI-2 contain genes responsible for the virulence of *Salmonella* serovars. SPI-1 genes encode type III secretion systems, which are essential for intestinal invasion and intracellular survival (210–214). SPI-2 genes, on the other hand, are required for invasive infection and survival within macrophages (212,213,215). The genes within the SPIs are critical to the pathogenicity of the bacteria and mutations within these regions results in a significant decrease in virulence (44,187). Thus, these regions are highly conserved within *Salmonella* serovars (216).

Salmonella Enteritidis contains the smallest of all *Salmonella* virulence plasmids, at only 58 kilobase pairs (kb) in size (217,218). The plasmid phylogeny mirrors that of the chromosome, suggesting the stable maintenance of the plasmid by each lineage and diversification with the chromosome (218). This stability is possibly due to the inability for self-transmission of the plasmid by conjugation (219). Although this is common amongst NTS serovars, the large deletions within the *tra* operons on the *Salmonella* Enteritidis plasmid that prevent self-transmission are unique (218,220).

1.5.2. Evolution of a genome

Salmonella Enteritidis is the most genetically homogenous *Salmonella* serovar (221–223). However, the loss and gain of mobile elements, such as prophage regions, significantly increases the genetic diversity of the accessory genome (224). Bacteriophages, viruses that employ bacterial cells to replicate, transfer useful cargo genes to their host in order to perpetuate (225,226). These novel genes introduced to the bacterial genome confer levels of specialisation (227). Certain genes may increase the virulence of individual isolates which facilitates competitiveness (227), while others may enable the colonisation of novel niches within the host (226).

Of particular interest is the genomic evolution associated with niche adaptation towards systemic infection (216). Distinct genomic signatures differentiate the gastrointestinal and invasive lifestyles of *Salmonella* Enteritidis (228). This has occurred through focused genomic degradation whereby functions required for accelerated growth in the gut are lost when the lineage becomes invasive (229–231). A decrease in the functional gene pool within the *Salmonella* Enteritidis genome may be indicative of a shift towards a more specialised host range, similar to that of *Salmonella* Typhi (190).

Furthermore, acquisition of MDR genes increases the competitive nature of novel isolates allowing for clonal expansion (229–231). Extrachromosomal elements, such as MDR- and enterotoxin-encoding genes, are often associated with virulence plasmids. The inability for self-transmission of plasmids in most *Salmonella* serovars suggests that MDR genes were acquired multiple times from bacteriophages or point mutations (216). Although the virulence plasmid provides an adaptive advantage, its presence is not exclusively associated with a more invasive lifestyle (44).

Genomic transformations have previously been described in the closely related *Salmonella* Typhimurium sequence type 313, which was implicated in multiple African iNTS disease outbreaks (232,233). The persistence of this particular sequence type within the sSA human population was attributed to changes in metabolic functioning and evasion of the gut inflammatory response through surface protein modification (233). Degradation of the *Salmonella* Typhimurium genome as well as the contribution of prophage genes facilitated these adaptive responses (233,234), similar to what has been observed in the current *Salmonella* Enteritidis population.

1.5.3. Population structure

Increased incidence of uncharacteristically severe invasive NTS disease in Africa suggested that distinct phage types were circulating in these countries (69). The significant difference in *Salmonella* Enteritidis niche colonization prompted investigation into the global genetic diversity of this particular serovar (216). Whole-genome sequence comparison of *Salmonella* Enteritidis isolates from across the world identified four definitive clades: Outlier, Global, East/Central African and West African (**Figure 1.5.1**) (216). Three of these clades were implicated in epidemics: the Global Clade was linked to poultry-associated human gastroenteritis (235), while the two African clades were related to severe invasive disease (200,236). The most notable differences between these clades were the pattern of genomic degradation and acquisition of unique prophage regions and MDR virulence (218,230,231).

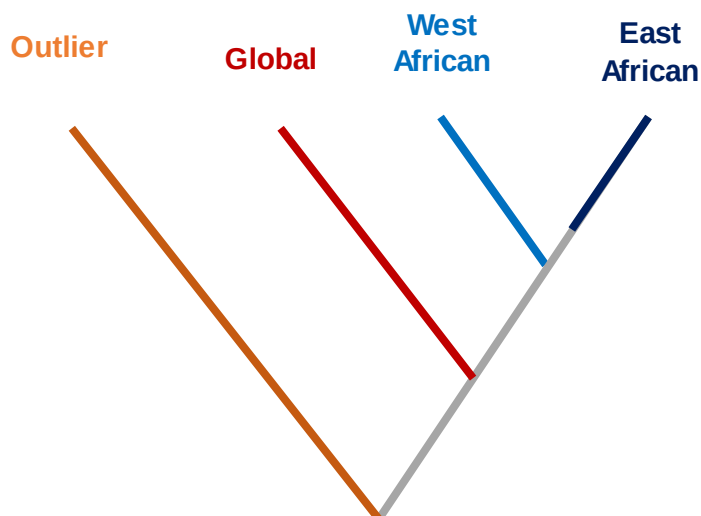


Figure 1.5.1 A cladogram representing the evolutionary relationship between the four *Salmonella* Enteritidis clades. Diagram created by the author, Sarah Gallichan.

The Global clade is the most geographically diverse of all four clades. This clade emerged from the paraphyletic Outlier cluster in 1918 and rapidly expanded in 1976 (216), corresponding with the significant increase in the number of *Salmonella* Enteritidis outbreaks in the US, EU and South America (237–240). The worldwide expansion of this lineage correlated with the acquisition of a prophage region (Φ SE20) (**Figure 1.5.2**) containing genes essential for the colonisation of chicken reproductive organs (241). Furthermore, the implementation of a vaccine for *Salmonella* Gallinarum, a prominent fowl pathogen, enabled the exploitation of this host by individuals of the Global clade (242). The specific requirements for the colonisation and survival within poultry may have placed selective pressure on the clonal expansion of *Salmonella* Enteritidis lineages (243–245). Thus, the Φ SE20 prophage region is relatively well conserved among isolates belonging to the Global clade (216).

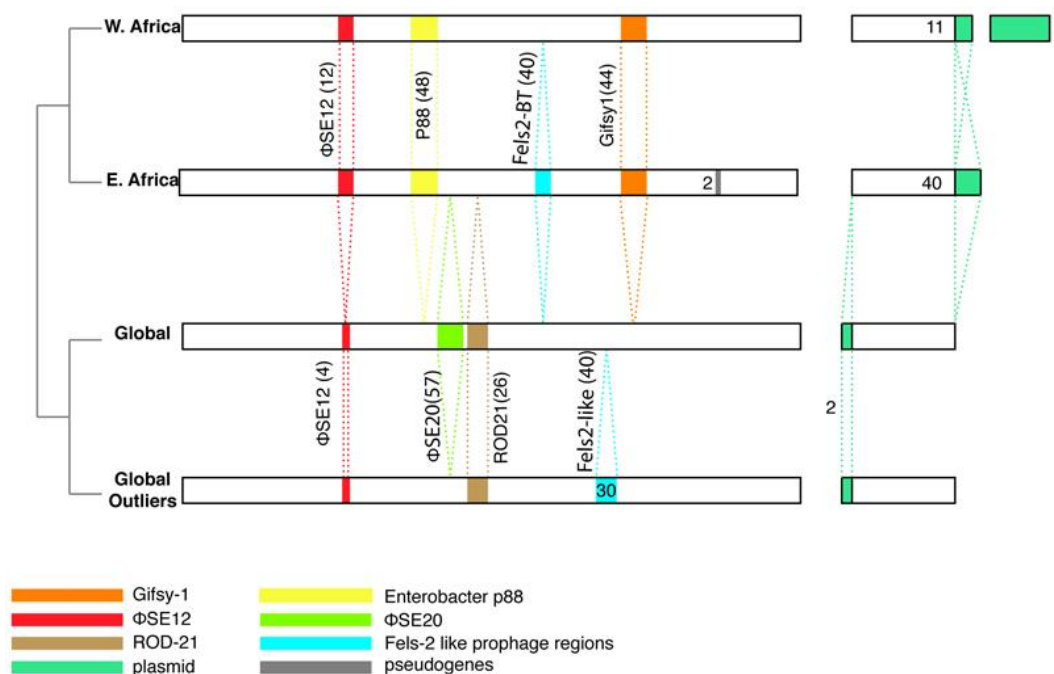


Figure 1.5.2 Diagrammatic representation of the differences between the accessory genomes of *Salmonella* Enteritidis clades. Adapted from Feasey *et al.* 2016. (218)

West and East/Central African clades are endemic to sSA and, as the naming suggests, are geographically distinct (216). These clades emerged from the Outlier cluster in 1933 and 1945 respectively (216). Reports of high iNTS disease incidence in Eastern Africa (Kenya, Malawi, Tanzania, Uganda) and Western Africa (Cote d'Ivoire, Nigeria, Ghana) correspond with the circulation of the two distinct African clades within these populations (200). Independent acquisition of unique prophage regions has enabled the development of these African lineages towards a more invasive lifestyle within human hosts (216).

The Gifsy-1 prophage region (**Figure 1.5.2**), ubiquitous in African *Salmonella* Enteritidis lineages, encodes factors related to the pathogenicity and increased competitiveness of the bacteria (225,246). Selective advantage conferred by genes in these loci has ensured the favourable maintenance of the region (216,246). A new Fels2-like prophage region (**Figure 1.5.2**) was acquired exclusively by East African isolates (216), however the functional significance of this region is unknown.

Most notably, the African clades contain virulence plasmids that are double the expected size in part due to the independent acquisition of MDR genes (216). The plasmids within these lineages are not novel; rather the different resistance genes were acquired through mobile genetic elements or clonal expansion in each region (229–231). East African isolates contained *catA2* chloramphenicol resistance gene and the *tet(A)* tetracycline resistance genes on the MDR plasmids while the West African lineage contained a different chloramphenicol resistance gene, *catA1* (216). The size and proportion of immunocompromised populations, driven primarily by the HIV pandemic in the afflicted African countries, facilitated the proliferation of these invasive MDR *Salmonella* Enteritidis lineages (75,247).

1.6. Study Objectives

Despite the recognition of different *Salmonella* Enteritidis clades, the distribution and epidemiology of these clades across Africa is poorly understood (187,190). Challenges with the availability of sensitive and affordable diagnostic facilities result in a limited number of reported NTS cases (165). Efficient, sensitive molecular diagnostic techniques are required to gain a better understanding of the epidemiology of specific *Salmonella* serovars implicated in NTS disease and outbreaks (165). In order to address these issues our study objectives were as follows:

- Identify cargo genes within the distinct prophage regions that can be used as molecular markers to distinguish between the four *Salmonella* Enteritidis clades (Global, Outlier, West African and East African)
- Develop a sensitive multiplex real-time PCR assay using the selected molecular markers to classify *Salmonella* Enteritidis isolates into clades
- Gain a retrospective understanding of the prevalence of each of the four distinct clades (Global, Outlier, West Africa and East Africa) present in South Africa over a two year period (2012 to 2013), by analysis of isolates stored in the GERMS-SA repository
- Determine the geographic distribution of each of the four *Salmonella* Enteritidis clades (Global, Outlier, West Africa and East Africa) across four South African provinces (Gauteng, Western Cape, KwaZulu-Natal, Mpumalanga) using the repository GERMS-SA surveillance data for *Salmonella* Enteritidis isolates from 2012 and 2013
- Statistically explore the association between *Salmonella* Enteritidis clades (Global, Outlier, West Africa and East Africa) and invasive disease, as well as the association with antimicrobial resistance

CHAPTER TWO

Development of a clade-classifying real-time PCR assay for *Salmonella* Enteritidis

2.1. Introduction

Non-typhoidal *Salmonella* (NTS) were estimated to be responsible for 95.1 million infections and 50 771 deaths are attributed to NTS gastroenteritis (2017 estimate) (248). In sub-Saharan Africa, invasive NTS (iNTS) disease is of greater concern, thought to be responsible for 421 600 illnesses and 49 440 deaths every year (also a 2017 estimate) (181). *Salmonella* Enteritidis is the most commonly reported serovar, and is implicated in both gastroenteritis and iNTS disease (96,178,179). Although incident rates remain lower than they were during the 1980s epidemic, the number of annual infections in both Europe and the USA have remained constant for more than 10 years despite multiple efforts to reduce infection (179).

Sub-Saharan Africa (sSA) accounts for the majority of the global iNTS infection burden (249). Confirmed cases from this region are often associated with multidrug-resistance (MDR) and a high mortality rate (73–75,218). In the case of *Salmonella* Enteritidis, the distinct characteristics of sSA iNTS infections have been associated with two geographically-defined lineages within these countries (218). Through genomic degradation, phage cargo gene transduction and the acquisition of MDR genes the African clades (West African and East African clade) have become genetically distinct from the worldwide epidemic lineage (Global clade) (218). Despite the recognition of different *Salmonella* Enteritidis clades, the distribution and epidemiology of these clades across Africa is poorly understood (187,190).

The paucity of sSA *Salmonella* Enteritidis information is largely due to resource limitations, in particular regarding diagnostic microbiology facilities. Challenges with accurate diagnoses arise from the lack of infrastructure, trained personnel and sensitive tools in the developing countries (69,189,190,194). The genetic homogeneity of the *Salmonella* Enteritidis serovar further complicates the investigation of outbreaks (221–223). Ideally, regional public health laboratories need access to robust, accurate and cost-effective tests with a rapid turnaround time that is capable of differentiating between genetically similar isolates. This will enable a better understanding of epidemiology of *Salmonella* Enteritidis and provide meaningful data for public health action.

In this chapter, we describe the development of a clade-classifying real-time PCR assay using genes within discrete phage regions as molecular markers. This high quality and cost-effective tool is suitable for implementation in regional public health laboratories throughout sSA and other low- and middle-income countries (LMIC) where whole-genome sequencing (WGS) may not be readily available.

2.2. Materials and methods

2.2.1. Workflow for the development of a clade-classifying real-time PCR assay

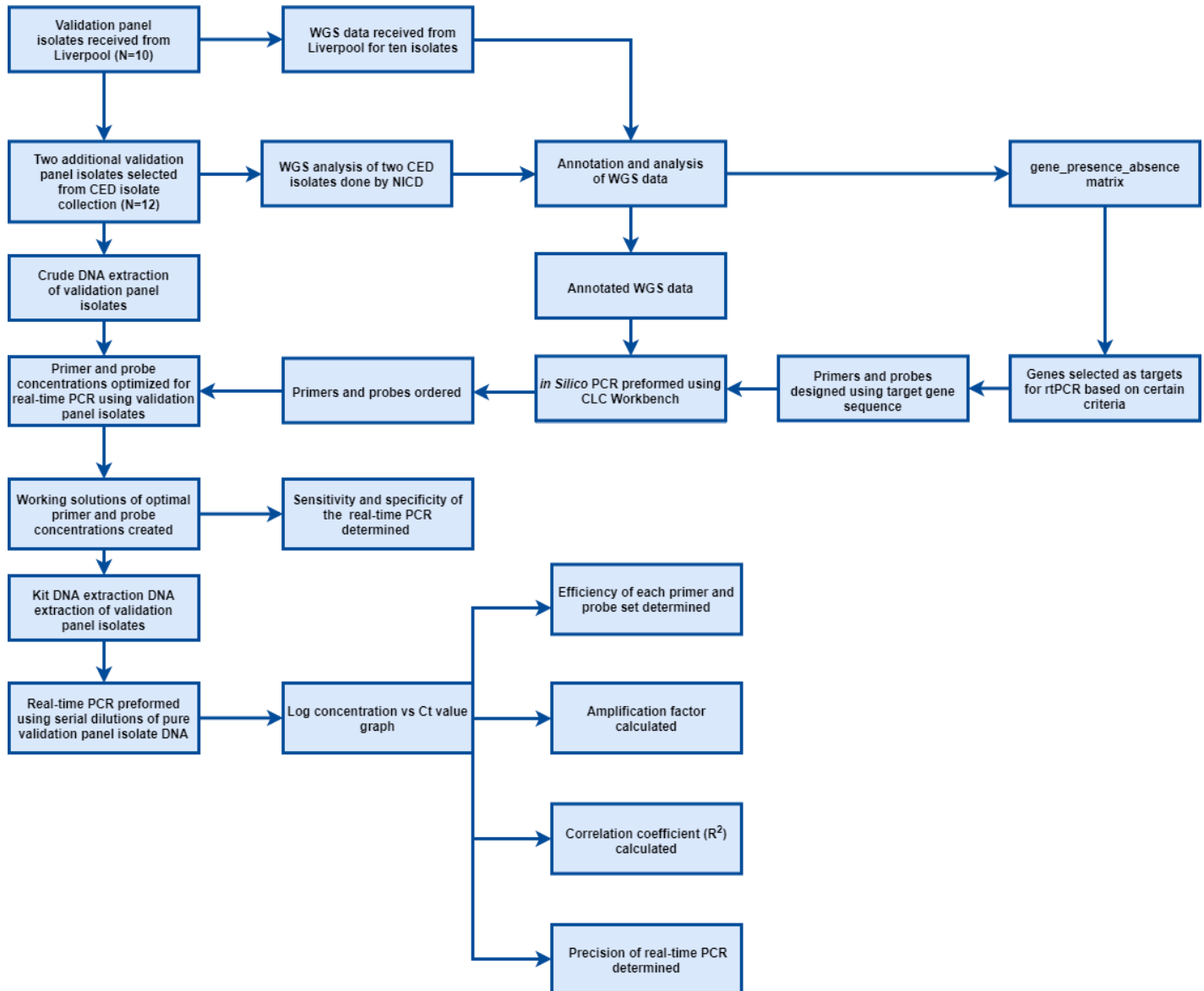


Figure 2.2.1 Flow chart showing the workflow for the identification of *Salmonella* Enteritidis genes for use as clade molecular markers and the development and validation of a clade-classifying real-time PCR assay

2.2.2. Research facility

All work related to this study was conducted at the Centre for Enteric Diseases (CED), National Institute for Communicable Diseases (NICD), Sandringham, Gauteng, South Africa.

2.2.3. Real-time PCR assay validation panel isolates

Selection of isolates for the validation panel was based on the predetermined hierBAPS (Bayesian Analysis of Population Structure) algorithm clade classification of worldwide *Salmonella* Enteritidis isolates (215). Ten isolates (N=10) were selected to represent the East African (n=3), West African (n=3), Global (n=3) and Outlier (n=1) clades (**Table 2.2.1**). The University of Liverpool, England, made these selected isolates (as used in the Feasey *et al.* 2016 (218) study) available to the authors.

Table 2.2.1 Validation panel isolates sent from the University of Liverpool

Isolate name	Origin	Clade	Category
P125109	UK	Global	Global/ Global Outlier
A1636	Malawi	Global	
1320	Uganda	Global	
791	Uganda	Global Outlier	
D7795	Malawi	East African	African
255	DRC	East African	
6396	Uganda	East African	
10136/01	Gambia	West African	
0527/01	Gambia	West African	
8078/01	Gambia	West African	

2.2.4. Additional samples for validation panel

Two additional Outlier isolates (**Table 2.2.2**) were added to the validation panel (N=12) to bring the total number of Outlier clade isolates to three (n=3). Selection of the two Outlier isolates was based on the predetermined hierBAPS algorithm clade classification of South African *Salmonella* Enteritidis isolates (215), which consisted of selected isolates from the CED isolate collection. The CED participates in the GERMS-SA surveillance program. GERMS-SA consists of a nationwide network of public and private clinical laboratories that participate in active surveillance of pathogens important to public health. Laboratories from all nine provinces have submitted confirmed *Salmonella* Enteritidis isolates to the CED since 2006.

Table 2.2.2 Additional validation panel isolates from GERMS-SA surveillance repository

Isolate name	Origin	Clade	Category
672246	South Africa	Global Outlier	Global/ Global Outlier
672632	South Africa	Global Outlier	

2.2.5. Whole-genome sequencing

The ten Liverpool isolates were sequenced by the Institute of Integrative Biology at the University of Liverpool using Illumina MiSeq (Illumina, California, USA) and PacBio (Pacific Biosciences, California, USA) equipment. Whole-genome scaffolds were received from the University of Liverpool for further analysis. The two additional CED validation panel isolates were sequenced using the Illumina MiSeq equipment at the sequencing facility on the NICD campus. Whole-genome sequences of all twelve isolates were annotated using the Prokka software tool (250). The resulting twelve annotated genomes were analysed together using the ROARY pipeline (251). This produced a gene_presence_absence matrix containing information on the pan genome of the validation panel isolates.

2.2.6. Selection of category- and clade-specific genes

To differentiate between the two broad categories (African and Global/ Outlier), distinct genes for each of the two categories were chosen according to the following criteria: the gene was shared exclusively by the six isolates in the category and the gene sequence was conserved among the six isolates. To distinguish between the four clades (East African, West African, Global and Outlier clades) within the two categories (African and Global/ Outlier), distinct genes were selected for one clade in each category (ie. East African and Global clade). This was done according to the following criteria: the gene was shared exclusively by the three isolates in the clade and the gene sequence was conserved among the three isolates. Thus, isolates were classified according to the presence and absence of particular genes (Table 2.2.3).

Table 2.2.3 Classification of isolates into clades according to real-time PCR assay results

Clade	Real-time PCR assay results			
	African category gene	Global/ Global Outlier category gene	East African clade gene	Global clade gene
East African	+	-	+	-
West African	+	-	-	-
Global	-	+	-	+
Global Outlier	-	+	-	-

2.2.7. Gene selection from whole-genome sequencing data

Selection of genes for both the category and clade differentiation was done according to the same methodology. The gene_presence_absence matrix was analysed using Microsoft Excel 2010 (Microsoft Corporation, Redmond, USA). A list of genes shared by the isolates within each category/ clade was created from the matrix. Using the gene annotation, the sequences of the relevant genes were located within the annotated Whole-genome of each isolate using CLC Genomics Workbench version 11 (QIAGEN, Hilden, Germany) sequence search bar. For each listed gene, the variations from each individual isolate were extracted and saved as separate sequences. These were then aligned using the Create Alignment tool. Genes with highest sequence conservation among the isolates were favoured. A consensus sequence was extracted using the Extract Consensus Sequence tool from the resulting alignment and exported as a text file (.txt). In this study the small toxic peptide gene, antitoxin pemI gene, TIR protein domain gene and Chitinase Class I gene were selected to represent the African Category, East African Clade, Global/ Outlier Category and Global clade, respectively.

2.2.8. Primer and probe design for real-time PCR assay

Sequence text (from the .txt file) of each selected gene consensus sequence was copied into the online PrimerQuest tool (Integrated DNA Technology). Primer and probe criteria (**Table 2.2.4**) set by the manufacturers of the 7500 Real-Time PCR System (Applied Biosystems, ThermoFisher Scientific, California, USA) were used to determine the most suitable Assay Sets (**Table 2.2.5**) from the PrimerQuest tool results. Primer and probe sequences were saved in separate text files (.txt).

Table 2.2.4 Optimal criteria for primer and probe characteristics

Characteristic	Primer criteria	Probe criteria
Melting temperature	60 °C – 64 °C	6 °C – 8 °C higher than primers
Annealing temperature	No more than 5 °C below melting temperature	No more than 5 °C below lowest primer melting temperature
GC content	35% – 65%	
Amplicon length	No more than 150 base pairs	

Table 2.2.5 Selected Assay Sets for primers and probes (including the fluorophores and quenchers) of target genes

	Sequence (5' to 3')	Length	Tm	GC%
Assay Set for African Category small toxic peptide gene (Amplicon length: 82 bp)				
Forward primer	TTGTATTGCGGTGGTACTCATA	22	62	40.9
Probe	56-FAM -TTACGCGGTTTCGTTATGCGAGCTA- 3IABkFQ	24	68	50
Reverse primer	AAACTCCGCACCTCCTAATC	20	62	50
Assay Set for East African Clade antitoxin pemI gene (Amplicon length: 101 bp)				
Forward primer	CTGTCGCTGGGTACAGATAATG	22	62	50
Probe	56-FAM -TGATAATGGCCGGCTGATTGTGGA- 3IABkFQ	24	68	50
Reverse primer	AACAGCTCAGCCAGTGAATAC	21	62	47.6
Assay Set for Global/ Global Outlier Category TIR protein domain gene (Amplicon length: 126 bp)				
Forward primer	CTCGGTTTGAGTTGTTGTTT	21	61	42.9
Probe	5CY5 -TGA CTGCTAGAGAGATGAGCGGTGA- 3IAbRQSp	25	68	52
Reverse primer	CGTGCCAGATAGGCAGTATTA	21	61	47.6
Assay Set for Global Clade Chitinase Class I gene (Amplicon length: 85 bp)				
Forward primer	TTTCTGTCAGCCAGTCCATTC	21	62	47.6
Probe	5CY5 -TGC GTTACACGGACAACATCACCT- 3IAbRQSp	24	68	50
Reverse primer	TACGTGGTTGCCTGATGTATTC	22	62	45.5

2.2.9. *In silico* PCR assay

CLC Genomics Workbench was used to perform *in silico* PCR assay. Primer and probe sequences were imported using the Standard Import tool. The tool was then used to run *in silico* PCR assay on each of the twelve validation panel Whole-genome sequences. The primers and probes that annealed to and amplified only the desired gene region were selected for use in the real-time PCR assay (Supplementary Table 1).

2.2.10. Preparation of primers and probes

Inqaba Biotechnology Industries (Pretoria, South Africa) synthesized the primers designed in this study. Fluorogenic probes with quenchers were synthesized by Integrated DNA Technologies (Iowa, USA). Primers and probes were received as desiccated oligonucleotides in 2 µl Safe-Lock tubes (Eppendorf, Hamburg, Germany). Pellets were resuspended in nuclease-free water (Ambion, ThermoFisher Scientific, California, USA), according to the volumes recommended by the relative manufacturing companies, for a final concentration of 100 µM. Primary stock solutions were then stored in a sterile, plastic container at -20 °C.

2.2.11. Crude genomic DNA extraction of validation panel isolates

The ten Liverpool validation panel isolates were received at the CED in the recommended packaging of Dangerous Goods in accordance with UN3373 regulations. The *Salmonella* Enteritidis bacterial colonies were received on solid transport media. The two additional validation panel isolates were stored in 500 µL Tryptic Soy broth (TS broth) (1 L distilled water, 17 g casein, 5 g sodium chloride, 3 g soytone, 2.5 g dextrose, 2.5 g dipotassium phosphate, adjusted to pH 7.3) at -70 °C in plastic containers marked with the year of sample collection. Approximately 10 µl of culture was collected from the transport media or TS broth using a 10 µL plastic classic Microloop (Medical Wire and Equipment Co. Ltd, Wiltshire, UK). Collected culture was then streaked on 5 % blood agar (Diagnostic Media Products, Johannesburg, South Africa) and incubated overnight in an IN 750 incubator (Mettler, Schwabach, Germany) set at 37 °C. Single colonies were picked from the plate and placed in 400 µl of 10 X TE buffer (800 mL distilled water, 2.92 g Tris, 15.76 g EDTA (pH 8)) in a 1.2 mL Safe-Lock tube (Eppendorf, Hamburg, Germany). The solution was then thoroughly mixed using a Vortex-Genie 2 (Lasec SA (Pty) Ltd., Midrand, South Africa) and boiled in a Dry Bath Plus heating block (Lasec SA (Pty) Ltd., Midrand, South Africa) set at 95 °C for 25 minutes. Boiled cell solutions were then centrifuged using the Centrifuge 5415R (Eppendorf, Hamburg, Germany) at 22 673 g and MiniSpin (Eppendorf, Hamburg, Germany) at 11 337 g. From the supernatant, 20 µl was transferred into 80 µl of 10 X TE buffer. The concentrations, 260 nm and 280 nm readings of each crude DNA solution was determined using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, California, USA). Crude DNA of the 12 validation panel

isolates was used throughout this study to validate the primers and probes and as positive controls for the South African isolate screening.

2.2.12. Validation of real-time PCR assay

All primers and probes were diluted to a concentration of 20 μ M using 500 μ L of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) and 100 μ L of the primary stock to make up a final volume of 600 μ L. These were used as working solutions. A real-time PCR assay was then set up using 25 μ L of TaqMan Gene Expression Master Mix (Applied Biosystems, ThermoFisher Scientific, California, USA), 17.8 μ L of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA), 2.5 μ L of the forward primer, 2.5 μ L of the reverse primer and 1 μ L of the relevant probe in each well of the MicroAmp® Optical 96-well reaction plate (Applied Biosystems, Life Technologies™, California, USA). Approximately 1.2 μ L of crude DNA from each of the 12 validation panel isolates was then transferred into four wells, each well containing a set of primers and a probe for amplification of the four different genes (**Table 2.2.6**). The wells were then sealed with a MicroAmp® Optical Adhesive Film (Applied Biosystems, Life Technologies™, California, USA) and placed in the Allegra™ X-22R Centriuge (Beckman Coulter™, California, USA) to ensure all reagents were concentrated at the bottom of the wells. The plate was then loaded into the drawer of the 7500 Real Time PCR System (Applied Biosystems, Life Technologies™, California, USA). Using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA), the relevant probes were selected, the plate layout was labelled and the PCR run was initiated using the conditions in **Table 2.2.7**. This test was repeated three times. Real-time PCR results from all assays in this study were taken at a set threshold of 0.002 delta Rn (fluorescence of a reporter dye divided by the fluorescence of a passive reference dye). Positive cycle threshold (Ct) values were designated as any Ct value less than 30. Ct values equal to or greater than 30 were negative.

Table 2.2.6 Set up of validation of primers and probes with validation panel isolates in a 96-well plate

African Category (Primers and Probes)	A1 NTC	A2 P125 109	A3 A1636	A4 1320	A5 D7795	A6 255	A7 6396	A8 10136 /01	A9 0527 /01	A10 8078	A11 791	A12 672246
	B1 NTC	B2 672632	B3	B4	B5	B6	B7	B8	B9	B10	B11	B12
East African Clade (Primers and probes)	C1 NTC	C2 P125 109	C3 A1636	C4 1320	C5 D7795	C6 255	C7 6396	C8 10136 /01	C9 0527 /01	C10 8078	C11 791	C12 672246
	D1 NTC	D2 672632	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
Global/ Global Outlier Category (Primers and Probes)	E1 NTC	E2 P125 109	E3 A1636	E4 1320	E5 D7795	E6 255	E7 6396	E8 10136 /01	E9 0527 /01	E10 8078	E11 791	E12 672246
	F1 NTC	F2 672632	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
Global Category (Primers and Probes)	G1 NTC	G2 P125 109	G3 A1636	G4 1320	G5 D7795	G6 255	G7 6396	G8 10136 /01	G9 0527 /01	G10 8078	G11 791	G12 672246
	H1 NTC	H2 672632	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

Table 2.2.7 7500 Real-time PCR System run conditions

Stage	Number of Cycles	Temperature (°C)	Time per cycle (Minutes)
1	1	50.0 °C	2:00
2	1	95.0 °C	10:00
3	40	95.0 °C	0:15
	∞	60.0 °C	1:00

Using the real-time PCR assay results, the sensitivity (proportion of correctly identified positive isolates from validation panel) of the test was calculated as follows:

$$\text{Percentage Sensitivity} = \frac{\text{No. of true positives}}{\text{No. of true positives} + \text{No. of false negatives}} \times 100 \quad \text{Equation 2.2.1}$$

Specificity (proportion of correctly identify negative isolates from validation panel) of the test was calculated as follows:

$$\text{Percentage Specificity} = \frac{\text{No. of true negatives}}{\text{No. of true negatives} + \text{No. of false positives}} \times 100 \quad \text{Equation 2.2.2}$$

2.2.13. Optimisation of primer concentrations for singleplex real-time PCR assay

To determine the optimal concentration of primers required to amplify the selected gene region, four validation panel isolates representing each of the four clades (Isolates: P125109, D7795, 10136, 791) were used. Dilutions of 15 µM, 10 µM, 5 µM and 1 µM concentrations were prepared for each of the four sets of primers (forward and reverse) using the 20 µM working solutions and nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) (**Table 2.2.8**). A real-time PCR assay was then set up using 25 µL of TaqMan Gene Expression Master Mix (Applied Biosystems, Life Technologies™, California, USA), 17.8 µL of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA), 1 µL of the relevant probe and 1.2 µL of crude DNA in each well of the MicroAmp® Optical 96-well reaction plate. Additionally, 5 µL of the four 15 µM primer set dilutions was added to each well, so that each of the four validation panel isolates received each set of primers (**Figure 2.2.2**). The same was done for the 10 µM, 5 µM and 1 µM dilutions. The wells were then sealed with a MicroAmp® Optical Adhesive Film (Applied Biosystems, Life Technologies™, California, USA) and placed in the Allegra™ X-22R Centriuge (Beckman Coulter™, California, USA) to ensure all reagents were concentrated at the bottom of the wells. The plate was then loaded into the drawer of the 7500 Real Time PCR System (Applied Biosystems, Life Technologies™, California, USA). Using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA), the relevant probes were selected, the plate layout was labelled and the PCR run was initiated using the conditions in **Table 2.2.7**. Results from the real-time PCR assay were analysed using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA) and the lowest primer concentrations yielding Ct values between 20 and 25 were accepted for use in further reactions.

Table 2.2.8 Volumes of 20 μ M primer Stock Solution and nuclease-free water required to make up different concentrations of primer

Concentration (μ M)	Volume of 20 μ l Stock Solution needed (μ l)	Volume of nuclease-free water needed (μ l)
15	15	5
10	10	10
5	5	15
2	2	18
1	1	19

	15 μ M		10 μ M		5 μ M		2 μ M		1 μ M		
A1 NTC	A2 P125 109	A3 P125 109	A4 P125 109	A5 P125 109	A6 P125 109	A7 P125 109	A8 P125 109	A9 P125 109	A10 P125 109	A11 P125 109	A12
B1 NTC	B2 D7795	B3 D7795	B4 D7795	B5 D7795	B6 D7795	B7 D7795	B8 D7795	B9 D7795	B10 D7795	B11 D7795	B12
C1 NTC	C2 10136 /01	C3 10136 /01	C4 10136 /01	C5 10136 /01	C6 10136 /01	C7 10136 /01	C8 10136 /01	C9 10136 /01	C10 10136 /01	C11 10136 /01	C12
D1 NTC	D2 791	D3 791	D4 791	D5 791	D6 791	D7 791	D8 791	D9 791	D10 791	D11 791	D12
E1 NTC	E2 P125 109	E3 P125 109	E4 P125 109	E5 P125 109	E6 P125 109	E7 P125 109	E8 P125 109	E9 P125 109	E10 P125 109	E11 P125 109	E12
F1 NTC	F2 D7795	F3 D7795	F4 D7795	F5 D7795	F6 D7795	F7 D7795	F8 D7795	F9 D7795	F10 D7795	F11 D7795	F12
G1 NTC	G2 10136 /01	G3 10136 /01	G4 10136 /01	G5 10136 /01	G6 10136 /01	G7 10136 /01	G8 10136 /01	G9 10136 /01	G10 10136 /01	G11 10136 /01	G12
H1 NTC	H2 791	H3 791	H4 791	H5 791	H6 791	H7 791	H8 791	H9 791	H10 791	H11 791	H12

African Category primers and probe

East African Clade primers and probe

Global/ Global Outlier Category primers and probe

Global Clade primers and probe

Figure 2.2.2 96-well plate layout for optimal primer concentrations to amplify clade-defining genes in *Salmonella* Enteritidis isolates in a real-time PCR assay

2.2.14. Optimisation of probe concentrations for singleplex real-time PCR assay

To determine the optimal concentration of probes required to amplify the selected gene region, four validation panel isolates representing each of the four clades (Isolates: P125109, D7795, 10136, 791) were used. Dilutions of 5 μ M and 2 μ M concentrations were prepared for each of the four probes using the 20 μ M working solutions and nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) (**Table 2.2.9**). A real-time PCR assay was then set up using 25 μ L of TaqMan Gene Expression Master Mix, 17.8 μ L of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA), 5 μ L of the relevant primer set (concentration determined in step 2.14) and 1.2 μ L of crude DNA in each well of the MicroAmp® Optical 96-well reaction plate. Additionally, 1 μ L of each of the four 5 μ M probe dilutions was added to four wells, so that each validation panel isolates received every primer (**Figure 2.2.3**). The same was done for the 2 μ M dilutions. The wells were then sealed with a MicroAmp® Optical Adhesive Film (Applied Biosystems, Life Technologies™, California, USA) and placed in the Allegra™ X-22R Centriuge (Beckman Coulter™, California, USA) to ensure all reagents were concentrated at the bottom of the wells. The plate was then loaded into the drawer of the 7500 Real Time PCR System (Applied Biosystems, Life Technologies™, California, USA). Using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA), the relevant probes were selected, the plate layout was labelled and the PCR run was initiated using the conditions set out in **Table 2.2.7**. Results from the real-time PCR were analysed using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA) and the lowest probe concentrations yielding Ct values between 22 and 25 were accepted for use in further reactions.

Table 2.2.9 Volumes of 20 μ M probe Stock Solution and nuclease-free water required to make up different concentrations of probe

Concentration (μ M)	Volume of 20 μ M Stock Solution needed (μ L)	Volume of nuclease-free water needed (μ L)
5	5	15
2	2	18

5 μ M	A1 NTC	A2 P125 109	A3 D7795	A4 10136	A5 791	A6	A7	A8	A9	A10	A11	A12
	B1 NTC	B2 P125 109	B3 D7795	B4 10136	B5 791	B6	B7	B8	B9	B10	B11	B12
2 μ M	C1 NTC	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
	D1 NTC	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12
	E1 NTC	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
	F1 NTC	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12
	G1 NTC	G2	G3	G4	G5	G6	G7	G8	G9	G10	G11	G12
	H1 NTC	H2	H3	H4	H5	H6	H7	H8	H9	H10	H11	H12

 African Category primers and probe	 East African Clade primers and probe
 Global/ Global Outlier Category primers and probe	 Global Clade primers and probe

Figure 2.2.3 96-well plate layout for optimal probe concentrations to amplify clade-defining genes in *Salmonella* Enteritidis isolates in a real-time PCR assay

2.2.15. Master mixes for multiplex real-time PCR assays

Using the optimal primer and probe concentrations determined in the previous steps (step 2.2.13 and 2.2.14 respectively), four master mixes for the two multiplex real-time PCR assays (category-classifying and clade-classifying) were made. The constituents of Master Mix One through Four are summarized in Table 2.2.10

Table 2.2.10 Constituents of the four Master Mixes used in the category-classifying (Master Mix One and Two) clade-classifying (Master Mix Three and Four) real-time PCR

	Master Mix One	Master Mix Two	Master Mix Three	Master Mix Four
	African Category	Global/ Global Outlier Category	East African Clade	Global Clade
Volume added (μ l)				
Forward primer	20	50	20	20
Reverse primer	20	50	20	20
Probe	50	50	50	50
Nuclease-free water	110	50	110	110

2.2.16. Pure whole-genome DNA extraction from isolates using kits

The QIAamp DNA Mini Kit (QIAGEN, Hilden, Germany) was used to extract pure genomic DNA from the validation panel isolates according to the instructions provided by the manufacturers. DNA concentrations, 260 nm and $^{260 \text{ nm}} / ^{280 \text{ nm}}$ readings were determined using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, California, USA).

2.2.17. Analytical verification of the multiplex real-time PCR assay

To determine the efficiency of the multiplex real-time PCR assay, serial dilutions of pure DNA solutions from four validation panel isolate were prepared. This was done by diluting the pure DNA of each isolate with nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) to a final concentration of 100 μM in a final volume of 50 μL (**Table 2.2.11**). Volumes of DNA and nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) required for each dilution was calculated using the following formula:

$$(Initial \text{ Concentration})(Initial \text{ Volume}) = (Final \text{ Concentration})(Final \text{ Volume}) \quad \text{Equation 2.2.3}$$

NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, California, USA) measurements gave the initial concentrations of the pure DNA. From each stock solution, 1 μL was transferred to a new 0.5 mL Safe-Lock tube (Eppendorf, Hamburg, Germany) with 9 μL of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) to make up a DNA concentration of 10 μM . Further dilutions were made by transferring 1 μL of each dilution into a new 0.5 mL Safe-Lock tube with 9 μL of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA) until the final concentration of DNA was 0.001 μM . The final 1 μL taken from this dilution was then discarded (**Figure 2.2.4**). The concentration of each dilution in the series was then measured using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, California, USA).

Table 2.2.11 Volume of water and DNA required for each validation panel isolate to obtain a final DNA concentration of 100 μM in a volume of 50 μL

Validation panel isolate	Initial DNA Concentration (ng/ μL)	Volume DNA needed (μL)	Volume nuclease-free water (μL)
791	151.5	33.00	17.00
10136/01	110.6	45.21	4.79
255	134.3	37.23	12.77
P125109	154.5	32.36	17.64

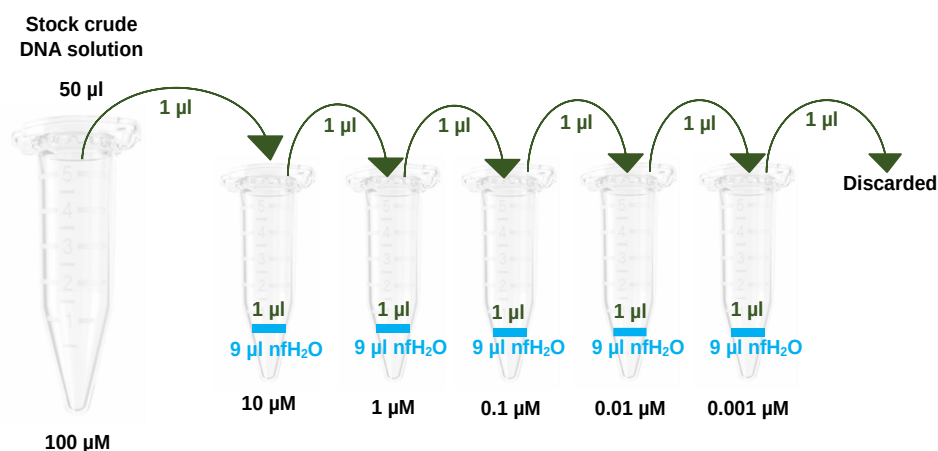


Figure 2.2.4 Serial dilutions of pure DNA used to test the efficiency of the primers and probes in the clade-classifying PCR assay

A real-time PCR assay was then set up using 25 µL of TaqMan Gene Expression Master Mix, 17.8 µL of nuclease-free water (Ambion, ThermoFisher Scientific, California, USA), 6 µL of the respective Master Mix (**Table 2.2.10**) and 1.2 µL of each DNA dilution to each well of the MicroAmp® Optical 96-well reaction plate (Applied Biosystems, ThermoFisher Scientific, California, USA). The wells were then sealed with a MicroAmp® Optical Adhesive Film (Applied Biosystems, Life Technologies™, California, USA) and placed in the Allegra™ X-22R Centrifuge (Beckman Coulter™, California, USA) to ensure all reagents were concentrated at the bottom of the wells. The plate was then loaded into the drawer of the 7500 Real Time PCR System (Applied Biosystems, Life Technologies™, California, USA). Using the 7500 Real Time PCR System software (Applied Biosystems, Life Technologies™, California, USA), the relevant probes were selected, the plate layout was labelled and the PCR run was initiated using the conditions in **Table 2.2.7**. The DNA concentration yielding the highest Ct value below 30 was determined to be the limit of detection for that particular primer and probe set. All Ct values from this run were plotted against the logarithm of DNA concentration. The slope of the resulting graph was then used to determine the efficiency:

$$Efficiency = 10^{-1/slope} - 1 \quad \text{Equation 2.2.4}$$

To calculate the amplification factor of each primer and probe set, the percentage efficiency was entered into the ThermoFisher Scientific online qPCR Efficiency Calculator (252). To determine the accuracy with which the Ct value could be used to quantify the number of genes in a sample, the determination coefficient (R^2) was calculated for the plot points of the triplicate assays for each primer and probe set using the CORREL function in Microsoft Excel 2010.

2.3. Results

2.3.1. Whole-Genome Sequence analysis of the validation panel isolates

Analysis of the annotated whole-genome sequences of the 12 validation panel isolates determined that the pan genome consisted of 5 321 genes. Approximately 4 174 (78.44%) of these genes constituted the core genome and were common to all validation panel isolates. Shell genes (628 (11.80%) genes) were shared between 15% and 95% of the isolates. Cloud genes (the remaining 519 (9.75%) genes) were shared by less than 15% of the isolates.

2.3.2. Selection of category- and clade-specific genes

Using the gene_presence_absence matrix, category- and clade-specific shell genes were selected based on the specified criteria. For the category-specific multiplex assay, the small toxic peptide was selected to identify the African category since it was shared among isolates from both West African (10136/01, 0527/01, 8078/01) and East African (D7795, 255, 6396) clades. To identify the Global/Outlier category, the TIR protein domain gene was selected since it was present in isolates from both the Global (P125109, A1636, 1320) and Outlier (791, 672246, 672632) clades. For the clade-specific multiplex assay the *PemI* gene was chosen to identify the East African clade and the Chitinase Class I gene was selected to identify the Global clade since each respective gene was exclusive to the isolates in each clade. Categorisation of *Salmonella* Enteritidis isolates by category- and clade-specific genes is summarised in **Figure 2.3.1**.

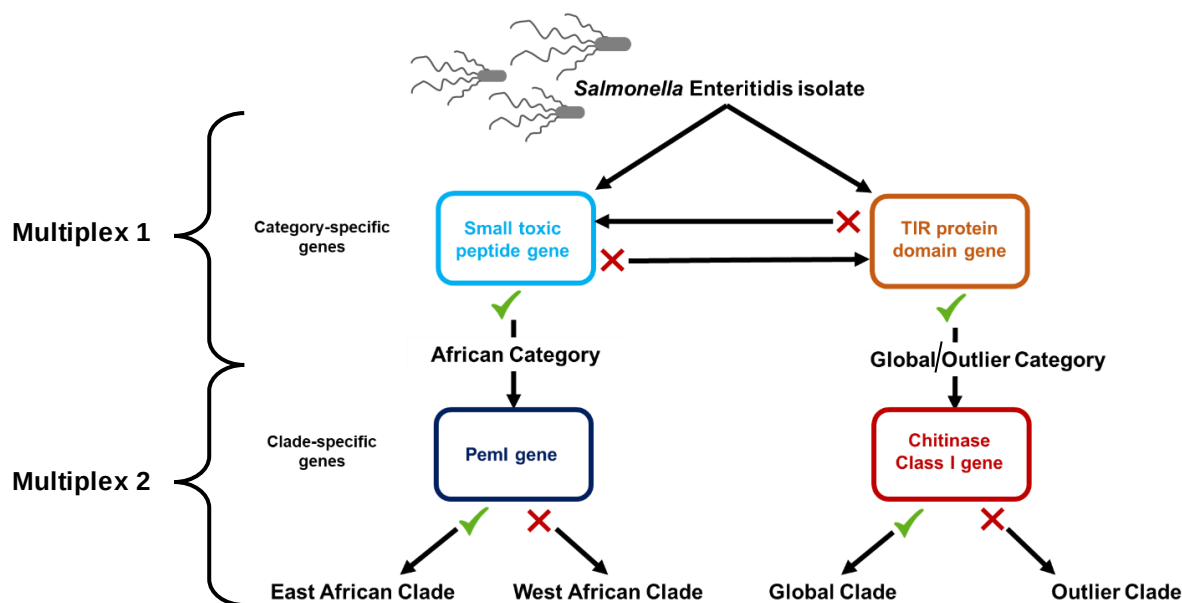


Figure 2.3.1 Workflow depicting the clade classification of a *Salmonella* Enteritidis isolate based on the presence (green tick) or absence (red cross) of genes targeted by the clade-classifying real-time PCR assay. Diagram created by the author, Sarah Gallichan

2.3.3. *In silico* primer and probe validation

Primers and probes were designed to amplify the selected genes using the online PrimerQuest tool. In order to validate the primer and probe sets, *in silico* PCR was performed on the whole-genome sequences of the 12 validation panel isolates using CLC Workbench software. The African primer set amplified an 82 base pair (bp) fragment of the small toxic peptide gene for all six African category isolate sequences (10136/01, 0527/01, 8078/01, D7795, 255, 6396) (**Supplementary Table 1**). The Global/Outlier category primer set amplified a 126 bp fragment of the TIR protein domain gene for all six Global/Outlier category isolate sequences (P125109, A1636, 1320, 791, 672246, 672632) (**Supplementary Table 1**). The East African clade primer set amplified a 101 bp fragment of the small toxic peptide gene from the East African clade isolate sequences (D7795, 255, 6396) (**Supplementary Table 1**). The Global clade primer set amplified an 85 bp fragment of the Chitinase Class I gene from the Global clade isolate sequences (P125109, A1636, 1320) (**Supplementary Table 1**). The corresponding probes bound specifically within the amplified region for each target gene (**Supplementary Table 2**). Thus, the designed primers and probes were accepted for synthesis and use in practical real-time PCR validation.

2.3.4. Sensitivity and specificity of the real-time PCR assay

Practical validation of the primer and probe sets was done using the 12 validation panel isolates to perform three technical replicates of the real-time PCR assays. The category- and clade-specific primer and probe sets were calculated to be 100% sensitive and specific for the respective target genes. No false positive or false negative results were recorded for any of the assays (**Table 2.3.1**).

Table 2.3.1 Average Ct values from the three technical replicate real-time PCR assays performed using the validation panel isolates. Target gene Ct values shaded in green indicate a positive result (Ct<30) and values in red indicate a negative result

Target genes						
Isolate name	Expected Clade	Small toxic peptide	TIR protein domain	Antitoxin pemi	Chitinase class 1	Final clade result (as per rtPCR)
P125109	Global	-	19.31	-	17.72	Global
A1636	Global	-	19.26	-	18.12	Global
1320	Global	-	18.54	-	16.67	Global
791	Global Outlier	-	19.90	-	-	Global Outlier
672246	Global Outlier	-	17.67	-	-	Global Outlier
672632	Global Outlier	-	18.25	-	-	Global Outlier
D7795	East African	17.66	-	16.92	-	East African
255	East African	18.12	-	16.81	-	East African
6396	East African	17.05	-	16.60	-	East African
10136/01	West African	18.59	-	-	-	West African
0527/01	West African	18.74	-	-	-	West African
8078	West African	20.26	-	-	-	West African

2.3.5. Primer and probe concentration optimisation

The optimal primer concentrations were determined by performing the real-time PCR assays with differing primer concentration (15 μ M, 10 μ M, 5 μ M, 2 μ M, 1 μ M). Primers for the African category, Global and East African clade genes performed optimally at a concentration of 2 μ M (**Supplementary Figure 1**). The Global/ Outlier category primers performed optimally at 5 μ M (**Supplementary Figure 2**). The optimal probe concentrations were determined by performing the real-time PCR assays with differing primer concentration (5 μ M, 2 μ M, 1 μ M). Probes for all category- and clade-specific assays performed optimally at a concentration of 5 μ M. The use of primers and probes at optimal concentrations resulted in true positives with Ct values between 22 and 25.

2.3.6. Analytical verification of primers and probe sets within multiplex assays

Note on the use of ThermoFisher Scientific guidelines: Within this dissertation, ThermoFisher Scientific guidelines (**reference for handbook**) were used as standards for the optimal performance of the real-time PCR assay. These guidelines were chosen since ThermoFisher Scientific developed the real-time PCR system used throughout this study.

To determine the efficiency of the primer and probe sets designed for this test, three technical replicates of 10-fold dilutions were performed on two validation panel isolates. Results from these tests were used to plot linear regressions from which the efficiencies of each primer and probe set could be resolved. According to the ThermoFisher Scientific real-time PCR standards, an assay should perform with an efficiency between 90.00% and 110.00%. The category multiplex assay that contained African and Global/ Outlier primer and probe sets performed at an average efficiency of 105.82% and 100.92% respectively (**Figure 2.3.2**). The clade multiplex assay that contained East African clade and Global clade primer and probe sets performed at an average efficiency of 106.76% and 107.23% respectively (**Figure 2.3.3**). Thus, all primer and probe sets designed in this study ran at optimal efficiency within the multiplex assays.

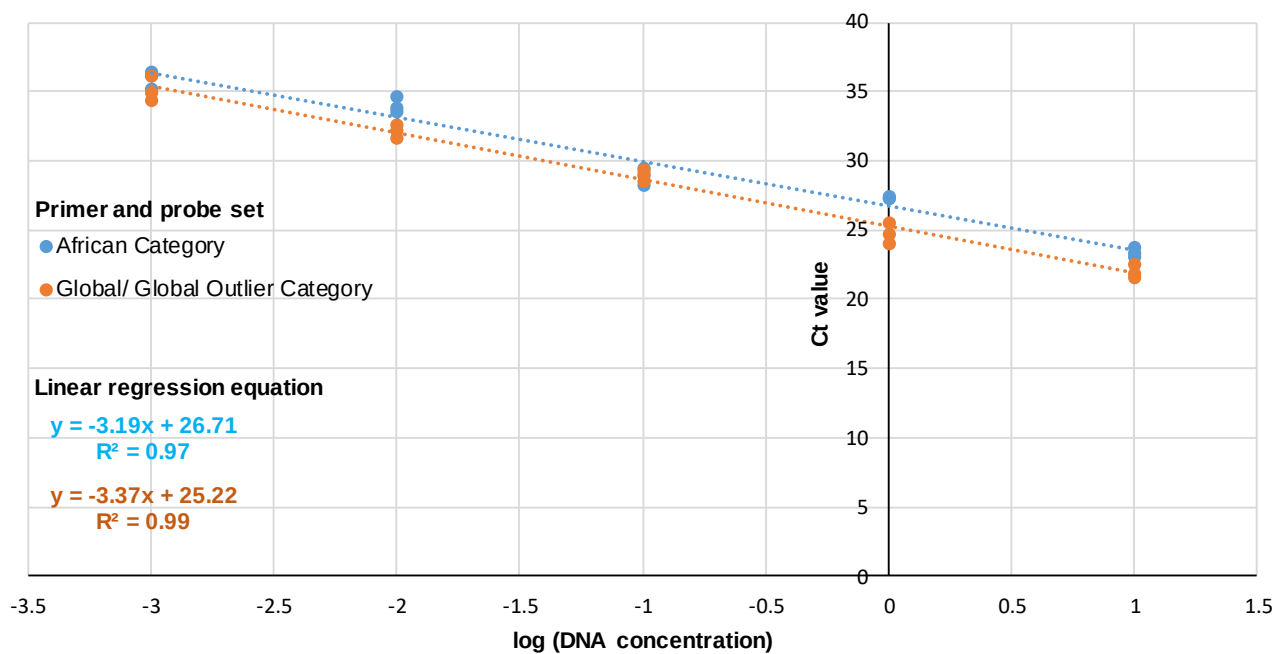


Figure 2.3.2 Logarithmic efficiency plot depicting the change in cycle threshold value with the change in DNA concentration for the African and Global/ Outlier category primer and probe sets within the category-specific real-time PCR assay

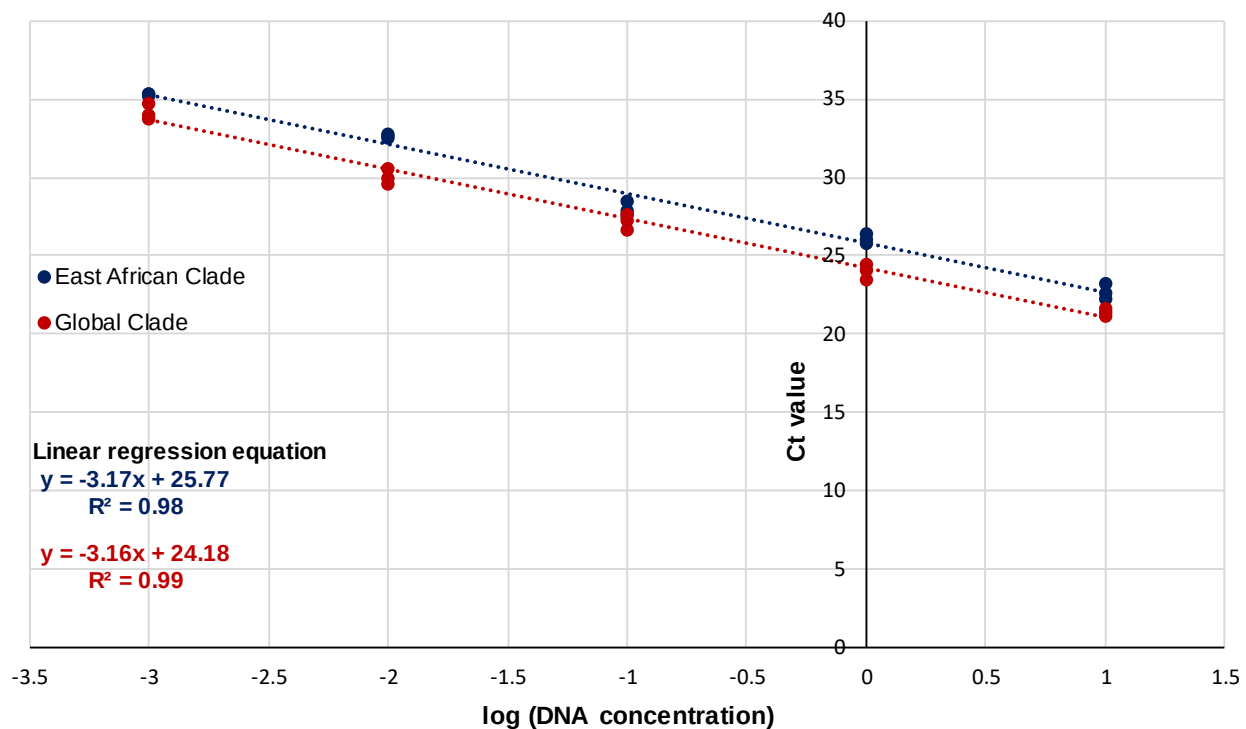


Figure 2.3.3 Logarithmic efficiency plot depicting the change in cycle threshold value with the change in DNA concentration for the East African and Global clade primer and probe sets within the clade-specific real-time PCR assay

Theoretically, a real-time PCR performing at 100% efficiency will result in the duplication of the target gene with every cycle. The amplification factor for each primer and probe set was calculated using the percentage efficiencies. The category multiplex assay that contained African and Global/ Outlier primer and probe sets had an approximate amplification factor of 2.06 and 2.01 genes per cycle respectively. The clade multiplex assay that contained East African clade and Global clade primer and probe sets both had an estimated amplification factor of 2.07 genes per cycle.

The linear association between the cycle threshold value (Ct value) and log-transformed DNA concentration determined the accuracy with which the target gene quantity can be mathematically deduced. Linear association was measured using the determination coefficient (R^2). ThermoFisher Scientific real-time PCR standards suggest that real-time PCR efficiency assays have an $R^2 \geq 0.98$. The category multiplex assay that contained African and Global/ Outlier primer and probe sets had determination coefficient values of 0.98 and 1.00 respectively (**Figure 2.3.2**). The clade multiplex assay that contained East African clade and Global clade primer and probe sets both had determination coefficients of 0.99 (**Figure 2.3.3**). Thus, determination coefficients for all primer and probe sets complied with the required $R^2 \geq 0.98$.

Using the dilution series, the limit of detection was determined as the lowest DNA concentration resulting in a true positive (Ct value below 30). The limit of detection for these assays was 0.1 μM (**Table 2.3.2**). Although the Global clade primer and probe set was capable of amplifying genes from a 0.01 μM DNA solution, the lack of technical reproducibility meant the true limit of detection was 0.1 μM as with the other primer and probe sets.

Table 2.3.2 Ct values from the three technical replicate real-time PCR assays performed with varying DNA concentrations. Values shaded in green indicate a positive result (Ct<30) and values in red indicate a negative result (Ct≥30)

African category primer and probe set					
Crude DNA concentration (μM)					
Replicate	10	1	0.1	0.01	0.001
1	23.7	27.32	28.28	34.6	36.14
2	23.02	27.42	28.94	33.86	36.45
3	23.27	27.27	29.45	33.54	35.36

Global/ Global Outlier category primer and probe set					
Crude DNA concentration (μM)					
Replicate	10	1	0.1	0.01	0.001
1	22.49	25.52	29.07	32.56	36.14
2	21.54	24.76	28.53	31.64	34.94
3	21.82	24.05	29.29	32.16	34.37

East African clade primer and probe set					
Crude DNA concentration (μM)					
Replicate	10	1	0.1	0.01	0.001
1	22.29	25.82	27.6	32.75	35.36
2	23.25	26.38	27.87	32.6	35.2
3	22.66	26.03	28.43	32.52	35.33

Global clade primer and probe set					
Crude DNA concentration (μM)					
Replicate	10	1	0.1	0.01	0.001
1	21.08	23.4	26.62	29.61	33.73
2	21.6	24.09	27.26	29.97	33.96
3	21.41	24.47	27.58	30.56	34.67

2.4. Discussion

Successful amplification of the relevant target genes was observed for all 12 validation panel isolates, and each isolate was classified into the appropriate clade. These results corresponded with the hierBAPS clade classification predicted from whole-genome sequence data of a large collection of *Salmonella* Enteritidis isolates from which the validation panel were selected ((218), Supplementary Table 1). Furthermore, the amplicon sizes defined by the PrimerQuest Assay set corresponded with those predicted by the *in silico* PCR assay. Therefore, together with the calculated specificity of 100%, we have successfully demonstrated that this clade-classifying PCR assay is highly specific in targeted amplification of the selected genes from *Salmonella* Enteritidis DNA extracts.

Both technical and biological reproducibility was shown throughout the clade-classifying real-time PCR assay validation process. Optimisation of the primer and probe concentrations for each set contributed greatly to the robustness of the assay since this step ensured the optimal amplification of target genes within the true positive range (i.e. Ct <30). Consequently, the assays had high efficiencies and optimal assay precision across technical replicates. Biological reproduction was also achieved for each primer and probe set, with the successful amplification of each target gene in three validation panel isolates. Additionally, true positive results were technically reproduced from the lowest DNA concentration (i.e. limit of detection) across four biologically different isolates.

Evidently, the assay was capable of accurately amplifying 0.1 ng/μL of pure whole-genome DNA extract in all biological and technical replicates. Optimal efficiencies observed for all primer and probe sets demonstrated a strong linear relationship between the logarithmic concentration of the DNA and the Ct value. Thus, the target gene was amplified and accurately detected for a range of DNA concentrations (10 ng/μL to 0.001 ng/μL), indicating a high sensitivity. This corresponded well with the calculated sensitivity of 100%.

It must be noted that this clade-classifying assay is limited to previously confirmed *Salmonella* Enteritidis isolates. Although the assay is highly specific to *Salmonella* Enteritidis, it is unsuitable for molecular subtyping of *Salmonella* serotypes. The necessity for two molecular tests adds additional cost and processing time per isolate, which is not ideal for implementation into routine surveillance laboratories. However, this problem could be solved through the addition of primers and probes to the multiplex real-time PCR, which amplify a *Salmonella* Enteritidis specific gene.

We have effectively demonstrated that this clade-classifying real-time PCR assay is a sensitive, specific and highly accurate gene-profiling assay for the discrimination of *Salmonella* Enteritidis clades. In doing so, the first two study objectives (identification of cargo genes for use as molecular markers and the development of a sensitive multiplex real-time PCR) were successfully met. Additionally, by fulfilling the

optimisation and validation requirements outlined in literature (253,254), this clade-classifying assay could readily be implemented in both research and public health laboratories. This method provides a rapid and sensitive test capable of distinguishing between otherwise genetically homogenous clades.

Improving the availability of high quality molecular tools in developing countries is a global priority (255). However, due to the resource limitations within these countries the implementation of such molecular tools has proven difficult. Most serological diagnostic tools, especially those used for NTS, have high running costs, present significant challenges in maintaining quality and often have low sensitivity and specificity (31). At present, equipment costs and specialised personnel necessary for the implementation of WGS hinders the availability of such tools in LMIC (256,257). Thus, the low running cost (approximately R8.45 or US\$0.59 per sample) together with the high quality of this clade-classifying real-time PCR assay makes this an ideal tool for use in public health laboratories, especially where WGS is not readily available.

CHAPTER THREE

Epidemiological investigation of *Salmonella* Enteritidis clades in South Africa

3.1. Introduction

Non-typhoidal *Salmonella* (NTS) is among the most commonly isolated foodborne pathogens worldwide (31). Annually, approximately 93 million infections and 155 000 deaths occur due to gastroenteritis (67). The major impact salmonellosis has on animal, human and environmental health sectors has resulted in the placement of NTS on a list of zoonotic diseases of most concern in the United States (US) (258). Furthermore, the World Health Organization has prioritized the strengthening of capacity for surveillance of foodborne pathogens globally (65).

Whilst *Salmonella* gastroenteritis accounts for the bulk of morbidity due to NTS, invasive NTS (iNTS) disease is of greater concern in terms of mortality, resulting in an estimated 4.26 million Disability Adjusted Life Years (DALYs) (181). iNTS disease is a particular threat in sub-Saharan Africa (sSA), which accounts for more than half of the globally estimated iNTS cases (178,181). It is thought that the large immunocompromised population within the sSA region has resulted in the adaptation of NTS to a more invasive lifestyle (194,218). The high prevalence of iNTS disease in this region, where the majority of the countries classify as low- to lower-middle-income (LMIC), is detrimental to the socioeconomic and healthcare systems (195).

Salmonella Enteritidis is one of the two most common NTS serovars, causing gastroenteritis as well as invasive disease (98,191–193). Phylogenetic analysis of *Salmonella* Enteritidis from around the world has shown that there are four distinct clades in circulation (218). Three of these clades have been implicated in outbreaks of epidemic proportions: the Global clade has been linked with poultry-associated gastroenteritis and the two African clades (East and West African clades) that have been associated with severe invasive disease (218).

Despite the recognition of the distinct *Salmonella* Enteritidis clades in sSA, the prevalence and transmission of these clades within South Africa has not been well described. *Salmonella* Enteritidis has consistently been the most commonly isolated serovar in South Africa since 2011 (198,259,260). The increased consumption and importation of poultry places the country at risk for an increase in foodborne outbreaks associated with the Global clade, as has been observed in the USA, Europe and Brazil (235,239,261,262). Thus, routine surveillance of clade distribution in South Africa together with the

monitoring of changes in clade proportions could assist with more targeted interventions for the prevention of outbreaks.

The aim of this chapter was to gain a retrospective understanding of the prevalence, geographical distribution and characteristics (association with invasive disease and antimicrobial resistance) of *Salmonella* Enteritidis clades in four South African provinces (Gauteng, KwaZulu-Natal, Mpumalanga and Western Cape) over a two year period (2012 and 2013) using the clade-classifying real-time PCR assay. The assay proved an effective tool in the investigation of clade proportions in the South African public health setting.

3.2. Materials and methods

3.2.1. Workflow for the epidemiological investigation of *Salmonella* Enteritidis clades in South Africa

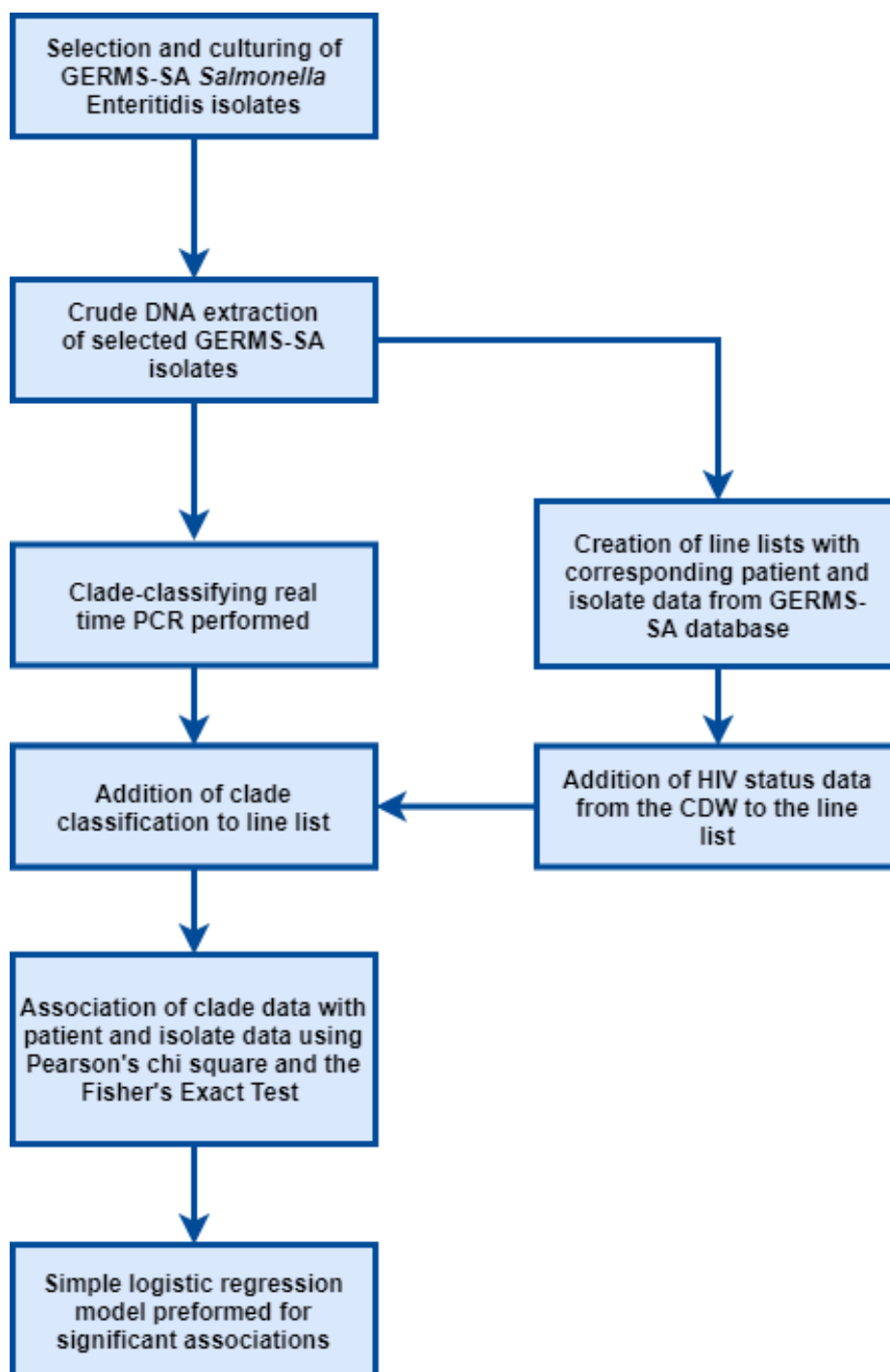


Figure 3.2.1 Flow chart showing the workflow for the statistical analysis of clade-classifying assay results and corresponding patient and isolate data

3.2.2. GERMS-SA laboratory surveillance program

Since 2003, approximately 50 public and private clinical laboratories across all provinces of South Africa have been participating in the GERMS-SA active laboratory-based surveillance program. Pathogens of public health concern (AIDS-related opportunistic infections, epidemic-prone diseases, vaccine-preventable diseases, nosocomial infections) are submitted to the National Institute for Communicable Diseases (NICD) for confirmation and further characterisation. All clinical and epidemiological information collected from laboratory confirmed cases are captured and stored on the GERMS-SA database. The data is regularly reviewed to ensure all referred isolates were compliant with the GERMS-SA case definitions. Approximately 83.1% of the South African population have access to public healthcare and are therefore under surveillance by the GERMS-SA laboratory network.

3.2.3. GERMS-SA non-typhoidal *Salmonella* isolates repository

Among other pathogens, the received laboratory non-typhoidal *Salmonella* isolates from GERMS-SA enhanced surveillance site laboratories were submitted to the CED. These isolates were serotyped using the White-Kauffman-Le Minor scheme and had antimicrobial susceptibility testing performed according to the Clinical and Laboratory Standards Institute (CLSI) guidelines and breakpoints. Minimum inhibitory concentration (MIC) was determined for each isolate against ampicillin, chloramphenicol, trimethoprim-sulfamethoxazole, tetracycline, ciprofloxacin and ceftriaxone using E-strips according to manufacturer's recommendations (Bio-Mérieux, Marcy-l'Etoile, France). These results, as well as the patient demographic and clinical data corresponding to each isolate, were recorded on the Access 2010 (Microsoft Corporation, Redmond) GERMS-SA database. The isolates were then stored at -80°C in Tryptic Soy broth (1 L distilled water, 17 g casein, 5 g sodium chloride, 3 g soytone, 2.5 g dextrose, 2.5 g dipotassium phosphate, adjusted to pH 7.3) in the GERMS-SA repository at the CED, NICD.

3.2.4. Isolate referral bias

It is important to note that the NTS isolates received by the CED represent a subset NTS cases in the South African population. Passive laboratory-based surveillance programs are limited by the health-seeking behaviour of an ill person, prevailing specimen collection practices, diagnostic practices at laboratories, and voluntary submission of isolates (**Figure 3.2.2**) (263). In particular, the difference in stool collection practices when investigating adults versus children suffering from diarrhoea in healthcare facilities affects the number of confirmed NTS gastroenteritis cases reported (264). Furthermore, stool collection is more likely for patients seen in private and tertiary hospitals, which have the available resources. We acknowledge that the invasive and non-invasive NTS surveillance isolates received by the NICD do not represent the entire exposed population and therefore the results from this study could not be used to infer risk factors or incident rates.

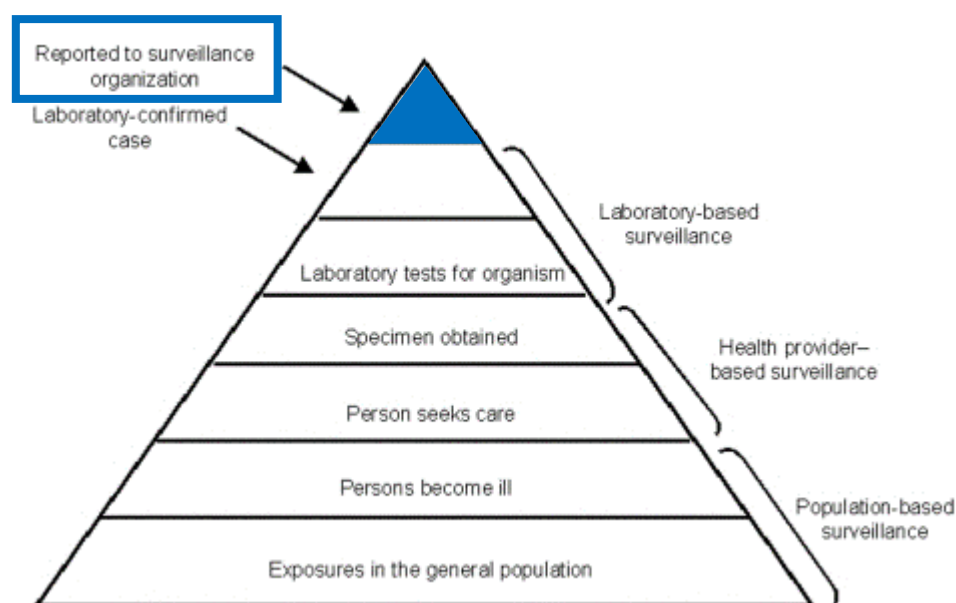


Figure 3.2.2 A diagram representing the proportion of potential pathogens present at different stages. Different types of surveillance programs sample pathogens at different stages, therefore have varying sensitivities and specificities. The blue area of the pyramid represents the proportion of isolates reported to surveillance organizations as with GERMS-SA. (263)

3.2.5. Selection of South African *Salmonella* Enteritidis isolates

Isolates used in this study were selected from the GERMS-SA frozen isolate repository to accurately represent the invasive (isolated from blood and cerebral spinal fluid) and non-invasive (isolated from stool) *Salmonella* Enteritidis isolates referred to the NICD by Gauteng, KwaZulu-Natal, Mpumalanga and Western Cape laboratories for the years 2012 and 2013. During this time period (2012 and 2013) the number of *Salmonella* Enteritidis isolates collected surpassed all other serovars and isolates collected during these two years were therefore of interest. Laboratories from Gauteng, KwaZulu-Natal, Mpumalanga and the Western Cape contributed the majority (94.58%, 1 866/1 973) of all *Salmonella* Enteritidis isolates received in 2012 and 2013. The number of isolates selected from each province (**Table 3.2.1**) was done according to the proportion of total *Salmonella* Enteritidis isolates contributed by the province relative to the other three provinces used in this study. Northern Cape, North West, Limpopo, Eastern Cape and Free State provinces collectively contributed only a small proportion of *Salmonella* Enteritidis isolates (5.42%, 107/1 973) in 2012 and 2013, and were therefore excluded from the study.

Table 3.2.1 Number of invasive and non-invasive *Salmonella* Enteritidis isolates selected from the 2012 and 2013 GERMS-SA archives to represent four South African provinces (Gauteng, Western Cape, KwaZulu-Natal, Mpumalanga)

		2012		2013		Total
	Invasive	Yes	No	Yes	No	
Province	Mpumalanga	9	12	3	11	35
	Kwazulu Natal	35	11	28	11	85
	Gauteng	90	63	83	94	330
	Western Cape	31	67	32	38	168
						618

3.2.6. Crude genomic DNA extraction

Approximately 10 µL of culture was collected from the trypticase soy broth using a 10 µL plastic classic Microloop (Medical Wire and Equipment Co. Ltd, Wiltshire, UK). Collected culture was then streaked on 5% blood agar (Diagnostic Media Products, Johannesburg, South Africa) and incubated overnight in an IN 750 incubator (Mettler, Schwabach, Germany) set at 37°C. Single colonies were collected from the plate and placed in 400 µL of 10 X TE buffer (800 mL distilled water, 2.92 g Tris, 15.76 g EDTA (pH 8)) in a 1.5 mL tube. The solution was then thoroughly mixed using a Vortex-Genie 2 (Lasec SA (Pty) Ltd., Midrand, South Africa) and boiled in a Dry Bath Plus heating block (Lasec SA (Pty) Ltd., Midrand, South Africa) set at 95°C for 25 minutes. Boiled cell solutions were then centrifuged using the Centrifuge 5415R (Eppendorf, Hamburg, Germany) at 22 673 g and MiniSpin (Eppendorf, Hamburg, Germany) at 11 337 g. From the supernatant, 20 µL was transferred into 80 µL of 10 X TE buffer (800 mL distilled water, 2.92 g Tris, 15.76 g EDTA (pH 8)). The concentrations, 260 nm and $^{260\text{ nm}}/_{280\text{ nm}}$ readings of each crude DNA solution were determined using the NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific, California, USA). A minimum DNA concentration of 10 ng/µL was accepted for use in further experimentation.

3.2.7. Classification of isolates into clades using clade-classifying real-time PCR assay

Clade classifying real-time PCR was performed using the crude DNA extracts from each of the selected South African isolates. Two validation panel isolates, from the East African and Global clade, were used as positive controls for each of the target genes in the rtPCR. Approximately 2 µL of crude DNA was loaded in MicroAmp® Optical 96-well reaction plate (Applied Biosystems, Life Technologies™, California, USA) from left to right, omitting the first three wells that were used for a non-template (negative) control (2 µL distilled water included in place of DNA) and the two positive template controls.

3.2.8. Line list of epidemiological data variables for *Salmonella* Enteritidis isolates

A line list containing the *Salmonella* Enteritidis isolate data (identification number, invasive/non-invasive, year of collection and antimicrobial susceptibility) and corresponding patient information (age, gender) from the GERMS-SA database was created in an Excel 2010 (Microsoft Corporation, Redmond) spreadsheet. Age was converted into ordinal age group categories consisting of 5 year groups. Additional information, such as HIV status, was requested from the Central Data Warehouse (CDW) of the National Health Laboratory Service (NHLS). HIV status ("Positive" or "Negative") was derived from a composite of HIV-1/HIV-2 antibody screens, CD4 (cluster of differentiation) counts and the HIV viral load data for each patient. Patients without HIV test data were given the status "Unknown". Lastly, the results from the real-time PCR assay were entered. Data entries were matched based on the unique identity number of each isolate as well as the date of birth of the patient to ensure data quality.

3.2.9. Bivariate statistical analysis of epidemiological data

Epidemiological data was analysed in a case-to-case method where variables associated with Global clade isolates were compared with those associated with Outlier isolates. The aim was to determine whether there were clade-specific characteristics. All epidemiological data variables for isolates were analysed using Stata version 14 (StataCorp LD, Texas, USA) (**Supplementary Figure 3**). The line list of data variables was imported from Excel into Stata, using the first row as variable name. Data for the isolate classified as West African Clade was removed from the dataset as it was an outlier. To facilitate analysis, all variables were converted from string (character) to numerical (number) variables. The cut-off point of significance for all statistical testing was set at 5% ($p < 0.05$). In order to assess geographical distribution of the *Salmonella* Enteritidis isolate clades, the Outlier clade was set to the numerical value “0” for use as the reference. The Global clade was then set to the numerical value “1”. These observation values were labelled to aid in the interpretation of results. Difference in clade proportions between the four provinces was tested using the Pearson’s chi square. Similarly, reference groups for the isolate variables were set to the numerical value “0”. All numerical variables were labelled. To assess whether patient variables (age groups and gender) and isolate variables (invasive disease, using “non-invasive disease” as a reference, and drug resistance, using “susceptible” as a reference) were associated with a particular clade, Pearson’s chi square and the Fisher’s Exact tests (for variables with two groups) were performed. Simple bivariate logistic regression models (with clade and a significantly associated variable) were then used to obtain odds ratios (OR) and confidence intervals (CI) for each association.

3.2.10. Bivariate statistical analysis of HIV data subset

In order to determine whether clade associations with antimicrobial resistance and invasive disease were confounded by HIV status, statistical analysis of the association between HIV status and clade, antimicrobial resistance and invasive disease was performed. *Salmonella* Enteritidis isolates with corresponding HIV test result data were analysed separately from those with an “Unknown” status. This reduced dataset was imported into Stata from Excel (**Supplementary Figure 4**). All HIV data were converted from string (character) to numerical (number) variables. HIV negative statuses were set to the numerical value “0” for use as a reference. The positive statuses were set to the numerical value “1”. These numerical observations were then labelled. The association between HIV status and clade and antimicrobial resistance was tested using Pearson’s chi square and simple bivariate logistic regression modelling to obtain OR and CI. In order to assess the association between HIV status and invasive disease, the 0 - 4 year age group was removed since children under the age of 5 are highly susceptible to invasive disease for multiple factors other than HIV, and thus would skew the data. HIV and invasive disease association was then tested using logistic regression with age groups stratified, since both variables differed according to age group.

3.2.11. Ethics statement

Permission to access the GERMS-SA database was obtained from Dr Vanessa Quan, head of the GERMS-SA surveillance program, Division of Public Health Surveillance and Response, NICD (approved 6 August 2018) (**Supplementary Figure 6**). Ethical clearance for all laboratory-based surveillance and research (approved 12 November 2018) was obtained from the University of Witwaterstrand, Johannesburg, South Africa (Wits protocol no. M140159) by the Centre for Enteric Diseases, NICD (**Supplementary Figure 5**). Access to additional National Health Laboratory Service data from the CDW was granted by Dr Babatyi Malope-Kgokong, National Manager of Academic Affairs and Research, NHLS (approved 22 March 2019) (**Supplementary Figure 7**). The sensitive data used within this study (ie. HIV status) was anonymised through the removal of patient names.

3.3. Results

3.3.1. Description of the study cohort

A total of 618 confirmed *Salmonella* Enteritidis isolates were selected from the GERMS-SA 2012 and 2013 repository. From the accompanying patient data, we determined that the cohort consisted of similar proportions of male (329/618, 53.24%) and female (284/618, 45.95%) patients. The age range of these patients spanned 0-92 years (**Figure 3.3.1**), with the highest proportion of cases in children under the age of five years (27.99%, 173/618). HIV test result data was available for 228 (228/618, 36.89%) of the patients. A total of 165 (165/228, 72.37%) patients were HIV positive and 63 (63/228, 27.63%) were HIV negative. The outstanding HIV status data for the remaining 390 patients were regarded as unknown.

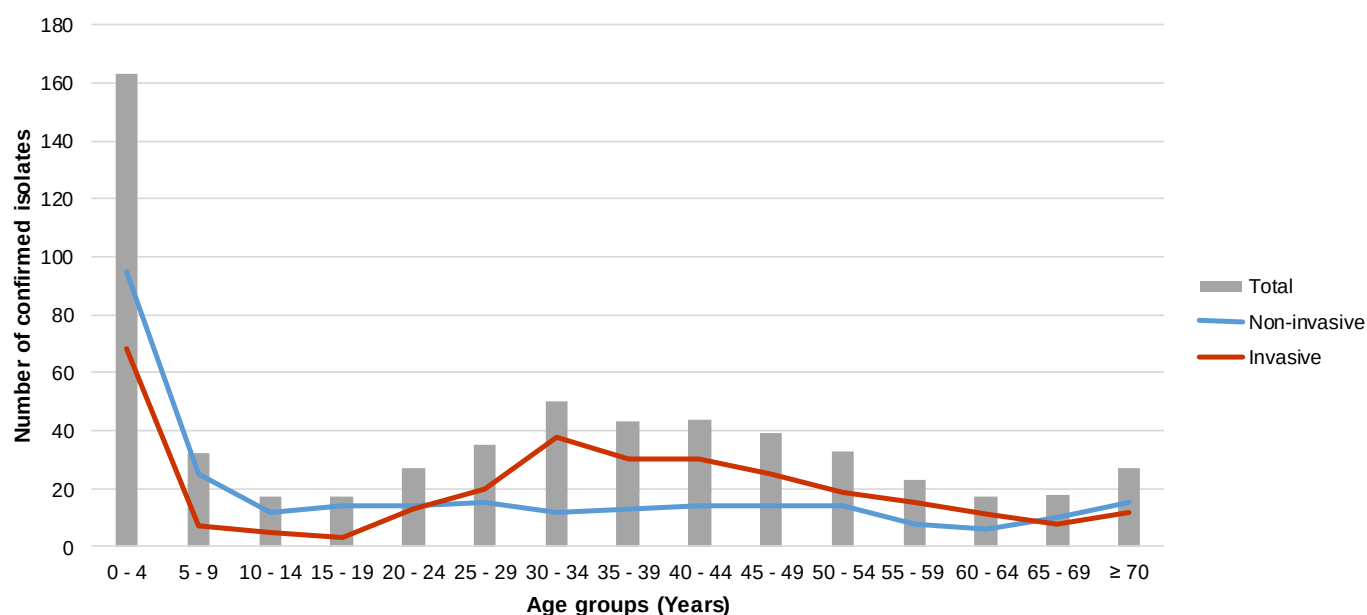


Figure 3.3.1 Frequency (grey bars) of the total *Salmonella* Enteritidis isolates, invasive isolates (red line) and non-invasive isolates (blue line) collected from patients in each age group.

3.3.2. Clade classification of *Salmonella* Enteritidis isolates across South Africa

All 618 *Salmonella* Enteritidis clinical isolates were successfully screened using the clade-categorising real-time PCR assay. The majority of the isolates were classified within the Outlier clade (377/618, 61.00 %). Fewer isolates were categorized in the Global clade (240/618, 38.83 %) and only one isolate was classified in the West African clade (1/618, 0.16 %). Seven isolates used in this study were previously grouped into clades by the hierBAPS (Bayesian Analysis of Population Structure) algorithm using whole-genome sequences. The results from the clade-classifying assay were concordant with the hierBAPS classification obtained previously ((218), Supplementary Table 1).

The proportions of each clade differed according to province (**Figure 3.3.2**). The Western Cape had the largest proportion of Global Clade isolates (81/168, 48.21%, OR: 1.49, CI: 1.02 – 2.17). Similar proportions of the Global Clade were observed in Gauteng (127/330, 38.48%, OR: 0.53, CI: 0.501 – 0.781) and KwaZulu-Natal (27/85, 31.36%, OR: 0.74, CI: 0.45 – 1.24). Mpumalanga had the smallest proportion of the Global Clade (5/35, 14.28%, OR: 0.28, CI: 0.10 – 0.73) and was the only province from which a Western African Clade isolate was identified.

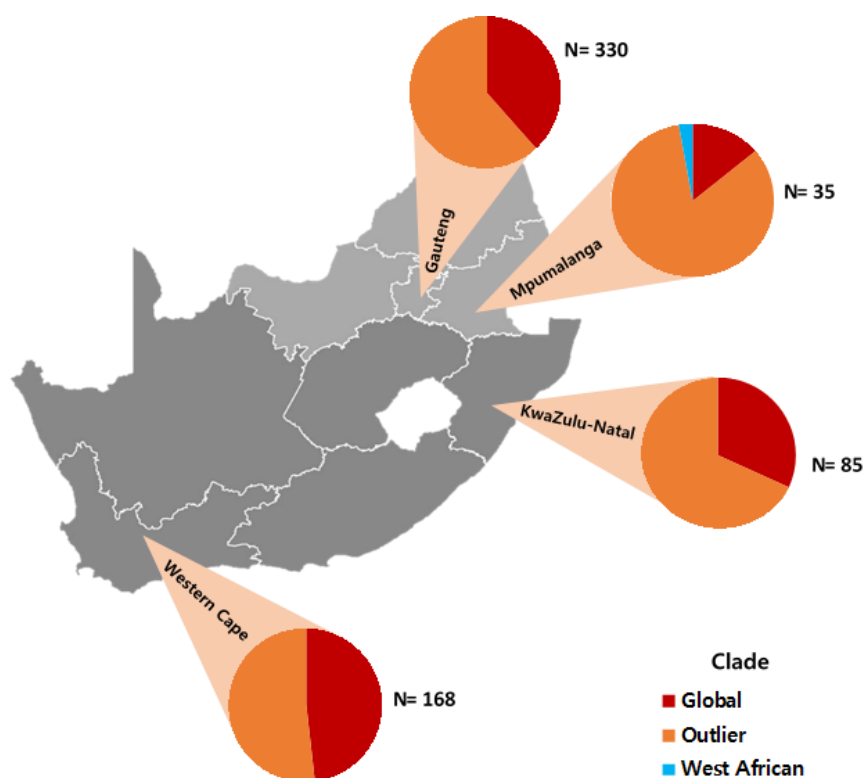


Figure 3.3.2 Graphical representation of the proportions of *Salmonella* Enteritidis isolates classified within the Outlier, Global and Western African Clades in four South African provinces (Gauteng, Mpumalanga, KwaZulu-Natal and Western Cape). The total number of isolates from each province (N) are included.

3.3.3. *Salmonella* Enteritidis isolates associated with invasive non-typhoidal *Salmonella* disease in patients 5 years and older

Of the 618 *Salmonella* Enteritidis isolates, 311/618 (50.32%) were associated with iNTS disease. A significant correlation was observed between invasive isolates and the age groups (**Figure 3.3.1**). Invasive salmonellosis was highest in patients aged 25 to 64 years (**Supplementary Table 6**). Similarly, when the smaller HIV data subset was analysed, a significant correlation was observed between HIV status and age group. The 25 to 34 years age group had the highest number of HIV positive patients (44/47, 93.6%, OR: 10.48, CI: 2.04 – 53.93) (**Supplementary Table 7**). This was followed by the 35 – 44 years age group (38/41, 92.7%, OR: 9.05, CI: 1.75 – 46.77). There was no evidence to support an association between invasive disease and HIV for all patients 5 years and above (126/168, 75.0%, OR: 2.4, CI: 0.84 – 6.88). However, when the data was stratified by age group, a significant association between HIV and invasive disease was observed in ages 25 – 34 years (OR: 8.67, CI: 1.54 – 48.70) and 35 – 44 years (OR: 7.78, CI: 1.38 – 43.85) (**Figure 3.3.3**).

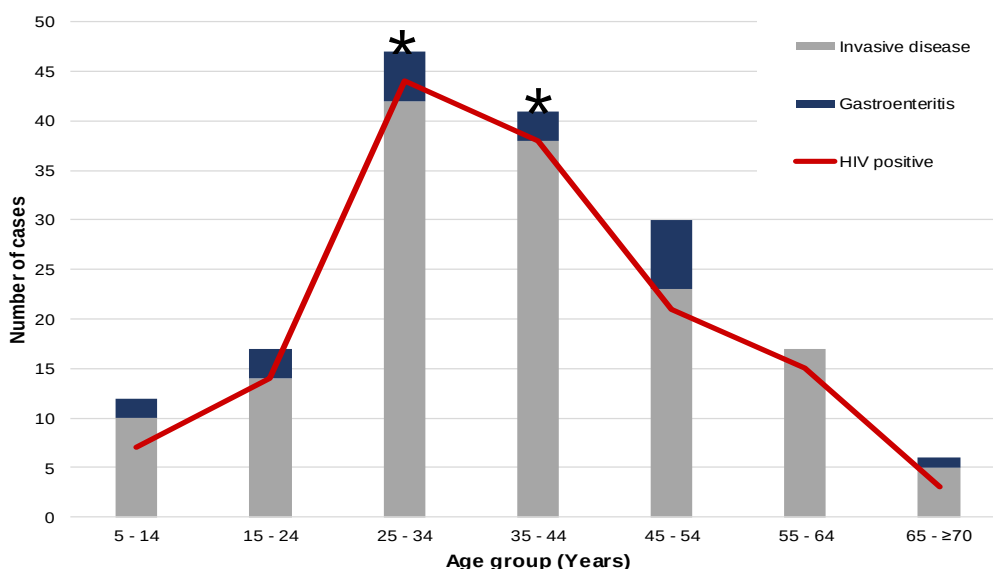


Figure 3.3.3 The number of invasive disease cases with the corresponding number of HIV positive cases across all age groups 5 years and above. The asterisks (*) indicate the age groups where there was evidence for an association between invasive disease and HIV status

The proportion of invasive isolates differed significantly between the four South African provinces (**Table 3.3.3**). The greatest proportion of invasive isolates were from KwaZulu-Natal (63/85, 74.12%, OR: 2.60, CI: 1.528 – 4.421), while, the Western Cape contributed the smallest proportion of invasive isolates (63/168, 37.50%, OR: 0.54, CI: 0.372 – 0.796). Isolates classified within the Global clade were significantly associated with invasive disease (140/240, 58.33%, OR: 1.57, CI: 1.13 – 2.17) (**Figure 3.3.1**).

Table 3.3.1 Association between invasive *Salmonella* Enteritidis isolates and South African provinces from which the *Salmonella* Enteritidis isolates were collected

Province	Proportion (%)	Odds Ratio (95% CI)
Gauteng	173/330 (52.42)	1.10 (0.888 - 1.368)
Kwazulu Natal	63/85 (74.12)	2.60 (1.528 - 4.421)
Western Cape	63/168 (37.50)	0.54 (0.372 - 0.796)
Mpumalanga	19/34 (55.88)	1.15 (0.565 - 2.400)

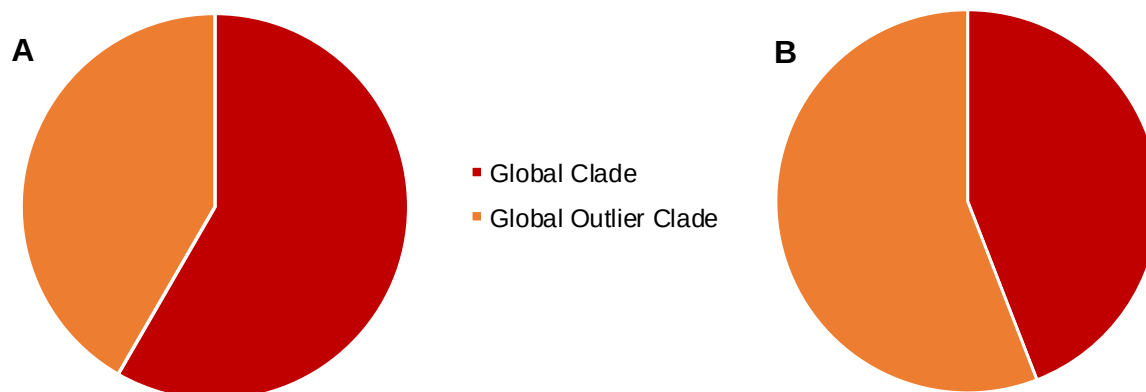


Figure 3.3.4 Proportion of invasive (**A**) and non-invasive (**B**) *Salmonella* Enteritidis isolates classified within each of the two clades (Outlier and Global Clade)

3.3.4. Antimicrobial resistance in *Salmonella* Enteritidis isolates

Forty percent (244/618) of the isolates were resistant to sulphamethoxazole (**Figure 3.3.5**). Similar proportions of sulphamethoxazole resistance were reported across the four South African provinces (**Table 3.3.2**). On analysis by E-test, 570/618 (92.23%) of the isolates were susceptible to ciprofloxacin (MIC ≤ 0.06 $\mu\text{g/mL}$) and 46/618 (7.44%) showed decreased ciprofloxacin susceptibility (DCS), with MIC ≥ 0.12 mg/L. None were classified as resistant isolates (MIC ≥ 4 $\mu\text{g/mL}$). The isolates with DCS (46/618, 7.44%) and nalidixic acid resistance (44/618, 7.12%), were concentrated in the Gauteng province (**Table 3.3.2**).

A small number of isolates were resistant to tetracycline (9/618, 1.46%), trimethoprim (9/618, 1.46%), ampicillin (7/618, 1.13%), chloramphenicol (6/618, 0.97%), cotrimoxazole (6/618, 0.97%) and amoxicillin-clavulanate (5/618, 0.81%). Few isolates were resistant to ceftazidime (3/618, 0.49%), ceftriaxone (3/618, 0.49%), streptomycin (3/618, 0.49%) and kanamycin (2/618, 0.32%). All isolates were fully susceptible to imipenem.

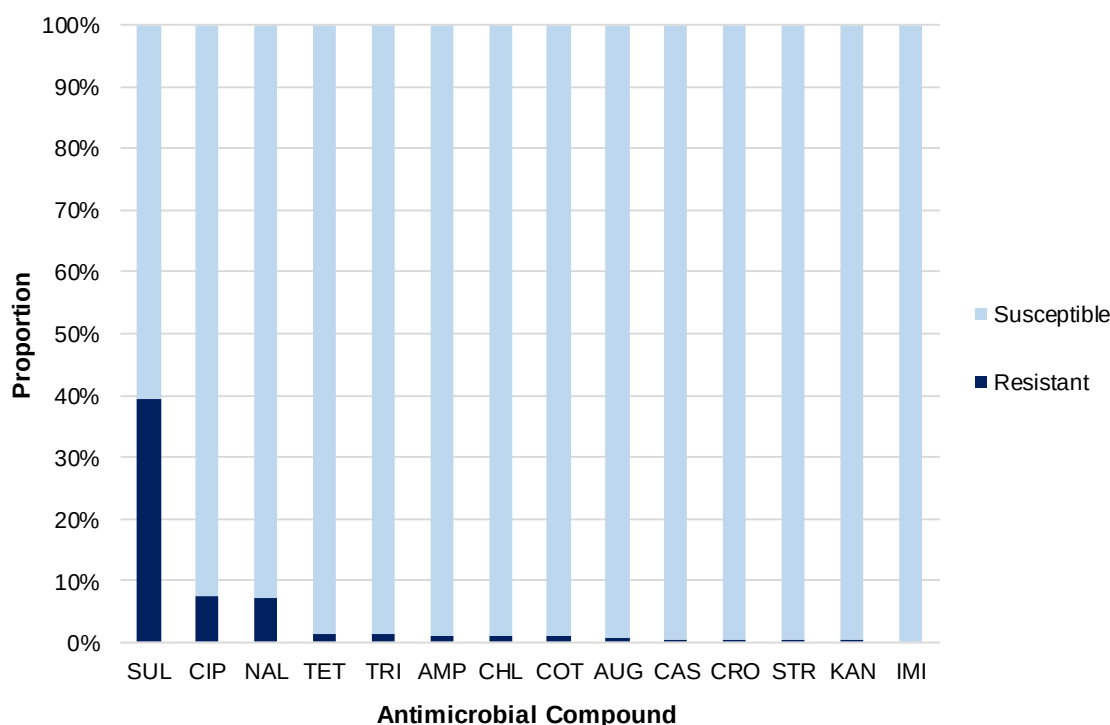


Figure 3.3.5 The proportion of *Salmonella* Enteritidis isolates that are resistant and susceptible to various antimicrobial compounds. (SUL= sulphamethoxazole; CIP= ciprofloxacin; NAL= nalidixic acid; TET= tetracycline; TRI= trimethoprim; AMP= ampicillin; CHL= chloramphenicol; COT= cotrimoxazole; AUG= amoxicillin-clavulanate; CAS= ceftazidime; CRO= ceftriaxone; STR= streptomycin; KAN= kanamycin; IMI= imipenem)

Table 3.3.2 Association between antimicrobial resistance of *Salmonella* Enteritidis isolates and the South African province from which the isolates were collected

					p-value	
	Gauteng (%)	Kwazulu Natal (%)	Mpumalanga (%)	Western Cape (%)	X ²	Fisher's Exact
Sulphamethoxazole						
Susceptible	190 (57.58)	56 (65.88)	22 (64.71)	105 (62.50)	0.433	0.446
Resistant	140 (42.42)	29 (34.12)	12 (35.29)	63 (37.50)		
Ciprofloxacin						
Susceptible	292 (88.48)	83 (97.65)	32 (94.12)	164 (97.62)	0.001	0.000
DCS	38 (11.52)	2 (2.35)	2 (5.88)	4 (2.38)		
Nalidixic Acid						
Susceptible	294 (89.09)	83 (97.65)	33 (97.06)	163 (97.62)	0.002	0.001
Resistant	36 (10.91)	2 (2.35)	1 (2.94)	5 (2.98)		

No evidence was found to support an association between sulphamethoxazole resistance and a particular clade. However, isolates classified within the Global Clade were strongly associated with resistance to nalidixic acid (28/240, 11.67%, OR: 11.63, CI: 4.671 – 28.968) and DCS (37/240, 15.42%, OR: 7.45, CI: 3.526 – 15.751) (**Table 3.3.3**). No evidence to support an association between a positive HIV status and nalidixic acid resistance (3/228, 1.32%, OR: 0.58, CI: 0.16 – 2.11) or DCS (4/228, 1.75%, OR: 0.73, CI: 0.23 – 2.30) was found.

Table 3.3.3 Association between antimicrobial resistance and Global Clade isolates

Antimicrobial Compound	Proportion Resistance (%)	Odds ratio (95% CI)
Ampicillin	3/240 (1.25)	1.18 (0.261 - 5.328)
Amoxicillin-clavulanate	3/240 (1.25)	2.37 (0.393 - 14.350)
Cotrimoxazole	3/240 (1.25)	1.58 (0.315 - 7.896)
Trimethoprim	3/240 (1.25)	0.79 (0.194 - 3.163)
Sulphamethoxazole	89/240 (37.08)	0.84 (0.605 - 1.178)
Chloramphenicol	4/240 (1.67)	3.18 (0.575 - 17.558)
Nalidixic Acid	28/240 (11.67)	11.63 (4.835 - 27.985)
Ciprofloxacin*	37/240 (15.42)	7.45 (3.526 - 15.751)
Tetracycline	1/240 (0.42)	0.19 (0.024 - 1.563)
Ceftriaxone	2/240 (0.83)	3.16 (0.284 - 35.173)
Ceftazidime	2/240 (0.83)	3.17 (0.285 - 35.322)

*Decreased susceptibility (MIC ≥ 0.12 mg/L)

3.4. Discussion

In this chapter, 618 *Salmonella* Enteritidis isolates were selected from the GERMS-SA frozen isolate repository and assigned clades using the clade-classifying real-time PCR assay. Corresponding epidemiological and clinical data were then statistically analysed to determine whether isolates within distinct clades were associated with particular characteristics (invasive disease and antimicrobial resistance). The results indicated that the proportion of Global and Outlier clade isolates differed significantly between provinces. Furthermore, there was strong evidence for the association of DCS and iNTS disease with Global clade isolates.

The frequency of invasive isolates collected from different age groups in this study (**Figure 3.3.2**) mirrored a high disease burden profile (**Figure 3.4.1**) that has previously been described in sub-Saharan Africa (sSA) (93,181). The majority of the invasive *Salmonella* Enteritidis isolates in the study cohort were collected from patients under the age of five years. This corresponds with the greatest proportion of globally reported iNTS cases occurring among children under the age of five (90,93,181). A second peak in iNTS disease incidence was observed in the 21 to 55 year age groups of the study cohort (**Figure 3.3.2**). The most likely explanation for this is the high HIV prevalence in this age range in sSA. A previous study observed the highest incidence of iNTS disease in HIV positive adults between the ages of 25 and 40 years in Africa (265). This corresponds well with the associations observed in this study between HIV positive status and iNTS disease in the 25 – 44 year age group. In high-income settings, such as Europe and the US, a second peak in iNTS disease incident is only observed in individuals over the age of 65 when the immune system naturally declines due to senescence (**Figure 3.4.1**) (181).

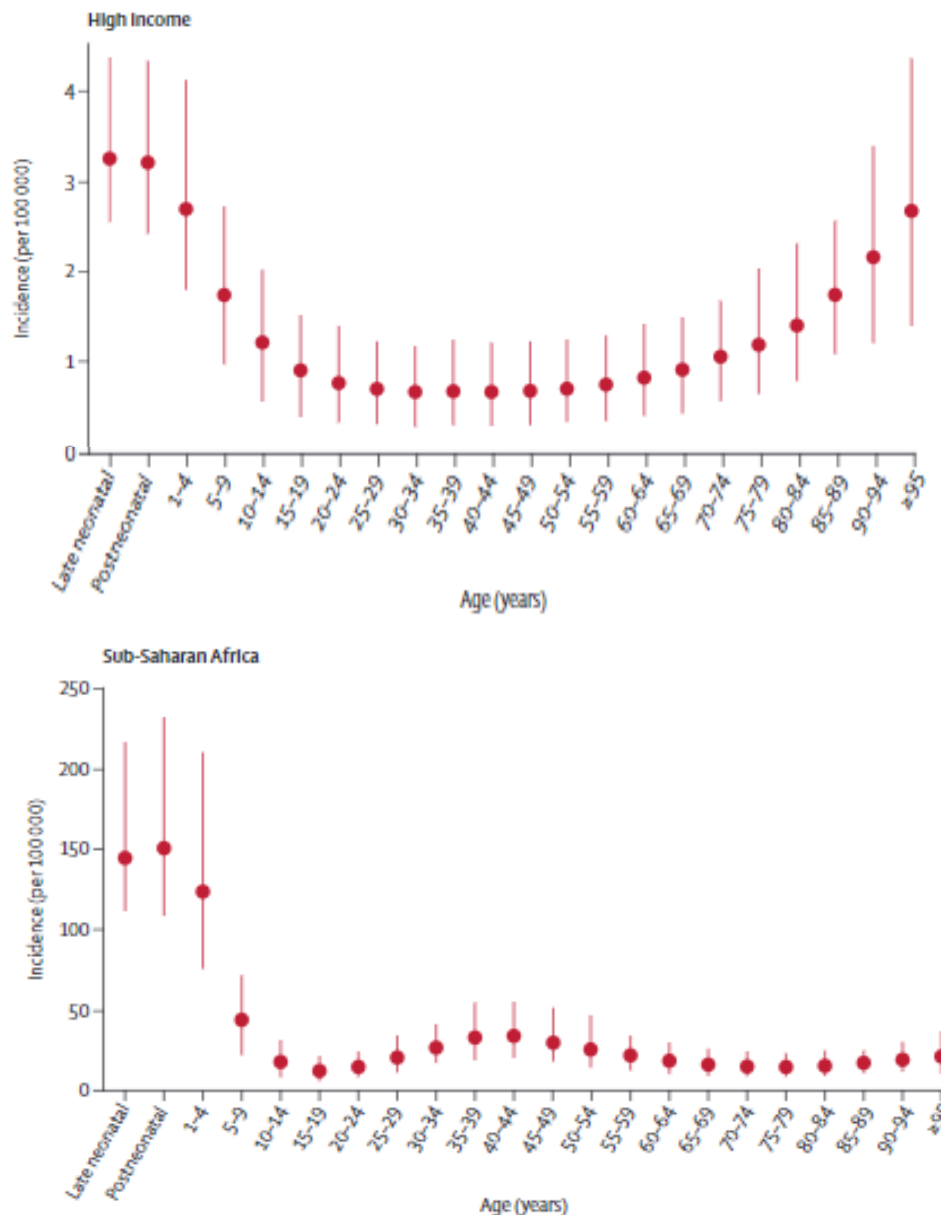


Figure 3.4.1 Proportion of invasive *Salmonella* isolates, by age group, from high-income and low- to middle-income settings in 2017. (181)

The clade-classifying assay successfully identified the clade of all 618 *Salmonella* Enteritidis isolates. Majority of the isolates in this study were from the Outlier clade. This particular clade is the oldest and largest paraphyletic cluster of *Salmonella* Enteritidis isolated worldwide (218,266) and therefore this was expected. Whilst few studies have associated the Outlier clade isolates with epidemics (218,261), there have been recent reports of focal outbreaks (261,267).

The results from this study have shown that the large proportions of Global clade isolates in three of South Africa's major economic provinces, Gauteng, KwaZulu-Natal and Western Cape, are of great concern. The Global clade has previously been connected with the epidemics of poultry-associated gastroenteritis across the US, Europe and Brazil in the 1970s and 1980s (235,239,261). The change of

predominant clade from Outlier to Global was suggested as the reason *Salmonella* Enteritidis become the most commonly isolated serovar in Brazil (261). It has been suggested that the SE20 phage region within the Global clade may have conferred an enhanced ability to colonise chickens, increasing the prevalence of *Salmonella* Enteritidis in the food chain and causing an increase in human infections (261). This could hypothetically also explain the rapid increase in the number of *Salmonella* Enteritidis cases observed in South Africa from 2006 to 2013 (**Figure 1.4.2.**). However, the change in proportions of the Global clade would require a longer surveillance period, than the two years covered in this study, to investigate this hypothesis. The clade-classifying assay, developed in this study, offers a low cost tool for this kind of surveillance.

The Global clade was thought to have originated in Europe and spread to the US and Brazil through the commercial exchange of poultry between these countries (261,266,268). The genetic homogeneity of the clade suggests rapid clonal expansion in each country (218,266). South Africa's increased poultry consumption between 2003 and 2013 (**Figure 3.4.2**) and subsequent increase in the import of poultry, after 2010, was the most likely source of Global clade introduction and spread within the country (262).

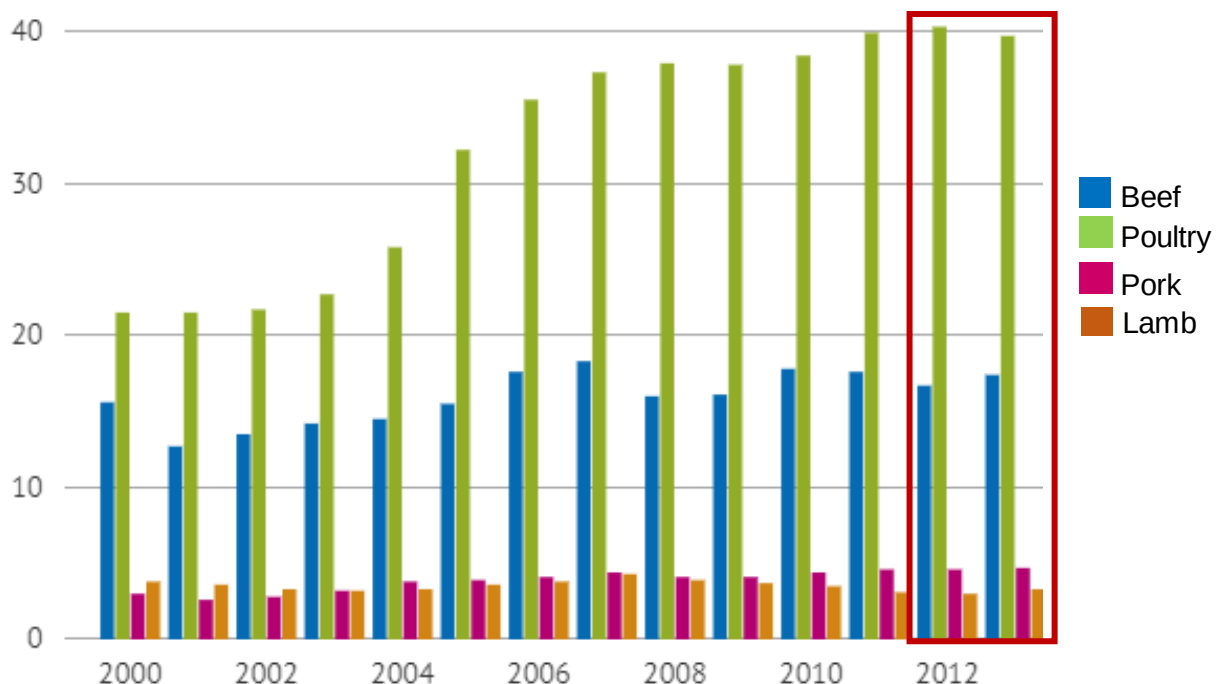


Figure 3.4.2 Consumption per capita of four different meat types (kg) in South Africa from 2000 to 2013. The two years (2012 and 2013) covered in this study are highlighted in the red box. (262)

The significant difference in clade proportions across the four South African provinces (Gauteng, KwaZulu-Natal, Western Cape and Mpumalanga) may be as a result of the number isolates selected from each province. Mpumalanga isolates were underrepresented since there was only one enhanced *Salmonella* surveillance site in the province. This may account for the significantly smaller proportion of Global clade within this province. However, the difference may also reflect the different opportunities for *Salmonella* Enteritidis transmission in various settings. The transmission dynamics of *Salmonella* Enteritidis within South Africa are not well described. The complexity of poultry distribution as well as differences in food safety practices between rural and urban settings presents a network of transmission opportunities (206). Therefore, the monitoring of clades present in the environment, livestock and health sector (a One Health integrated surveillance approach) will provide vital insight into the spread of *Salmonella* Enteritidis along the farm-to-fork continuum for more targeted prevention measures.

The absence of East and West African clade isolates, with the exception of the single West African isolate from Mpumalanga, may be due to the fundamental differences in the water, sanitation and hygiene infrastructure in South Africa compared with the rest of the continent. NTS can survive for long periods in water sources (65,269). Untreated sewage and environmental water have been shown to contain the same *Salmonella* Enteritidis strains as those implicated in iNTS disease (270). However, a recent study suggested that the genomic degradation within NTS lineages found in Africa and associated with invasive disease may have resulted in the loss of ability in these clades to form biofilm, the so-called RDAR phenotype (271). Subsequently, the iNTS strains are less likely to be found on food products and may have adopted alternate modes of transmission, such as human-to-human (271). Thus, greater access to quality sanitation and safe water provisions in South Africa may have partially protected the population from the spread of highly invasive, MDR Africa associated clades.

Of note, this study found evidence for the association of South African Global clade isolates with iNTS disease. This association was also described in a Ghanaian WGS study that clustered the majority of invasive *Salmonella* Enteritidis isolates within the Global clade (272). In contrast, Global clade isolates were not significantly associated with invasive disease in the European or the US epidemics (218); however, this is likely due to the much higher prevalence of immunosuppressive disease (importantly, HIV) in sSA, which predisposes individuals to iNTS disease (181,187). This was observed in KwaZulu-Natal, which has the highest HIV prevalence and subsequently had the highest number of invasive *Salmonella* Enteritidis isolates in this study (Table 3.3.1) (273). However, the weak association between invasive *Salmonella* Enteritidis infection and a positive HIV status observed in this and other studies may suggest that there are additional factors contributing to the risk of iNTS disease in South Africa (94).

The current recommended treatment for invasive *Salmonella* Enteritidis is a fluoroquinolone (such as ciprofloxacin) or a third-generation cephalosporin (such as ceftriaxone) (274). According to a multi-center surveillance study, DCS is uncommon in African iNTS isolates (275). However, since 2010 South Africa has reported an increase in DCS, with approximately 1% of all *Salmonella* Enteritidis surveillance isolates exhibiting DCS (276). This study also found a small proportion of isolates (7.44%) showed DCS for the years 2012 and 2013. Clinical evidence has shown that DCS may render the recommended treatment of iNTS ineffective (277). In cases of DCS, patients are generally treated with longer courses of alternative intravenous antimicrobials (274), which increases the cost of treatment, which is a major problem in developing countries where resources are limited.

The majority of South Africa's imported poultry is sourced from Brazil (262). Maximum residue limits for veterinary drugs have not been implemented in Brazil, therefore residual ciprofloxacin is still present in poultry intended for human consumption (278). In 2010, the highest ciprofloxacin resistance incidence were reported in turkey and chicken meat originating from Brazil (278,279). Thus, the increase in DCS in South Africa is possibly due to the increase in imported poultry from Brazil where DCS has previously been reported (280). Alternatively, the DCS may have developed independently due to extensive use in the poultry and healthcare industries within South Africa (274,281). To better understand the origins of DCS, routine monitoring of imported and local poultry is essential to implement targeted preventive measures.

The DCS observed in this study was strongly associated with Global clade isolates. The same association was also observed in the WGS study of Ghanaian *Salmonella* Enteritidis isolates exhibiting DCS, which clustered exclusively within the Global clade (272). This may suggest that the increase in DCS isolates was facilitated by the clonal expansion of the Global clade (282,283). This further substantiates the need to monitor clade distribution across South Africa to limit the spread of DCS *Salmonella* Enteritidis.

The identification of discrete *Salmonella* Enteritidis clades within sSA, associated with antimicrobial resistance and severe iNTS disease, highlights the necessity for surveillance and monitoring of changes in clades (218). Better understanding of clade trends and characteristics will allow for improved treatment and implementation of more targeted preventive measures. The clade-classifying assay provides a sensitive, affordable method for regular *Salmonella* Enteritidis clade surveillance in a public health setting.

CHAPTER FOUR

Conclusions

4.1. Summary of work

Four distinct *Salmonella* Enteritidis clades (Global, Outlier, West African and East African clades) were previously identified using whole genome sequences of isolates from across the world (218). The African clades differed genetically due to genomic degradation and the acquisition of unique prophage regions and multidrug-resistance (218,230,231). Using genes within the unique prophage regions as molecular markers, a real-time PCR, capable of distinguishing between the four clades, was developed (CHAPTER TWO). The first category-specific multiplex real-time PCR targeted the small toxic peptide gene and TIR domain gene, which distinguished between the African category (East and West African clades) and the Global/ Outlier category (Global and Outlier clades) respectively. The second clade-specific multiplex real-time PCR targeted the PemI gene and Chitinase Class I gene, which distinguished between the East African clade and Global clade respectively. The clade-classifying real-time PCR assay fulfilled all optimisation and validation requirements, and was determined to be a sensitive, specific and highly accurate gene-profiling assay for the discrimination of *Salmonella* Enteritidis clades.

In order to gain a retrospective understanding of the prevalence, geographical distribution and characteristics of clades in South Africa, the clade-classifying assay was used to classify 618 *Salmonella* Enteritidis isolates (CHAPTER THREE). These isolates were selected from the GERMS-SA frozen repository, which were collected from four South African provinces (Gauteng, KwaZulu-Natal, Mpumalanga and the Western Cape) in 2012 and 2013. The assay was successful in classifying all 618 isolates; the majority of the isolates were classified within the Outlier clade (61.00%), with fewer classified within the Global clade (38.83%) and only one isolate classified within the West African Clade (0.16%). The emergence of the Global clade in South Africa was of particular concern, since isolates within this clade were strongly associated with decreased ciprofloxacin susceptibility and invasive disease.

4.2. Limitations

The clade-classifying real-time PCR assay is limited to confirmed *Salmonella* Enteritidis isolates. The necessity for a serotyping test prior to the clade classifying of the isolate increases the cost per isolate as well as the processing time. However, this can be solved through the addition of a serovar-specific gene to the multiplex real-time PCR.

Due to surveillance bias the epidemiology investigation in this study could not infer risk factors or incident rates. Health-seeking behavior of people suffering from gastroenteritis and invasive disease differs greatly. Gastroenteritis, resulting from NTS infection, is self-limiting and usually does not require antibiotic treatment. Therefore, people generally do not visit hospitals when suffering from gastroenteritis, which limits the number of stool samples collected from patients and resulted in the overall underrepresentation of NTS. iNTS disease, although more severe than gastroenteritis and more likely to cause hospital-seeking behavior in diseased people, may be difficult to distinguish from other non-specific febrile illness. iNTS is more likely to be recognised in tertiary care settings where blood cultures are more likely to be collected from patients with non-specific febrile syndromes. Therefore, surveillance systems are reliant on the better funded academic hospitals to send isolates associated with iNTS disease. The distribution of academic hospitals in the country differs greatly between provinces, resulting in some provinces being underrepresented in surveillance programs.

The sample size of *Salmonella* Enteritidis isolates used in Chapter Three of this study was too small to use for significant healthcare measures. Approximately a third (618/1 973, 31.32 %) of all *Salmonella* Enteritidis isolates submitted to the NICD in 2012 and 2013 were classified in this study. This number of isolates was sufficient to demonstrate the trial and application of the clade-classifying assay on South African surveillance isolates. However, in order to gain a better, more complete understanding of the *Salmonella* Enteritidis clades in South Africa, all isolates submitted to the NICD would need to be classified according to clade.

Furthermore, the presence of a West African clade isolate in Mpumalanga is of particular concern. To gain a more accurate understanding of the prevalence of the African clades in South Africa, provinces underrepresented in this study, such as Mpumalanga, Eastern Cape, Northern Cape and the Free State, require more intensive surveillance. This may be done by establishing sentinel surveillance sites or incentivising public and private healthcare facilities to submit *Salmonella* isolates to the NICD for further classification.

Another constraint of this study was the limited HIV data available for the selected isolates. There was no data available for whether the patient was receiving HIV treatment at the time of infection with *Salmonella* Enteritidis. This restricted the interpretation of the association between iNTS disease and HIV since treatment may have reduced the patient's susceptibility to iNTS, thus concealing the true association. In future studies, a more complete HIV dataset should be established that includes test results, status and treatment regimes.

4.3. Implication of this work

DNA fingerprinting methodologies, such as MLVA and PFGE, are insufficient for distinguishing between *Salmonella* Enteritidis groups due to the high degree of clonality (31). Therefore, these tests were not suitable for use in epidemiological investigations to determine the relatedness of isolates. The clade-classifying real-time PCR developed in this study offers a high quality, cost-effective method of distinguishing between the four phylogenetic groups within the *Salmonella* Enteritidis serovar in a setting where WGS may not be available. The assay can be used to cluster isolates from an outbreak investigation by clade, which could be especially helpful in tracing transmission of *Salmonella* Enteritidis within South Africa.

The association of the Global clade with invasive disease and DCS in South Africa is important to note. The Global clade has an enhanced ability to infect poultry, which has placed the country at an ever increasing risk of *Salmonella* Enteritidis infections as poultry consumption continues to increase. This necessitates the surveillance of clades in South Africa in order to monitor changes in clade proportions. Better understanding of the epidemiology and transmission of particular clades will help with the implementation of more targeted prevention measures.

CHAPTER FIVE

References

1. Thucydides (Translated by Smith W.) History of the Peloponnesian War. 1st Edit. London: Jones & Co.; 1831.
2. Papagrigorakis MJ, Yapijakis C, Synodinos PN, Baziotopoulou-Valavani E. DNA examination of ancient dental pulp incriminates typhoid fever as a probable cause of the Plague of Athens. *Int J Infect Dis*. 2006;10(3):206–14.
3. Percy G. Observations gathered out of a discourse of the plantation of the southern colonies in Virginia by the English, 1606. In: Purchas his Pilgrimes. 1625. p. 1685–90.
4. Smith J. The description of Virginia by captaine John Smith. In: Purchas his Pilgrimes. 1625. p. 1691–704.
5. Smart C. The medical and surgical history of the War of the Rebellion. Washington D.C.: Government Printing Office; 1879.
6. Sedgwick W, Hazen A. Typhoid fever in Chicago. *Eng News Am Railw J*. 1892;
7. Wright A. A short treatise on anti-typhoid inoculation; containing an exposition of the principles of the method and a summary of the results achieved by its application. 1st ed. London: His Majesty's Printers; 1904.
8. Center for Disease Control and Prevention (U.S.). Salmonella Typhimurium outbreak traced to a commercial apple cider-New Jersey. Vol. 24, Morbidity and Mortality Weekly Report. [U.S. Dept. of Health, Education, and Welfare, Public Health Service, Center for Disease Control]; 1975.
9. Ryan CA, Nickels MK, Hargrett-Bean NT, Potter ME, Endo T, Mayer L, et al. Massive outbreak of antimicrobial-resistant Salmonellosis traced to pasteurized milk. *J Am Med Assoc*. 1987;258(22):3269.
10. Bean N, Griffin P, Goulding J, Ivey C. Foodborne disease outbreaks, 5-year summary, 1983-1987. Atlanta; 1990.
11. Altekruse SF, Cohen ML, Swerdlow DL. Emerging foodborne diseases. *Emerg Infect Dis*. 1997;3(3):285–93.
12. Ricke S, Gast R. Overview of Salmonellosis and food-borne Salmonella: Historical and current perspectives. In: Producing safe eggs : Microbial ecology of Salmonella. Academic Press; 2016. p. 116.
13. Eberth C. Die Organismen in den Organen bei Typhus abdominalis. *Arch für Pathol Anat und Physiol und für Klin Med*. 1880;81(1):58–74.
14. Coats J. Eberth's Typhoid Bacillus. *Br Med J*. 1882;1(1108):421.
15. Salmon DE, Smith T. The bacterium of swine-plague. *Mon Microsc J*. 1886;7:204–5.
16. Salmonella Subcommittee of the Nomenclature Committee of the International Society for Microbiology TSS of the NC of the IS for. The Genus Salmonella Lignières, 1900. *J Hyg (Lond)*. 1934;34(3):333–50.
17. United States Bureau of Animal Industry. Hog Cholera: Its history, nature and treatment, as determined by the United States Bureau of Animal Industry. Hog Cholera. Washington: Government Printing Office; 1889.
18. Schultz M. Theobald Smith. *Emerg Infect Dis*. 2008;14(12):1940–2.
19. Kauffmann W. Salmonella subcommittee. *J Hyg (Lond)*. 1934;34:333–50.
20. Murray PR, Baron EJ, American Society for Microbiology. Manual of clinical microbiology [Internet]. 8th ed. Washington D.C.: ASM Press; 2003 [cited 2019 Feb 11]. Available from: <https://www.worldcat.org/title/manual-of-clinical-microbiology/oclc/300278480>

21. Agbaje, M, Begum, R, Oyekunle, M, Ojo, O, Adenubi O. Evolution of *Salmonella* nomenclature: a critical note. *Folia Microbiol (Praha)*. 2011;56(6):497–503.
22. Kauffmann F. *The Bacteriology of Enterobacteriaceae*. Copenhagen: Munksgaard; 1966.
23. Tortora GJ, Funke BR, Case CL. *Microbiology : an introduction*. 12th Edit. Harlow, United Kingdom: Pearson Education Limited; 2016.
24. Dwarkin, M, Alkow S. Proteobacteria; Gamma subclass. In: *The prokaryotes*. New York: Springer; 2006. p. 811-835.
25. Berkowitz F, Jerris R. *Practical medical microbiology for clinicians*. 1st Edit. Hoboken, New Jersey: Wiley-Blackwell; 2016.
26. Bern M, Goldberg D. Automatic selection of representative proteins for bacterial phylogeny. *BMC Evol Biol*. 2005;5(1):34.
27. Le Minor L, Popoff MY. Designation of *Salmonella enterica* sp. nov., nom. rev., as the type and only species of the genus *Salmonella*. *Int J Syst Bacteriol*. 1987;37(4):465–8.
28. Reeves MW, Evins GM, Heiba AA, Plikaytis BD, Farmer JJ. Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. *J Clin Microbiol*. 1989;27(2):313–20.
29. WHO Collaborating Centre for Reference and Research on *Salmonella*. *Antigenic Formulae of the Salmonella serovars*. 9th Edit. Grimont P, Weill F, editors. Geneva, Switzerland; 2007.
30. Tauxe, R, Pavia A. *Salmonellosis: nontyphoidal*. In: *Bacterial infections of humans: epidemiology and control*. 3rd Edit. New York: Plenum Medical Book Company; 1998. p. 613–30.
31. Arya G, Holtslander R, Robertson J, Yoshida C, Harris J, Parmley J, et al. Epidemiology, Pathogenesis, Genosertotyping, Antimicrobial Resistance, and Prevention and Control of Non-Typhoidal *Salmonella* Serovars. *Curr Clin Microbiol Reports* [Internet]. 2017;4(1):43–53. Available from: <http://link.springer.com/10.1007/s40588-017-0057-7>
32. Threlfall J, Ward L, Old D. Changing the nomenclature of *Salmonella*. *Commun Dis public Heal*. 1999;2(3):156–7.
33. Lynch MF, Tauxe R V. *Salmonellosis: Nontyphoidal*. In: *Bacterial Infections of Humans*. 4th Edit. Boston, MA: Springer US; 2009. p. 677–98.
34. Uzzau S, Brown DJ, Wallis T, Rubino S, Leori G, Bernard S, et al. Host adapted serotypes of *Salmonella enterica*. *Epidemiol Infect*. 2000;125(2):229–55.
35. Gray JT, Fedorka-Cray PJ, Stabel TJ, Kramer TT. Natural transmission of *Salmonella choleraesuis* in swine. *Appl Environ Microbiol*. 1996;62(1):141–6.
36. Kingsley RA, Bäumlér AJ. Host adaptation and the emergence of infectious disease: the *Salmonella* paradigm. *Mol Microbiol*. 2000;36(5):1006–14.
37. Schikora A, Carreri A, Charpentier E, Hirt H. The Dark Side of the Salad: *Salmonella Typhimurium* overcomes the innate immune response of *Arabidopsis thaliana* and shows an endopathogenic lifestyle. *Ojcius DM*, editor. *PLoS One*. 2008;3(5):e2279.
38. Giannella RA. *Salmonella*. Medical Microbiology. University of Texas Medical Branch at Galveston; 1996.
39. Gram H. Über die isolierte Färbung der Schizomyceten in Schnitt- und Trockenpräparaten. *Fortschr Med*. 1884;2:185–9.

40. Versalovic J, American Society for Microbiology. Manual of clinical microbiology. ASM Press; 2011.
41. Allen-Vercoe E, Woodward MJ. The Role of Flagella, But not Fimbriae, in the adherence of *Salmonella enterica* serotype Enteritidis to Chick Gut Explant. *J Med Microbiol.* 1999;48(8):771–80.
42. Parker CT, Guard-Petter J. Contribution of flagella and invasion proteins to pathogenesis of *Salmonella enterica* serovar enteritidis in chicks. *FEMS Microbiol Lett.* 2001;204(2):287–91.
43. Josenhans C, Suerbaum S. The role of motility as a virulence factor in bacteria. *Int J Med Microbiol.* 2002;291(8):605–14.
44. Shah, D, Xiaohui, Z, Addwebi, T, Davis, M, Orfe, L, Call, D, Guard, J, Besser T. Cell invasion of poultry-associated *Salmonella enterica* serovar Enteritidis isolates is associated with pathogenicity, motility and proteins secreted by the type III secretion system. *Microbiology.* 2011;157:1428–45.
45. Winter SE, Thiennimitr P, Winter MG, Butler BP, Huseby DL, Crawford RW, et al. Gut inflammation provides a respiratory electron acceptor for *Salmonella*. *Nature.* 2010;467(7314):426–9.
46. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, et al. *Salmonella enterica* serovar Typhimurium exploits inflammation to compete with the intestinal microbiota. *PLoS Biol.* 2007;5(10):2177–89.
47. Farr SB, Kogoma T. Oxidative Stress Responses in *Escherichia coli* and *Salmonella Typhimurium*. Vol. 55, *Microbiological Reviews.* 1991.
48. Hassett DJ, Cohen MS. Bacterial adaptation to oxidative stress: implications for pathogenesis and interaction with phagocytic cells. *FASEB J.* 1989;3(14):2574–82.
49. Gutnick D, Calvo JM, Klopotoski T, Ames BN. Compounds which serve as the sole source of carbon or nitrogen for *Salmonella Typhimurium* LT-2. *J Bacteriol.* 1969;100(1):215–9.
50. Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA. Brock biology of microorganisms. 14th ed. Thakur S, editor. Pearson Education Limited; 2015. 1030 p.
51. Barrett EL, Riggs DL. Evidence of a second nitrate reductase activity that is distinct from the respiratory enzyme in *Salmonella Typhimurium*. *J Bacteriol.* 1982;150(2):563–71.
52. Finlay BB, Falkow S. *Salmonella* interactions with polarized human intestinal Caco-2 epithelial cells. *J Infect Dis.* 1990;162(5):1096–106.
53. Lhocine N, Arena ET, Bomme P, Ubelmann F, Prévost M-C, Robine S, et al. Apical invasion of intestinal epithelial cells by *Salmonella Typhimurium* requires villin to remodel the brush border actin cytoskeleton. *Cell Host Microbe.* 2015;17(2):164–77.
54. Takeuchi A. Electron microscope studies of experimental *Salmonella* infection: Penetration into the intestinal epithelium by *Salmonella Typhimurium*. *Am J Pathol.* 1967;50(1):109–36.
55. Bäumler AJ, Tsois RM, Heffron F. Contribution of fimbrial operons to attachment to and invasion of epithelial cell lines by *Salmonella Typhimurium*. *Infect Immun.* 1996;64(5):1862–5.
56. Jepson MA, Collares-Buzato CB, Clark MA, Hirst BH, Simmons NL. Rapid disruption of epithelial barrier function by *Salmonella typhimurium* is associated with structural modification of intercellular junctions. *Infect Immun.* 1995;63(1):356–9.
57. Williams, Jonathan, Duckworth, Carrie, Burkitt, Michael, Pritchard D. Epithelial Cell Shedding and Barrier Function: A Matter of Life and Death at the Small Intestinal Villus Tip. *Vetrinary Pathol.* 2014;52(3).

58. Vazquez-Torres A, Jones-Carson J, Bäumlér AJ, Falkow S, Valdivia R, Brown W, et al. Extraintestinal dissemination of *Salmonella* by CD18-expressing phagocytes. *Nature*. 1999;401(6755):804–8.
59. Jones BD, Ghori N, Falkow S. *Salmonella* Typhimurium initiates murine infection by penetrating and destroying the specialized epithelial M cells of the Peyer's patches. *J Exp Med*. 1994;180(1):15–23.
60. Leung KY, Finlay BB. Intracellular replication is essential for the virulence of *Salmonella* Typhimurium. Vol. 88, *Proc. Natl. Acad. Sci. USA*. 1991.
61. Haraga A, Ohlson MB, Miller SI. *Salmonellae* interplay with host cells. *Nat Rev Microbiol*. 2008;6(1):53–66.
62. Giannella RA, Formal SB, Dammin GJ, Collins H. Pathogenesis of Salmonellosis: Studies of fluid secretion, mucosal invasion, and morphological reaction in the rabbit ileum. *J Clin Invest*. 1973;52(2):441–53.
63. CDC. *Salmonella* serotype Enteritidis infections among workers producing poultry vaccine, Maine, November–December 2006. 2007.
64. World Health Organisation (WHO). *Foodborne disease outbreaks: Guidelines for investigation and control*. Geneva, Switzerland; 2008.
65. World Health Organisation (WHO). *Salmonella (non-typhoidal)* [Internet]. 2018 [cited 2019 Oct 23]. Available from: [https://www.who.int/news-room/fact-sheets/detail/salmonella-\(non-typhoidal\)](https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal))
66. Brooks JT, Matyas BT, Fontana J, DeGroot MA, Beuchat LR, Hoekstra M, et al. An outbreak of *Salmonella* serotype Typhimurium infections with an unusually long incubation period. *Foodborne Pathog Dis*. 2012;9(3):245–8.
67. Majowicz SE, Musto J, Scallan E, Angulo FJ, Kirk M, O'Brien SJ, et al. The Global Burden of Nontyphoidal *Salmonella* Gastroenteritis. *Clin Infect Dis*. 2010;50(6):882–9.
68. Parry CM, Thomas S, Aspinall EJ, Cooke RP, Rogerson SJ, Harries AD, et al. A retrospective study of secondary bacteraemia in hospitalised adults with community acquired non-typhoidal *Salmonella* gastroenteritis. *BMC Infect Dis*. 2013;13(1):107.
69. Herrera-León S, Alonso PL, García V, Mandomando I, Ruiz J, Herrera-León S, et al. *Salmonella enterica* serovars Typhimurium and Enteritidis causing mixed infections in febrile children in Mozambique. *Infect Drug Resist*. 2018;11:195–204.
70. Gal-Mor O, Boyle EC, Grassl GA. Same species, different diseases: how and why typhoidal and non-typhoidal *Salmonella enterica* serovars differ. *Front Microbiol*. 2014;5:391.
71. Crump JA, Sjölund-Karlsson M, Gordon MA, Parry CM. Epidemiology, Clinical Presentation, Laboratory Diagnosis, Antimicrobial Resistance, and Antimicrobial Management of Invasive *Salmonella* Infections. *Clin Microbiol Rev*. 2015;28(4):901–37.
72. Vugia DJ, Samuel M, Farley MM, Marcus R, Shiferaw B, Shallow S, et al. Invasive *Salmonella* infections in the United States, FoodNet, 1996–1999: Incidence, Serotype Distribution, and Outcome. *Clin Infect Dis*. 2004;38(s3):S149–56.
73. Muthumbi E, Morpeth SC, Ooko M, Mwanzi A, Mwarumba S, Mturi N, et al. Invasive Salmonellosis in Kilifi, Kenya. *Clin Infect Dis*. 2015;61(Suppl 4):S290–301.
74. Tapia MD, Tennant SM, Bornstein K, Onwuchekwa U, Tamboura B, Maiga A, et al. Invasive nontyphoidal *Salmonella* infections among children in Mali, 2002–2014: Microbiological and Epidemiologic Features Guide Vaccine Development. *Clin Infect Dis*. 2015;61 Suppl 4(Suppl 4):S332–8.

75. Feasey NA, Everett D, Faragher EB, Roca-Feltre A, Kang'ombe A, Denis B, et al. Modelling the contributions of malaria, HIV, malnutrition and rainfall to the decline in paediatric invasive non-typhoidal *Salmonella* Disease in Malawi. Baker S, editor. *PLoS Negl Trop Dis*. 2015;9(7):e0003979.
76. Braden CR. *Salmonella enterica* serotype Enteritidis and eggs: A national epidemic in the United States. *Clin Infect Dis*. 2006;43(4):512–7.
77. St Louis ME, Morse DL, Potter ME, DeMelfi TM, Guzewich JJ, Tauxe R V, et al. The emergence of grade A eggs as a major source of *Salmonella* Enteritidis infections. New implications for the control of salmonellosis. *JAMA*. 1988;259(14):2103–7.
78. Ricke SC, Gast RK. *Producing safe eggs : microbial ecology of Salmonella*. 1st Edit. London, United Kingdom: Elsevier- Academic Press; 2017. 436 p.
79. Patrick ME, Adcock PM, Gomez TM, Altekruze SF, Holland BH, Tauxe R V., et al. *Salmonella* Enteritidis Infections, United States, 1985-1999. *Emerg Infect Dis*. 2004;10(1):1–7.
80. Wright AP, Richardson L, Mahon BE, Rothenberg R, Cole DJ. The rise and decline in *Salmonella enterica* serovar Enteritidis outbreaks attributed to egg-containing foods in the United States, 1973-2009. *Epidemiol Infect*. 2016;144(4):810–9.
81. Tier J. *Global Meat Production and Consumption Continue to Rise*. Worldwatch Institute (Press Center). 2011;
82. Food and Agricultural Organization of the United Nations. Animal Production and Health Division: Meat and Meat Products [Internet]. 2019. Available from: <http://www.fao.org/ag/aginfo/themes/en/meat/home.html>
83. Ward LR, Threlfall J, Smith HR, O'Brien SJ. *Salmonella* Enteritidis epidemic. *Science* (80-). 2000;287(5459):1753–6.
84. Dekker D, Krumkamp R, Eibach D, Sarpong N, Boahen KG, Frimpong M, et al. Characterization of *Salmonella enterica* from invasive bloodstream infections and water sources in rural Ghana. 2018;1–5.
85. Tadesse G, Gebremedhin EZ. Prevalence of *Salmonella* in raw animal products in Ethiopia: a meta-analysis. *BMC Res Notes*. 2015;8(1):163.
86. Sibhat B, Molla Zewde B, Zerihun A, Muckle A, Cole L, Boerlin P, et al. *Salmonella* serovars and antimicrobial resistance profiles in beef cattle, slaughterhouse personnel and slaughterhouse environment in Ethiopia. *Zoonoses Public Health*. 2011;58(2):102–9.
87. Kikui GM, Ombui JN, Mitema ES. Serotypes and antimicrobial resistance profiles of *Salmonella* isolates from pigs at slaughter in Kenya. *J Infect Dev Ctries*. 2010;4(4):243–8.
88. Kariuki S, Revathi G, Kariuki N, Kiiru J, Mwituria J, Muyodi J, et al. Invasive multidrug-resistant non-typhoidal *Salmonella* infections in Africa: zoonotic or anthroponotic transmission? *J Med Microbiol*. 2006;55(5):585–91.
89. Dekker D, Krumkamp R, Eibach D, Sarpong N, Boahen KG, Frimpong M, et al. Characterization of *Salmonella enterica* from invasive bloodstream infections and water sources in rural Ghana. *BMC Infect Dis*. 2018;18(1):1–5.
90. Boore AL, Hoekstra RM, Iwamoto M, Fields PI, Bishop RD, Swerdlow DL. *Salmonella enterica* infections in the United States and assessment of coefficients of variation: A Novel approach to identify epidemiologic characteristics of individual serotypes, 1996-2011. Chaturvedi V, editor. *PLoS One*. 2015;10(12):e0145416.
91. Barton Behravesh C, Jones TF, Vugia DJ, Long C, Marcus R, Smith K, et al. Deaths Associated With Bacterial Pathogens Transmitted Commonly Through Food: Foodborne Diseases Active Surveillance Network (FoodNet), 1996–2005. *J Infect Dis*. 2011;204(2):263–7.

92. Leekitcharoenphon P, Nielsen EM, Kaas RS, Lund O, Aarestrup FM. Evaluation of Whole Genome Sequencing for Outbreak Detection of *Salmonella enterica*. Chabalgoyt JA, editor. PLoS One. 2014;9(2):e87991.
93. Ao TT, Feasey NA, Gordon MA, Keddy KH, Angulo FJ, Crump JA, et al. Global Burden of invasive nontyphoidal *Salmonella* disease, 2010. Emerg Infect Dis. 2015;21(6):941–9.
94. Keddy KH, Takuva S, Musekiwa A, Puren AJ, Sooka A, Karstaedt A, et al. An association between decreasing incidence of invasive non-typhoidal salmonellosis and increased use of antiretroviral therapy, Gauteng Province, South Africa, 2003–2013. Vermund SH, editor. PLoS One. 2017;12(3):2003–13.
95. Mandomando I, Bassat Q, Sigaúque B, Massora S, Quintó L, Ácacio S, et al. Invasive *Salmonella* infections among children from rural Mozambique, 2001–2014. Clin Infect Dis. 2015;61(suppl 4):S339–45.
96. European Food Safety Authority (EFSA) EC for DP and C (ECDC). The European Union summary report on trends and sources of zoonoses agents and foodborne outbreaks in 2011. EFSA J. 2013;11:3129.
97. Public Health Laboratory Network. Burden and causes of foodborne disease in Australia: Annual report of the OzFoodNet network. Australian Government Department of Health and Ageing; 2005.
98. Gordon MA, Graham SM, Walsh AL, Wilson L, Phiri A, Molyneux E, et al. Epidemics of invasive *Salmonella enterica* serovar Enteritidis and *S. enterica* serovar Typhimurium infection associated with multidrug resistance among adults and children in Malawi. Clin Infect Dis. 2008;46(7):963–9.
99. Herrera-León S, Ramiro R, Arroyo M, Díez R, Usera MA, Echeita MA. Blind comparison of traditional serotyping with three multiplex PCRs for the identification of *Salmonella* serotypes. Res Microbiol. 2007;158(2):122–7.
100. Foley SL, Zhao S, Walker RD. Comparison of molecular typing methods for the differentiation of *Salmonella* foodborne pathogens. Foodborne Pathog Dis. 2007;4(3):253–76.
101. Wattiau P, Boland C, Bertrand S. Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: gold standards and alternatives. Appl Environ Microbiol. 2011;77(22):7877–85.
102. Ng PC, Kirkness EF. Whole Genome Sequencing. In Humana Press, Totowa, NJ; 2010. p. 215–26.
103. Ibrahim GM, Morin PM. *Salmonella* serotyping using whole genome sequencing. Front Microbiol. 2018;9:2750.
104. Statens Serum Institut. E. Coli Antisera. 2013.
105. McQuiston JR, Fields PI, Tauxe R V., Logsdon JM. Do *Salmonella* carry spare tyres? Trends Microbiol. 2008;16(4):142–8.
106. Koehn A. Technical modification of the swarming-plate method of Sven Gard in *Salmonella* diagnosis. Zentralblatt für Bakteriologie Parasitenkunde, Infekt und Hyg. 1970;215(4):449–55.
107. Schrader KN, Fernandez-Castro A, Cheung WKW, Crandall CM, Abbott SL. Evaluation of commercial antisera for *Salmonella* serotyping. J Clin Microbiol. 2008;46(2):685–8.
108. Anderson ES, Williams RE O. Bacteriophage typing of enteric pathogens and Staphylococci and its use in epidemiology. Vol. 9, J. clin. Path. 1956.
109. Hickman-Brenner FW, Stubbs AD, Farmer III JJ. Phage Typing of *Salmonella* Enteritidis in the United States. Vol. 99, Journal of Clinical Microbiology. 1991.
110. Baggesen DL, Sørensen G, Nielsen EM, Wegener HC. Phage typing of *Salmonella* Typhimurium - is it still a useful tool for surveillance and outbreak investigation? Euro Surveill. 2010;15(4):19471.

111. Rankin, S., Platt DJ. Phage conversion in *Salmonella enterica* serotype Enteritidis: implications for epidemiology. *Epidemiol Infect.* 1995;114:227–36.
112. Tankouo-Sandjong, B., Kinde, H., Wallace I. Development of a sequence typing scheme for differentiation of *Salmonella Enteritidis* strains. *FEMS Microbiol Lett.* 2012;(331):165–75.
113. Best EL, Hampton MD, Ethelberg S, Liebana E, Clifton-Hadley FA, Threlfall EJ. Drug-Resistant *Salmonella* Typhimurium DT 120: Use of PFGE and MLVA in a Putative International Outbreak Investigation. *Microb Drug Resist.* 2009;15(2):133–8.
114. Dionisi AM, Carattoli A, Luzzi I, Magistrali C, Pezzotti G. Molecular genotyping of *Salmonella enterica* Abortusovis by pulsed field gel electrophoresis. *Vet Microbiol* [Internet]. 2006;116(1–3):217–23. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/16624505>
115. Sivapalasingam S, Barrett E, Kimura A, Van Duyne S, De Witt W, Ying M, et al. A multistate outbreak of *Salmonella enterica* serotype Newport infection linked to mango consumption: Impact of water-dip disinfection technology. *Clin Infect Dis.* 2003;37(12):1585–90.
116. Threlfall EJ, Hampton MD, Ward LR, Richardson IR, Lanser S, Greener T. Pulsed field gel electrophoresis identifies an outbreak of *Salmonella enterica* serotype Montevideo infection associated with a supermarket hot food outlet. *Commun Dis public Heal.* 1999;2(3):207–9.
117. Ashton PM, Nair S, Peters TM, Bale JA, Powell DG, Painset A, et al. Identification of *Salmonella* for public health surveillance using whole genome sequencing. *PeerJ.* 2016;2016(4):e1752.
118. Schwartz DC, Cantor CR. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell.* 1984 May;37(1):67–75.
119. EU *Salmonella* typing quality assessed [Internet]. [cited 2020 May 17]. Available from: <https://www.foodnavigator.com/Article/2015/11/23/EU-Salmonella-typing-quality-assessed>
120. Center for Disease Control and Prevention (U.S.). Standard operating procedure for PulseNet PFGE of *Escherichia coli* O157:H7, *Escherichia coli* non-O157 (STEC), *Salmonella* serotypes, *Shigella sonnei* and *Shigella flexneri*. 2017.
121. Kérouanton A, Marault M, Lailier R, Weill F-X, Feurer C, Espié E, et al. Pulsed-Field Gel Electrophoresis Subtyping Database for Foodborne *Salmonella enterica* Serotype Discrimination. *Foodborne Pathog Dis.* 2007;4(3):293–303.
122. Mullis K, Faloona F, Scharf S, Saiki R, Horn G, Erlich H. Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction. *Cold Spring Harb Symp Quant Biol.* 1986;51:263–73.
123. Saiki RK, Scharf S, Faloona F, Mullis KB, Horn GT, Erlich HA, et al. Enzymatic amplification of beta-globin genomic sequences and restriction site analysis for diagnosis of sickle cell anemia. *Science.* 1985;230(4732):1350–4.
124. Higuchi R, Dollinger G, Walsh PS, Griffith R. Simultaneous amplification and detection of specific DNA sequences. *Biotechnology (N Y).* 1992;10(4):413–7.
125. Lindstedt B-A, 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. *J Microbiol Methods.* 2004 Nov 1;59(2):163–72.
126. Chiou C-S, Hung C-S, Torpdahl M, Watanabe H, Tung SK, Terajima J, et al. Development and evaluation of multilocus variable number tandem repeat analysis for fine typing and phylogenetic analysis of *Salmonella enterica* serovar Typhimurium. *Int J Food Microbiol.* 2010;142(1–2):67–73.

127. Witonski D, Stefanova R, Ranganathan A, Schutze GE, Eisenach KD, Cave MD. Variable-Number Tandem Repeats That Are Useful in Genotyping Isolates of *Salmonella enterica* subsp. *enterica* Serovars Typhimurium and Newport. *J Clin Microbiol*. 2006;44(11):3849–54.
128. Beranek A, Mikula C, Rabold P, Arnhold D, Berghold C, Lederer I, et al. Multiple-locus variable-number tandem repeat analysis for subtyping of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. *Int J Med Microbiol*. 2009 Jan;299(1):43–51.
129. Malorny B, Junker E, Helmuth R. Multi-locus variable-number tandem repeat analysis for outbreak studies of *Salmonella enterica* serotype Enteritidis. *BMC Microbiol*. 2008;8(1):84.
130. Ross IL, Heuzenroeder MW. A comparison of two PCR-based typing methods with Pulsed-Field Gel Electrophoresis in *Salmonella enterica* serovar Enteritidis. *Int J Med Microbiol*. 2009;299(6):410–20.
131. Boxrud D, Pederson-Gulrud K, Wotton J, Medus C, Lyszkowicz E, Besser J, 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. *J Clin Microbiol*. 2007;45(2):536–43.
132. Liu Y, Shi X, Li Y, Chen Q, Jiang M, Li W, et al. The evaluation and application of Multi-locus Variable-number Tandem Repeat Analysis (MLVA) for the molecular epidemiological study of *Salmonella enterica* subsp. *enterica* serovar Enteritidis infection. *Ann Clin Microbiol Antimicrob*. 2016 Dec 29;15(1):4.
133. Multiple Locus Variable-number Tandem Repeat Analysis (MLVA) | PulseNet Methods | PulseNet | CDC [Internet]. [cited 2020 May 17]. Available from: <https://www.cdc.gov/pulsenet/pathogens/mlva.html>
134. Torpdahl M, Sørensen G, Lindstedt B-A, Nielsen EM. Tandem repeat analysis for surveillance of human *Salmonella* Typhimurium infections. *Emerg Infect Dis*. 2007;13(3):388–95.
135. Hopkins KL, Peters TM, de Pinna E, Wain J. Standardisation of Multi-locus Variable-number Tandem Repeat Analysis (MLVA) for subtyping of *Salmonella enterica* serovar Enteritidis. *Eurosurveillance*. 2011;16(32):1–11.
136. Nadon CA, Trees E, Ng LK, Møller NE, Reimer A, Maxwell N et al., 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. *Euro Surveill*. 2013;18(35):20565.
137. Oppen MJH van, Rico C, Turner GF, Hewitt GM. Extensive homoplasy, non-stepwise mutations, and shared ancestral polymorphism at a complex microsatellite locus in Lake Malawi Cichlids. *Mol Biol Evol*. 2000;17(4):489–98.
138. Bertrand S, De Lamine de Bex G, Wildemaue C, Lunguya O, Phoba MF, Ley B, et al. Multi Locus Variable-Number Tandem Repeat (MLVA) typing tools improved the surveillance of *Salmonella* Enteritidis: A 6 Years Retrospective Study. *PLoS One*. 2015;10(2):e0117950.
139. Hilton AC, Banks JG, Penn CW. Random amplification of polymorphic DNA (RAPD) of *Salmonella*: strain differentiation and characterization of amplified sequences. *J Appl Bacteriol*. 1996;81(6):575–84.
140. Williams JG, Kubelik AR, Livak KJ, Rafalski JA, Tingey S V. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*. 1990;18(22):6531–5.
141. Welsh J, McClelland M. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res*. 1990;18(24):7213–8.
142. Lin AW, Usera MA, Barrett TJ, Goldsby RA. Application of Random Amplified Polymorphic DNA analysis to differentiate strains of *Salmonella* Enteritidis. Vol. 34, *Journal of Clinical Microbiology*. 1996.
143. Yan S, Pendrak ML, Abela-Ridder B, Punderson JW, Fedorko DP, Foley SL. An overview of *Salmonella* typing: Public health perspectives. *Clin Appl Immunol Rev*. 2004;4(3):189–204.

144. Hashemi A, Baghbani-arani F. The effective differentiation of *Salmonella* isolates using four PCR-based typing methods. *J Appl Microbiol* [Internet]. 2015 [cited 2019 Mar 13];118(6):1530–40. Available from: <http://doi.wiley.com/10.1111/jam.12805>
145. Versalovic J, Koeuth T, Lupski JR. Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes. 1991 [cited 2019 Mar 12];19(24). Available from: <http://www.ncbi.nlm.nih.gov/pubmed/1762913>
146. Wilson LA, Sharp PM. Enterobacterial Repetitive Intergenic Consensus (ERIC) Sequences in *Escherichia coli*: Evolution and Implications for ERIC-PCR. *Mol Biol Evol* [Internet]. 2006 [cited 2019 Mar 13];23(6):1156–68. Available from: <https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msj125>
147. Adzitey F, Huda N, Ali GRR. Molecular techniques for detecting and typing of bacteria, advantages and application to foodborne pathogens isolated from ducks. *3 Biotech* [Internet]. 2013 [cited 2019 Mar 13];3(2):97–107. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28324565>
148. Grimont F, Grimont PA. Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools. *Ann Inst Pasteur Microbiol*. 137B(2):165–75.
149. Stull TL, LiPuma JJ, Edlind TD. A broad-spectrum probe for molecular epidemiology of bacteria: ribosomal RNA. *J Infect Dis*. 1988;157(2):280–6.
150. Esteban E, Snipes K, Hird D, Kasten, And R, Kinde2 H. Use of Ribotyping for Characterization of *Salmonella* Serotypes. *JOURNAL OF CLINICAL MICROBIOLOGY*. 1993.
151. Pavic M, Griffiths MW. Principles, Applications, and Limitations of Automated Ribotyping as a Rapid Method in Food Safety. *Foodborne Pathog Dis*. 2009;6(9):1047–55.
152. Clark CG, Kruk TMAC, Bryden L, Hirvi Y, Ahmed R, Rodgers FG. Subtyping of *Salmonella enterica* serotype enteritidis strains by manual and automated *Pst*I-*Sph*I ribotyping. *J Clin Microbiol*. 2003;41(1):27–33.
153. Vatopoulos AC, Mainas, Evangelos Balis, E, Threlfall EJ, Kanelopoulou M, Kalapothaki V, Malamou-Lada H, et al. Molecular Epidemiology of Ampicillin-Resistant Clinical Isolates of *Salmonella Enteritidis*. Vol. 32, *J Clin Microbiol*. 1994.
154. Maiden MCJ, van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. *Nat Rev Microbiol* [Internet]. 2013 [cited 2019 Feb 11];11(10):728–36. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/23979428>
155. Kidgell C, Reichard U, Wain J, Linz B, Torpdahl M, Dougan G, et al. *Salmonella Typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect Genet Evol*. 2002;2(1):39–45.
156. Alikhan NF, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of *Salmonella*. *PLoS Genet*. 2018;14(4):1–13.
157. Enright MC, Day NP, Davies CE, Peacock SJ, Spratt BG. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of *Staphylococcus aureus*. *J Clin Microbiol*. 2000;38(3):1008–15.
158. Liu Y-Y, Chen C-C, Chiou C-S. Construction of a Pan-Genome Allele Database of *Salmonella enterica* Serovar Enteritidis for Molecular Subtyping and Disease Cluster Identification. *Front Microbiol* [Internet]. 2016 [cited 2019 Feb 11];7:2010. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28018331>
159. McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L, et al. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*. 2001;413(6858):852–6.

160. Köser CU, Ellington MJ, Cartwright EJP, Gillespie SH, Brown NM, Farrington M, et al. Routine Use of Microbial Whole Genome Sequencing in Diagnostic and Public Health Microbiology. Rall GF, editor. PLoS Pathog. 2012 Aug;8(8):e1002824.
161. Pallen MJ, Loman NJ, Penn CW. High-throughput sequencing and clinical microbiology: progress, opportunities and challenges. *Curr Opin Microbiol*. 2010;13(5):625–31.
162. Harbottle H. Advanced methods for detection of foodborne pathogens. In: Tang,Y, Stratton C, editor. *Advanced Techniques in Diagnostic Microbiology: Applications*. 3rd ed. Switzerland: Springer Nature; 2018.
163. Weymann D, Laskin J, Roscoe R, Schrader KA, Chia S, Yip S, et al. The cost and cost trajectory of whole-genome analysis guiding treatment of patients with advanced cancers. *Mol Genet Genomic Med* [Internet]. 2017 [cited 2019 Feb 11];5(3):251–60. Available from: <http://doi.wiley.com/10.1002/mgg3.281>
164. Ellis J. WGS makes its way to forefront [Internet]. Future Lab: Biocompare. 2017. Available from: <https://www.biocompare.com/Editorial-Articles/343809-WGS-Makes-Its-Way-to-Forefront/>
165. Niehaus AJ, Apalata T, Coovadia YM, Smith AM, Moodley P. An Outbreak of Foodborne Salmonellosis in Rural KwaZulu-Natal, South Africa. *Foodborne Pathog Dis*. 2011;8(6):693–7.
166. Belkum A Van, Tassios PT, Dijkshoorn L, Haeggman S, Cookson B, Fry NK, et al. Guidelines for the validation and application of typing methods for use in bacterial epidemiology. *Clin Microbiol Infect*. 2007;13:1–46.
167. Centre HPS. Notifiable Diseases and their respective causative pathogens. 2016.
168. Scallan E, Hoekstra RM, Angulo FJ, Tauxe R V, Widdowson M-A, Roy SL, et al. Foodborne illness acquired in the United States--major pathogens. *Emerg Infect Dis*. 2011;17(1):7–15.
169. Australian Government Department of Health and Ageing PHLN. *Salmonella Laboratory Case Definition (LCD)*. 2005;
170. Crim S, Henao O, Huang J, Luna-Gierke R, Mody R, Patrick M, et al. Foodborne Diseases Active Surveillance Network (FoodNet): FoodNet Surveillance Report for 2012 [Internet]. 2012 [cited 2019 Aug 19]. Available from: https://www.cdc.gov/foodnet/PDFs/2012_annual_report_508c.pdf
171. Health AGD of. OzFoodNet - Enhancing surveillance for foodborne disease in Australia.
172. Ashley EA, Shetty N, Patel J, van Doorn R, Limmathurotsakul D, Feasey NA, et al. Harnessing alternative sources of antimicrobial resistance data to support surveillance in low-resource settings. *J Antimicrob Chemother* [Internet]. 2019 [cited 2019 Mar 19];74(3):541–6. Available from: <https://academic.oup.com/jac/article/74/3/541/5245301>
173. Herikstad H, Motarjemi, Y, Tauxe R. *Salmonella surveillance: a global survey of public health serotyping*. *Epidemiol Infect*. 2002;129:1–8.
174. Mutorwa R. *Government Gazette of the Republic of Namibia*. Windhoek; 2013.
175. *Animal Diseases Act, 1984 (Act no. 35)*. Cape Town: Department of Agriculture, Forestry and Fisheries; 1984.
176. Muvhali M, Smith AM, Rakgantso AM, Keddy KH. Investigation of *Salmonella* Enteritidis outbreaks in South Africa using multi-locus variable-number tandem-repeats analysis, 2013-2015. *BMC Infect Dis* [Internet]. 2017 [cited 2019 Feb 8];17(1):1–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28969587>
177. Keddy KH, Musekiwa A, Sooka A, Karstaedt A, Nana T, Seetharam S, et al. Clinical and microbiological features of invasive nontyphoidal *Salmonella* associated with HIV-infected patients, Gauteng Province, South Africa. *Medicine (Baltimore)* [Internet]. 2017 [cited 2019 Feb 8];96(13):e6448. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28353576>

178. Ao TT, Feasey NA, Gordon MA, Keddy KH, Angulo FJ, Crump JA. Global Burden of invasive nontyphoidal *Salmonella* disease, 2010. *Emerg Infect Dis*. 2015;21(6):941–9.
179. Tack DM, Marder EP, Griffin PM, Cieslak PR, Dunn J, Hurd S, et al. Preliminary Incidence and Trends of Infections with Pathogens Transmitted Commonly Through Food — Foodborne Diseases Active Surveillance Network, 10 U.S. Sites, 2015–2018. *MMWR Morb Mortal Wkly Rep* [Internet]. 2019 [cited 2019 Apr 30];68(16):369–73. Available from: http://www.cdc.gov/mmwr/volumes/68/wr/mm6816a2.htm?s_cid=mm6816a2_w
180. de Jong B, Ekdahl K. The comparative burden of salmonellosis in the European Union member states, associated and candidate countries. *BMC Public Health*. 2006;6(1):4.
181. Stanaway JD, Parisi A, Sarkar K, Blacker BF, Reiner RC, Hay SI, et al. The global burden of non-typhoidal *Salmonella* invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis* [Internet]. 2019 [cited 2019 Oct 11]; Available from: <https://linkinghub.elsevier.com/retrieve/pii/S1473309919304189>
182. ANGULO FJ, JOHNSON KR, TAUXE R V., COHEN ML. Origins and Consequences of Antimicrobial-Resistant Nontyphoidal *Salmonella*: Implications for the Use of Fluoroquinolones in Food Animals. *Microb Drug Resist*. 2000;6(1):77–83.
183. Threlfall EJ, Frost JA, Ward LR, Rowe B. Increasing spectrum of resistance in multi-resistant *Salmonella* Typhimurium. *Lancet* (London, England). 1996;347(9007):1053–4.
184. Boyd D, Cloeckaert A, Chaslus-Dancla E, Mulvey MR. Characterization of variant *Salmonella* genomic island 1 multidrug resistance regions from serovars Typhimurium DT104 and Agona. *Antimicrob Agents Chemother*. 2002;46(6):1714–22.
185. Shane AL, Mody RK, Crump JA, Tarr PI, Steiner TS, Kotloff K, et al. 2017 Infectious Diseases Society of America Clinical Practice Guidelines for the Diagnosis and Management of Infectious Diarrhea. *Clin Infect Dis*. 2017 Nov;65(12):e45–80.
186. Cuypers WL, Jacobs J, Wong V, Klemm EJ, Deborggraave S, Van Puyvelde S. Fluoroquinolone resistance in *Salmonella*: insights by whole-genome sequencing. *Microb Genomics*. 2018;1–9.
187. Sánchez-Vargas FM, Abu-El-Haija MA, Gómez-Duarte OG. *Salmonella* infections: An update on epidemiology, management, and prevention. *Travel Med Infect Dis*. 2011;9(6):263–77.
188. Galanis E, Wong DMALF, Patrick ME, Binsztein N, Cieslik A, Chalermchaikit T, et al. Web-based surveillance and global *Salmonella* distribution, 2000–2002. *Emerg Infect Dis*. 2006;12(3):381–8.
189. Morpeth SC, Ramadhani HO, Crump JA. Invasive non-Typhi *Salmonella* disease in Africa. *Clin Infect Dis*. 2009;49(4):606–11.
190. Feasey NA, Dougan G, Kingsley RA, Heyderman RS, Gordon MA. Invasive non-typhoidal *Salmonella* disease: an emerging and neglected tropical disease in Africa. *Lancet* (London, England). 2012;379(9835):2489–99.
191. Sigaúque B, Roca A, Mandomando I, Morais L, Quintó L, Sacarlal J, et al. Community-Acquired Bacteremia Among Children Admitted to a Rural Hospital in Mozambique. *Pediatr Infect Dis J*. 2009 Feb;28(2):108–13.
192. Lepage P, Bogaerts J, Van Goethem C, Ntahorutaba M, Nsengumuremyi F, Hitimana DG, et al. Community-acquired bacteraemia in African children. *Lancet* (London, England). 1987;1(8548):1458–61.
193. Gordon MAA, Kankwatira AMKMK, Mwafulirwa G, Walsh ALL, Hopkins MJJ, Parry CMM, et al. Invasive Non-typhoid *Salmonellae* Establish Systemic Intracellular Infection in HIV-Infected Adults: An Emerging Disease Pathogenesis. *Clin Infect Dis*. 2010 Apr;50(7):953–62.

194. Balasubramanian R, Im J, Lee J-S, Jeon HJ, Mogeni OD, Kim JH, et al. The global burden and epidemiology of invasive non-typhoidal Salmonella infections. *Hum Vaccin Immunother* [Internet]. 2018 [cited 2019 Feb 19];1–6. Available from: <https://www.tandfonline.com/doi/full/10.1080/21645515.2018.1504717>
195. The World Bank. Current classification by income [Internet]. 2018 [cited 2019 Feb 19]. Available from: <https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups>
196. GERMS-SA. GERMS-SA Annual Report 2013. 2014; Available from: [http://www.nicd.ac.za/assets/files/GERMS-SA AR 2013\(1\)\(1\).pdf](http://www.nicd.ac.za/assets/files/GERMS-SA AR 2013(1)(1).pdf)
197. GERMS-SA. GERMS-SA Annual Report: 2012. 2013; Available from: [http://www.nicd.ac.za/assets/files/GERMS-SA AR 2012\(1\)\(1\).pdf](http://www.nicd.ac.za/assets/files/GERMS-SA AR 2012(1)(1).pdf)
198. GERMS-SA. GERMS-SA Annual Report 2014 [Internet]. 2015. Available from: [http://www.nicd.ac.za/assets/files/GERMS-SA AR 2014\(1\)\(1\).pdf](http://www.nicd.ac.za/assets/files/GERMS-SA AR 2014(1)(1).pdf)
199. Muvhali M, Smith AM, Rakgantso AM, Keddy KH. Investigation of Salmonella Enteritidis outbreaks in South Africa using multi-locus variable-number tandem-repeats analysis, 2013-2015. *BMC Infect Dis* [Internet]. 2017 [cited 2019 Feb 8];17(1):1–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28969587>
200. Uche IV, MacLennan CA, Saul A. A Systematic Review of the Incidence, Risk Factors and Case Fatality Rates of Invasive Nontyphoidal Salmonella (INTS) Disease in Africa (1966 to 2014). *PLoS Negl Trop Dis*. 2017;11(1):1–28.
201. Malorny B, Schroeter A, Helmuth R. Incidence of quinolone resistance over the period 1986 to 1998 in veterinary Salmonella isolates from Germany. *Antimicrob Agents Chemother*. 1999;43(9):2278–82.
202. Cheng AC, Turnidge J, Collignon P, Looke D, Barton M, Gottlieb T. Control of fluoroquinolone resistance through successful regulation, Australia. *Emerg Infect Dis*. 2012 Sep;18(9):1453–60.
203. Collignon PC, Conly JM, Andremon A, McEwen SA, Aidara-Kane A, World Health Organization advisory group BM on IS of AR (WHO-A, et al. World Health Organization ranking of antimicrobials according to their importance in human medicine: A critical step for developing risk management strategies to control antimicrobial resistance from food animal production. Griffin PM, editor. *Clin Infect Dis*. 2016;63(8):1087–93.
204. Boore AL, Hoekstra RM, Iwamoto M, Fields PI, Bishop RD, Swerdlow DL. Salmonella enterica infections in the United States and assessment of coefficients of variation: A Novel approach to identify epidemiologic characteristics of individual serotypes, 1996-2011. Chaturvedi V, editor. *PLoS One*. 2015;10(12):1–11.
205. WHO | Antiretroviral therapy and early mortality in South Africa. WHO. 2011;
206. Hawkes C, Ruel M. The links between agriculture and health: an intersectoral opportunity to improve the health and livelihoods of the poor. *Bull World Health Organ* [Internet]. 2006 [cited 2019 Feb 11];84(12):984–90. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/17242835>
207. Allard MW, Luo Y, Strain E, Pettengill J, Timme R, Wang C, et al. On the Evolutionary History, Population Genetics and Diversity among Isolates of Salmonella Enteritidis PFGE Pattern JEGX01.0004. *PLoS One*. 2013;8(1).
208. Grassl GA, Finlay BB. Pathogenesis of enteric Salmonella infections. *Curr Opin Gastroenterol*. 2008;24(1):22–6.
209. Gerlach RG, Jäckel D, Geymeier N, Hensel M. Salmonella pathogenicity island 4-mediated adherence is coregulated with invasion genes in Salmonella enterica. *Infect Immun*. 2007;75(10):4697–709.
210. Desin TS, Lam P-KS, Koch B, Mickael C, Berberov E, Wisner ALS, et al. Salmonella enterica serovar enteritidis pathogenicity island 1 is not essential for but facilitates rapid systemic spread in chickens. *Infect Immun*. 2009;77(7):2866–75.

211. Dieye Y, Ameiss K, Mellata M, Curtiss R, III. The Salmonella Pathogenicity Island (SPI) 1 contributes more than SPI2 to the colonization of the chicken by Salmonella enterica serovar Typhimurium. BMC Microbiol. 2009;9:3.
212. Jones MA, Hulme SD, Barrow PA, Wigley P. The Salmonella pathogenicity island 1 and Salmonella pathogenicity island 2 type III secretion systems play a major role in pathogenesis of systemic disease and gastrointestinal tract colonization of Salmonella enterica serovar Typhimurium in the chicken. Avian Pathol. 2007;36(3):199–203.
213. Li S, Zhang Z, Pace L, Lillehoj H, Zhang S. Functions exerted by the virulence-associated type-three secretion systems during Salmonella enterica serovar Enteritidis invasion into and survival within chicken oviduct epithelial cells and macrophages. Avian Pathol. 2009;38(2):97–106.
214. Rychlik I, Karasova D, Sebkova A, Volf J, Sisak F, Havlickova H, et al. Virulence potential of five major pathogenicity islands (SPI-1 to SPI-5) of Salmonella enterica serovar Enteritidis for chickens. BMC Microbiol. 2009;9(1):268.
215. Wisner ALS, Desin TS, Koch B, Lam P-KS, Berberov EM, Mickael CS, et al. Salmonella enterica subspecies enterica serovar Enteritidis Salmonella pathogenicity island 2 type III secretion system: role in intestinal colonization of chickens and systemic spread. Microbiology. 2010 Sep 1;156(9):2770–81.
216. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, et al. Distinct Salmonella Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nat Genet. 2016;48(10):1211–7.
217. Chu C, Hong S-F, Tsai C, Lin W-S, Liu T-P, Ou JT. Comparative Physical and Genetic Maps of the Virulence Plasmids of Salmonella enterica Serovars Typhimurium, Enteritidis, Choleraesuis, and Dublin. Vol. 67, INFECTION AND IMMUNITY. 1999.
218. Feasey NA, Hadfield J, Keddy KH, Dallman TJ, Jacobs J, Deng X, et al. Distinct Salmonella Enteritidis lineages associated with enterocolitis in high-income settings and invasive disease in low-income settings. Nat Genet. 2016;48(10):1211–7.
219. Rotger R, Casadesús J. The virulence plasmids of Salmonella. Vol. 2, Article in International Microbiology. 1999.
220. Rodríguez-Peña JM, Buisán M, Ibáñez M, Rotger R. Genetic map of the virulence plasmid of Salmonella Enteritidis and nucleotide sequence of its replicons. Gene. 1997;188(1):53–61.
221. Swaminathan B, Barrett TJ, Hunter SB, Tauxe R V., CDC PulseNet Task Force. PulseNet: The Molecular Subtyping Network for Foodborne Bacterial Disease Surveillance, United States. Emerg Infect Dis. 2001;7(3):382–9.
222. Hudson CR, Garcia M, Gast RK, Maurer JJ. Determination of close genetic relatedness of the major Salmonella Enteritidis phage types by pulsed-field gel electrophoresis and DNA sequence analysis of several Salmonella virulence genes. Avian Dis. 45(4):875–86.
223. Liebana E, Garcia-Migura L, Breslin MF, Davies RH, Woodward MJ. Diversity of Strains of Salmonella enterica Serotype Enteritidis from English Poultry Farms Assessed by Multiple Genetic Fingerprinting. J Clin Microbiol. 2001;39(1):154–61.
224. Karberg KA, Olsen GJ, Davis JJ. Similarity of genes horizontally acquired by Escherichia coli and Salmonella enterica is evidence of a supraspecies pangenome. Proc Natl Acad Sci U S A. 2011;108(50):20154–9.
225. Figueroa-Bossi N, Uzzau S, Maloriol DD, Bossi L. Variable assortment of prophages provides a transferable repertoire of pathogenic determinants in Salmonella. Mol Microbiol [Internet]. 2001 [cited 2019 Jan 30];39(2):260–71. Available from: <http://doi.wiley.com/10.1046/j.1365-2958.2001.02234.x>
226. Wagner PL, Waldor MK. Bacteriophage control of bacterial virulence. Infect Immun [Internet]. 2002 [cited 2019 Mar 13];70(8):3985–93. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12117903>

227. Thomson N, Baker S, Pickard D, Fookes M, Anjum M, Hamlin N, et al. The role of prophage-like elements in the diversity of *Salmonella enterica* serovars. *J Mol Biol.* 2004;339(2):279–300.
228. Nuccio S-P, Bäumler AJ. Reconstructing pathogen evolution from the ruins. *Proc Natl Acad Sci U S A.* 2015;112(3):647–8.
229. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res.* 2009;19(12):2279–87.
230. Feasey NA, Masesa C, Jassi C, Faragher EB, Mallewa J, Mallewa M, et al. Three epidemics of invasive multidrug-resistant *Salmonella* bloodstream infection in Blantyre, Malawi, 1998–2014. *Clin Infect Dis.* 2015;61 Suppl 4(Suppl 4):S363–71.
231. Reddy EA, Shaw A V, Crump JA. Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis. *Lancet Infect Dis.* 2010;10(6):417–32.
232. Kingsley RA, Msefula CL, Thomson NR, Kariuki S, Holt KE, Gordon MA, et al. Epidemic multiple drug resistant *Salmonella* Typhimurium causing invasive disease in sub-Saharan Africa have a distinct genotype. *Genome Res.* 2009;19(12):2279–87.
233. Okoro CK, Barquist L, Connor TR, Harris SR, Clare S, Stevens MP, et al. Signatures of adaptation in human invasive *Salmonella* Typhimurium ST313 populations from Sub-Saharan Africa. Yang R, editor. *PLoS Negl Trop Dis* [Internet]. 2015 [cited 2019 Mar 13];9(3):e0003611. Available from: <https://dx.plos.org/10.1371/journal.pntd.0003611>
234. Owen S V, Wenner N, Canals R, Makumi A, Hammarlöf DL, Gordon MA, et al. Characterization of the prophage repertoire of African *Salmonella* Typhimurium ST313 reveals high levels of spontaneous induction of novel Phage BTP1. *Front Microbiol* [Internet]. 2017 [cited 2019 Mar 13];8:235. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28280485>
235. Rodrigue DC, Tauxe R V, Rowe B. International increase in *Salmonella* Enteritidis: a new pandemic? *Epidemiol Infect.* 1990 Aug;105(1):21–7.
236. Keddy KH, Musekiwa A, Sooka A, Karstaedt A, Nana T, Seetharam S, et al. Clinical and microbiological features of invasive nontyphoidal *Salmonella* associated with HIV-infected patients, Gauteng Province, South Africa. *Med (United States)* [Internet]. 2017 [cited 2019 Feb 8];96(13):1–7. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/28353576>
237. Caffer MI, Eiguer T. *Salmonella* Enteritidis in Argentina. *Int J Food Microbiol.* 1994;21(1–2):15–9.
238. Roberts JA, Sockett PN. The socio-economic impact of human *Salmonella* enteritidis infection. *Int J Food Microbiol.* 1994;21(1–2):117–29.
239. Olsen SJ, Bishop R, Brenner FW, Roels TH, Bean N, Tauxe R V., et al. The changing epidemiology of *Salmonella*: Trends in serotypes isolated from humans in the United States, 1987–1997. *J Infect Dis.* 2001;183(5):753–61.
240. Rodrigue DC, Tauxe R V., Rowe B. International increase in *Salmonella* Enteritidis: A new pandemic? *Epidemiol Infect.* 1990;105(01):21–7.
241. Shah DH, Zhou X, Kim HY, Call DR, Guard J. Transposon mutagenesis of *Salmonella enterica* serovar Enteritidis identifies genes that contribute to invasiveness in human and chicken cells and survival in egg albumen. *Infect Immun.* 2012;80(12):4203–15.
242. Rabsch W, Tschäpe H, Bäumler AJ. Non-typhoidal salmonellosis: emerging problems. *Microbes Infect.* 2001;3(3):237–47.

243. Botteldoorn N, Van Coillie E, Goris J, Werbrouck H, Piessens V, Godard C, et al. Limited genetic diversity and gene expression differences between egg- and non-egg-related *Salmonella* Enteritidis strains. Vol. 57, *Zoo And Pub. H.* Blackwell Verlag; 2010. 345–357 p.
244. Liu F, Kariyawasam S, Jayarao BM, Barrangou R, Gerner-Smidt P, Ribot EM, et al. Subtyping *Salmonella enterica* serovar Enteritidis isolates from different sources by using sequence typing based on virulence genes and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPRs). *Appl Environ Microbiol.* 2011;77(13):4520–6.
245. Olson AB, Andrysiak AK, Tracz DM, Guard-Bouldin J, Demczuk W, Ng L-K, et al. Limited genetic diversity in *Salmonella enterica* serovar Enteritidis PT13. *BMC Microbiol.* 2007;7:87.
246. Bossi L, Fuentes JA, Mora G, Figueroa-Bossi N. Prophage contribution to bacterial population dynamics. *J Bacteriol* [Internet]. 2003 [cited 2019 Jan 30];185(21):6467–71. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/14563883>
247. Gordon MA, Banda HT, Gondwe M, Gordon SB, Boeree MJ, Walsh AL, et al. Non-typhoidal *Salmonella* bacteraemia among HIV-infected Malawian adults: high mortality and frequent recrudescence. *AIDS* [Internet]. 2002 [cited 2019 Jan 30];16(12):1633–41. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/12172085>
248. James SL, Abate D, Abate KH, Abay SM, Abbafati C, Abbasi N, et al. Global, regional, and national incidence, prevalence, and years lived with disability for 354 Diseases and Injuries for 195 countries and territories, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *Lancet.* 2018 Nov 10;392(10159):1789–858.
249. Gordon MA, Feasey NA, Nyirenda TS, Graham SM. Nontyphoid *Salmonella* Disease. *Hunter's Trop Med Emerg Infect Dis* [Internet]. 2020 [cited 2019 Jun 24];500–6. Available from: <https://www.sciencedirect.com/science/article/pii/B9780323555128000491>
250. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* [Internet]. 2014 [cited 2019 Oct 16];30(14):2068–9. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btu153>
251. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* [Internet]. 2015 [cited 2019 Oct 16];31(22):3691–3. Available from: <https://academic.oup.com/bioinformatics/article-lookup/doi/10.1093/bioinformatics/btv421>
252. qPCR Efficiency Calculator - ZA. [cited 2019 Apr 16]; Available from: <https://www.thermofisher.com/za/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/qpcr-efficiency-calculator.html>
253. Raymaekers M, Smets R, Maes B, Cartuyvels R. Checklist for Optimization and Validation of Real-Time PCR Assays. *J Clin Lab Anal* [Internet]. 2009 [cited 2019 Mar 28];23:145–51. Available from: www.interscience.wiley.com
254. Broeders S, Huber I, Grohmann L, Berben G, Taverniers I, Mazzara M, et al. Guidelines for validation of qualitative real-time PCR methods. *Trends Food Sci Technol.* 2014;37(2):115–26.
255. McNerney R. Diagnostics for Developing Countries. *Diagnostics* (Basel, Switzerland) [Internet]. 2015 [cited 2019 Jun 6];5(2):200–9. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/26854149>
256. Abou Tayoun AN, Burchard PR, Malik I, Scherer A, Tsongalis GJ. Democratizing Molecular Diagnostics for the Developing World. *Am J Clin Pathol* [Internet]. 2014 [cited 2019 Jun 20];141(1):17–24. Available from: <https://academic.oup.com/ajcp/article/141/1/17/1766064>
257. Helmy M, Awad M, Mosa K. Limited resources of genome sequencing in developing countries: Challenges and solutions. *Appl Transl Genomics.* 2016;9:15–9.

258. CDC. Prioritizing Zoonotic Diseases for Multisectoral, One Health Collaboration in the United States [Internet]. 2019. Available from: <https://www.cdc.gov/onehealth/pdfs/us-ohzdp-report-508.pdf>
259. Group for Enteric R and M disease S in SA. GERMS-SA Annual Report 2013 [Internet]. 2013 [cited 2019 Feb 12]. Available from: [http://www.nicd.ac.za/assets/files/GERMS-SA AR 2013\(1\).pdf](http://www.nicd.ac.za/assets/files/GERMS-SA AR 2013(1).pdf)
260. GERMS-SA. GERMS-SA Annual Report 2015. 2016.
261. Campioni F, Cao G, Kastanis G, Janies DA, Bergamini AMM, Rodrigues DDP, et al. Changing of the genomic pattern of Salmonella Enteritidis strains isolated in Brazil over a 48 year-period revealed by Whole Genome SNP Analyses. *Sci Rep*. 2018;8(1):1–7.
262. Department of Agriculture F and F. A profile of the South African broiler market value chain [Internet]. [cited 2019 Oct 16]. Available from: <https://www.nda.agric.za/doaDev/sideMenu/Marketing/Annual Publications/Commodity Profiles/field crops/Broiler market value chain profile 2015.pdf>
263. Crump JA, Youssef FG, Luby SP, Wasfy MO, Rangel JM, Taalat M, et al. Estimating the incidence of typhoid fever and other febrile illnesses in developing countries. *Emerg Infect Dis*. 2003 May 1;9(5):539–44.
264. World Health Organization: Food and Agricultural Organization of the United Nation. Risk Assessments of Salmonella in Eggs and Broiler Chickens. Second. Rome, Italy: WHO Library Cataloguing-in-Publication Data
265. Feasey NA, Archer BN, Heyderman RS, Sooka A, Dennis B, Gordon MA, et al. Typhoid fever and invasive nontyphoid salmonellosis, Malawi and South Africa. *Emerg Infect Dis*. 2010;16(9):1448–51.
266. Zheng J, Pettengill J, Strain E, Allard MW, Ahmed R, Zhao S, et al. Genetic Diversity and Evolution of Salmonella enterica Serovar Enteritidis Strains with Different Phage Types. *J Clin Microbiol* [Internet]. 2014 [cited 2019 Sep 5];52(5):1490. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3993623/>
267. Public Health England. Outbreak of Salmonella Enteritidis phage type 14b, May-September 2014 [Internet]. 2015 [cited 2019 Oct 16]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/432637/S_Enteritidis_PT14B_outbreak_report_gateway__2_.pdf
268. Silva E, Duarte A. Salmonella Enteritidis in Poultry: Retrospective in Brazil. *Rev Bras Ciência Avícola* [Internet]. 2002 [cited 2019 Sep 5]; 4(2):85–100. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-635X2002000200001&lng=pt&tlng=pt
269. Liu H, Whitehouse CA, Li B. Presence and persistence of Salmonella in water: The Impact on Microbial Quality of Water and Food Safety. *Front Public Heal*. 2018 May 30;6.
270. Landeras E, González-Hevia MA, Mendoza MC. Molecular epidemiology of Salmonella serotype Enteritidis. Relationships between food, water and pathogenic strains. *Int J Food Microbiol*. 1998 Aug 18;43(1–2):81–90.
271. MacKenzie KD, Wang Y, Musicha P, Hansen EG, Palmer MB, Herman DJ, et al. Parallel evolution leading to impaired biofilm formation in invasive Salmonella strains. *PLOS Genet*. 2019 Jun 24;15(6):e1008233.
272. Aldrich C, Hartman H, Feasey N, Chattaway MA, Dekker D, Al-Emran HM, et al. Emergence of phylogenetically diverse and fluoroquinolone resistant Salmonella Enteritidis as a cause of invasive nontyphoidal Salmonella disease in Ghana. Akullian A, editor. *PLoS Negl Trop Dis* [Internet]. 2019 Jun 20 [cited 2019 Sep 5];13(6):e0007485. Available from: <http://dx.plos.org/10.1371/journal.pntd.0007485>
273. Human Sciences Research Council. HIV impact assessment summary. 2018.
274. Essential Drugs Program. Primary Healthcare Standard Treatment Guidelines and Essential Medicines List [Internet]. 5th ed. Pretoria, South Africa: The National Department of Health; 2014 [cited 2019 Oct 14]. Available from: <https://sahivsoc.org/Files/Essential drug list.pdf>

275. Marks F, von Kalckreuth V, Aaby P, Adu-Sarkodie Y, El Tayeb MA, Ali M, et al. Incidence of invasive *Salmonella* disease in sub-Saharan Africa: a multicentre population-based surveillance study. *Lancet Glob Heal*. 2017 Mar 1;5(3):e310–23.
276. Crowther-Gibson, P, Govender, N, Lewis, D, Bamford, C, Brink, A, von Gottenberg A, Klugman, K, du Plessis, M, Fali, A, Harris, B, Keddy, K, Botha M. Human infections and antibiotic resistance. *South African Med J*. 2011 Jul 28;101(8):567–78.
277. The European Committee on Antimicrobial Susceptibility Testing. Breakpoint tables for interpretation of MICs and zone diameters [Internet]. 2019. Available from: <http://www.eucast.org>.
278. Gouvêa R, Santos F dos, Aquino M de, Pereira VL de A, Gouvêa R, Santos F dos, et al. Fluoroquinolones in industrial poultry production, bacterial resistance and food residues:a review. *Rev Bras Ciência Avícola* [Internet]. 2015 [cited 2019 Sep 13];17(1):1–10. Available from: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1516-635X2015000100001&lng=en&tlng=en
279. Agyare C, Etsiapa Boamah V, Ngofi Zumbi C, Boateng Osei F. Antibiotic Use in Poultry Production and Its Effects on Bacterial Resistance. In: *Antimicrobial Resistance - A Global Threat*. IntechOpen; 2019.
280. South African Poultry Association. South African Poultry Meat Imports: Tariff code report for January 2019. 2019.
281. Magwedere K, Rauff D, De Klerk G, Keddy KH, Dziva F. Incidence of nontyphoidal *Salmonella* in food-producing animals, animal feed, and the associated environment in South Africa, 2012-2014. *Clin Infect Dis*. 2015 Nov 1;61:S283–9.
282. Govender N. Molecular Epidemiology and Mechanism of Resistance of Invasive QuinoloneResistant South African Isolates of *Salmonella enterica*, 2004-2006 . 2009.
283. Kilmartin D, Morris D, O'Hare C, Corbett-Feeney G, Cormican M. Clonal expansion may account for high levels of quinolone resistance in *Salmonella enterica* serovar enteritidis. *Appl Environ Microbiol*. 2005 May;71(5):2587–91.

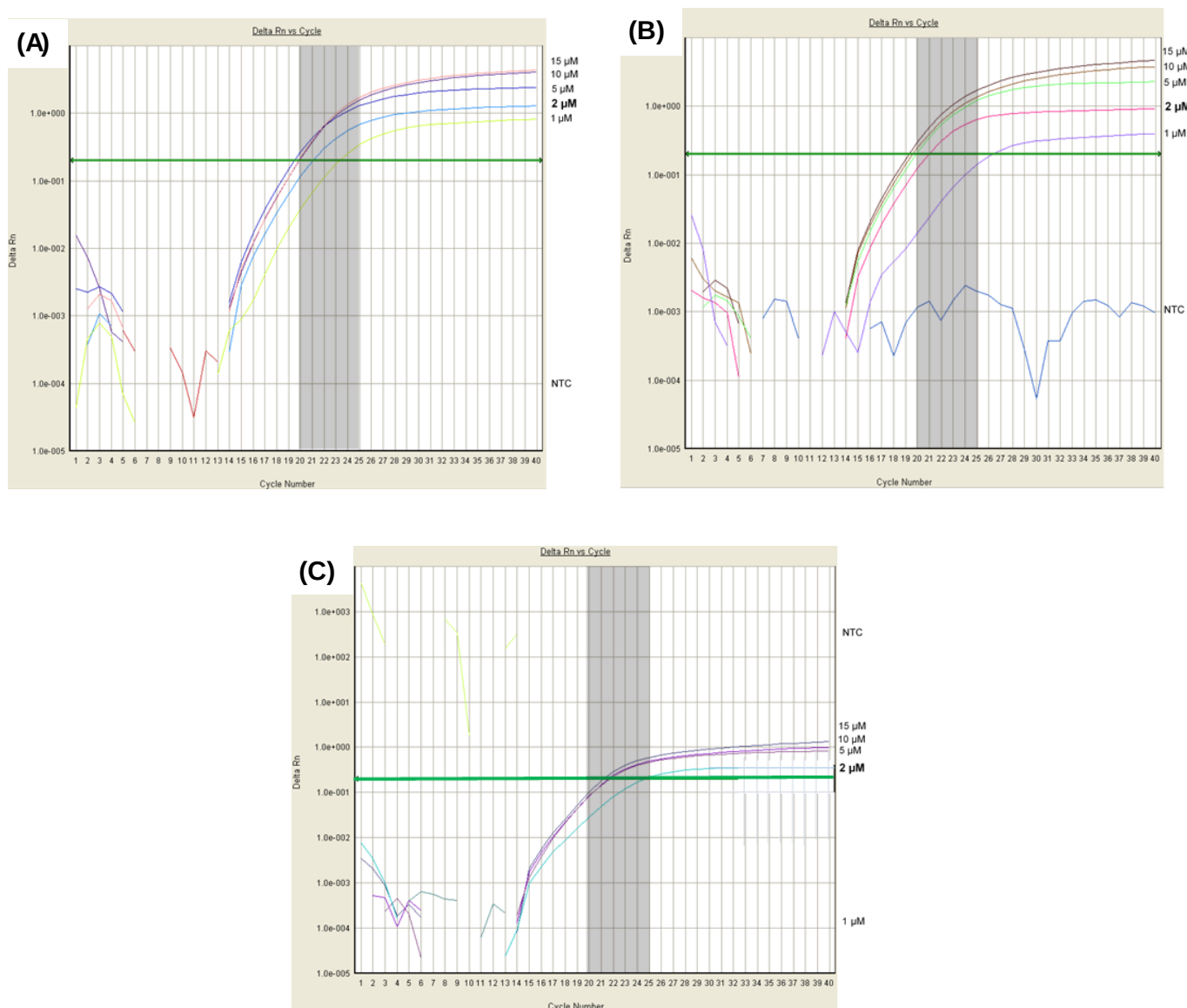
SUPPLEMENTARY MATERIAL

Supplementary Table 1 The specific region within the validation panel isolates targeted and amplified by the corresponding primer sets *in silico*

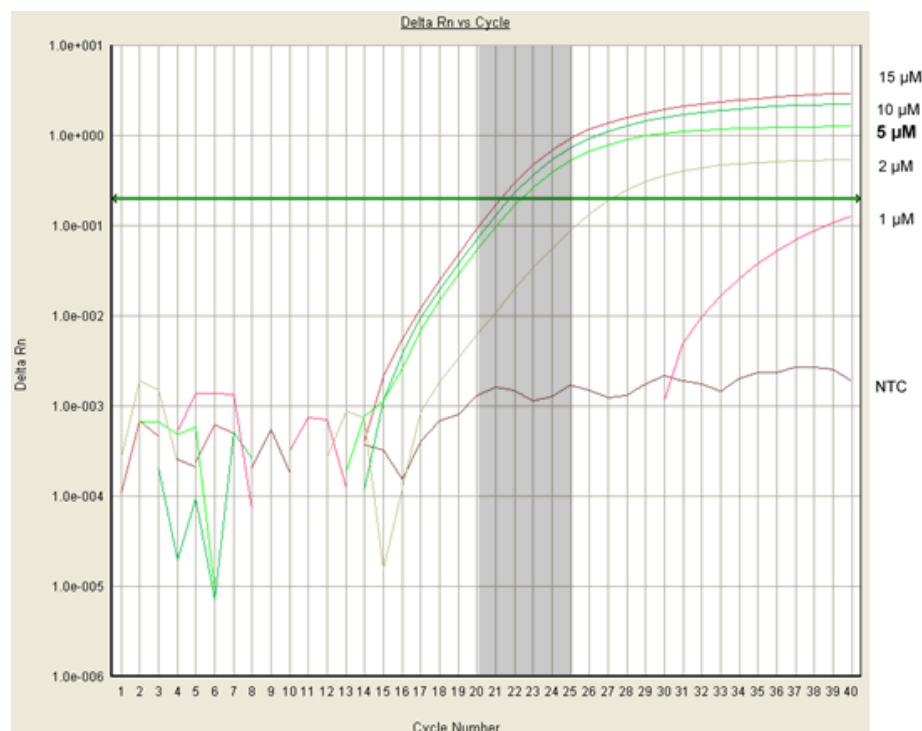
Isolate	Origin	Amplified Fragment	Length	Region
D7795	Malawi	African Clade small toxic peptide gene	82	3972971...3973052
		East African Clade antitoxin pemI gene	101	4803260...4803360
255 DRC	DRC	East African Clade antitoxin pemI gene	101	4918605...4918705
		African Clade small toxic peptide gene	82	1322611...1322692
		East African Clade antitoxin pemI gene	101	4797666...4797766
6396	Uganda	African Clade small toxic peptide gene	82	1980642...1980723
		East African Clade antitoxin pemI gene	101	4560050...4560150
10136/01	Gambia	African Clade small toxic peptide gene	82	1970326...1970407
0527/01		African Clade small toxic peptide gene	82	1981992...1982073
8078/01		African Clade small toxic peptide gene	82	1908167...1908248
P125109	UK	Global and Global Outlier Clade TIR protein domain gene	126	742300...742425
		Global Clade chitinase class I gene	85	716719...716803
A1636	Malawi	Global and Global Outlier Clade TIR protein domain gene	126	3675375...3675500
		Global Clade chitinase class I gene	85	3649794...3649878
1320	Malawi	Global and Global Outlier Clade TIR protein domain gene	126	2167896...2168021
		Global Clade chitinase class I gene	85	2142315...2142399
LSS-40	UK	Global and Global Outlier Clade TIR protein domain gene	126	2125549...2125674
791	Uganda	Global and Global Outlier Clade TIR protein domain gene	126	2165634...2165759

Supplementary Table 2 Specific probe binding site in each of the corresponding validation panel isolates *in silico*

	Validation panel isolate	Origin	Binding site
African Category Probe	D7795	Malawi	3973002...3973025
	255	DRC	1322642...1322665
	6396	Uganda	1980673...1980696
	10136/01	Gambia	1970357...1970380
	0527/01	Gambia	1982023...1982046
	8078/01	Gambia	1908198...1908221
East African Clade Probe	D7795	Malawi	4803298...4803321
	255	DRC	4797705...4797728
	6396	Uganda	4560089...4560112
Global/ Global Outlier Category Probe	P125109	UK	742372...742396
	A1636	Malawi	3675447...3675471
	1320	Uganda	2167968...2167992
	LSS-40	UK	2165706...2165730
	791	Uganda	2125621...2125645
Global Clade Probe	P125109	UK	716758...716781
	A1636	Malawi	3649833...3649856
	1320	Uganda	2142354...2142377



Supplementary Figure 1 Standard real-time PCR curves of the normalized reporting value (Delta Rn) of the (A) African Category assay, (B) East African Clade assay and (C) Global Clade assay with different primer concentrations plotted against the cycle number. The grey area depicts the acceptable cycle threshold (horizontal green line) value range for an optimal



Supplementary Figure 2 Standard real-time PCR curves of the normalized reporting value (Delta Rn) of the Global/ Outlier Category assay with different primer concentrations plotted against the cycle number. The grey area depicts the acceptable cycle threshold (horizontal green line) value range for an optimal assay.

Supplementary Table 3 Cycle threshold values of category and clade amplicons using different concentrations of the corresponding primer set

	Category		Clade	
	African	Global/ Global Outlier	East African	Global
Primer Concentration (µM)	Ct value			
15	19.44	21.75	19.23	21.17
10	19.87	21.25	19.56	21.6
5	20.92	22.36	19.88	21.67
2	21.06	27.2	20.97	22.88
1	23.33	Undetermined	26.37	24.67

Supplementary Table 4 Cycle threshold values of category and clade amplicons using different concentrations of the corresponding probe

	Category		Clade	
	African	Global/ Global Outlier	East African	Global
Probe Concentration (µM)	Ct value			
5	22.60	21.67	20.60	21.70
2	25.53	23.5	24.41	26.07

```

1 //import Excel spreadsheet of all epidemiological data and real-time PCR results
2 import excel "C:\Users\Sarah Gallichan\Documents\Masters\Results\Master results.xlsx", sheet("Sheet1") firstrow clear
3
4 //descriptive epidemiology data
5 tab GENDER
6 tab Agegroup
7
8 //generate numerical Clade data from string Clade data (rtPCR results)
9 encode Clade, generate(Clade_n)
10
11 //create binary Clade_n data
12 recode Clade_n 2=0
13
14 //label Clade_n data
15 label define L_clade 0 "Global Outlier" 1 "Global"
16 label value Clade_n L_clade
17
18 //generate numerical Invasive data from string Invasive data
19 encode Invasive, generate(Invasive_n)
20
21 //create binary Invasive_n data
22 recode Invasive_n 1=0 2=1
23
24 //label Invasive_n data
25 label define L_inyes 0 "No" 1 "Yes"
26 label value Invasive_n L_inyes
27
28 //generate numerical Provincial data from string PROVLAB data
29 encode PROVLAB, generate(Province_n)
30
31 //generate numerical NAL & CIP data from string NAL & CIP data
32 encode NAL, generate(NAL_n)
33 encode CIP, generate(CIP_n)
34 encode SUL, generate(SUL_n)
35 encode TET, generate(TET_n)
36 encode TRI, generate(TRI_n)
37 encode AMP, generate(AMP_n)
38 encode CHL, generate(CHL_n)
39 encode COT, generate(COT_n)
40 encode AUG, generate(AUG_n)
41 encode CAS, generate(CAS_n)
42 encode CRO, generate(CRO_n)
43 encode STR, generate(STR_n)
44 encode KAN, generate(KAN_n)
45 encode IMI, generate(IMI_n)
46
47 //create binary NAL_n & CIP_n data
48 recode NAL_n 2=0
49 recode CIP_n 2=0
50 recode SUL_n 2=0
51 recode TET_n 2=0
52 recode TRI_n 2=0
53 recode AMP_n 2=0
54 recode CHL_n 2=0
55 recode COT_n 2=0
56 recode AUG_n 2=0
57 recode CAS_n 2=0
58 recode CRO_n 2=0
59 recode STR_n 2=0
60 recode KAN_n 2=0
61 recode IMI_n 2=0
62
63 //label Invasive_n data
64 label define L_resistant 0 "Susceptible" 1 "Resistant"
65 label value NAL_n L_resistant
66 label value CIP_n L_resistant
67 label value SUL_n L_resistant
68 label value TET_n L_resistant
69 label value TRI_n L_resistant
70 label value AMP_n L_resistant
71 label value CHL_n L_resistant
72 label value COT_n L_resistant
73 label value AUG_n L_resistant
74 label value CAS_n L_resistant
75 label value CRO_n L_resistant
76 label value STR_n L_resistant
77 label value KAN_n L_resistant
78 label value IMI_n L_resistant
79
80 //generate numerical Agegroup data from string Agegroup data
81 encode Agegroup, generate(Agegroup_n)
82
83 //determine geographical association for clades and provide proportions by row
84 tab Province_n Clade_n, chi exact row
85
86 //determine association between isolate factors and clades and provide proportion by column
87 tab Invasive_n Clade_n, chi exact col
88 tab NAL_n Clade_n, chi exact col
89 tab CIP_n Clade_n, chi exact col
90 tab SUL_n Clade_n, chi exact col
91 tab TET_n Clade_n, chi exact col
92 tab TRI_n Clade_n, chi exact col
93 tab AMP_n Clade_n, chi exact col
94 tab CHL_n Clade_n, chi exact col
95 tab COT_n Clade_n, chi exact col
96 tab AUG_n Clade_n, chi exact col
97 tab CAS_n Clade_n, chi exact col
98 tab CRO_n Clade_n, chi exact col
99 tab STR_n Clade_n, chi exact col
100 tab KAN_n Clade_n, chi exact col
101 tab IMI_n Clade_n, chi exact col
102
103 //determine association between patient factors and clades and provide proportion by column
104 tab Agegroup_n Clade_n, chi col
105 tab Sex_n Clade_n, chi exact col
106
107 //assess potential association between province and significant isolate factors
108 tab Invasive_n Province_n, chi exact col
109 tab NAL_n Province_n, chi exact col
110 tab CIP_n Province_n, chi exact col
111 tab SUL_n Province_n, chi exact col
112 tab TET_n Province_n, chi exact col
113 tab TRI_n Province_n, chi exact col
114 tab AMP_n Province_n, chi exact col
115 tab CHL_n Province_n, chi exact col
116 tab COT_n Province_n, chi exact col
117 tab CIP_n Province_n, chi exact col
118
119 //perform logistical regression for significant isolate factors for the Global Clade, adjusting for stratified Province where necessary
120 logit Clade_n Invasive_n i.Province, or
121 logit Clade_n NAL_n i.Province, or
122 logit Clade_n CIP_n i.Province, or

```

Supplementary Figure 3 Stata Do file containing the commands run for the statistical analysis of the association between clade and patient (gender and age) and isolate (invasive disease and antimicrobial resistance) data

```

1 //import Excel spreadsheet of epidemiological data and real-time PCR results for isolates with HIV test results
2 import excel "C:\Users\Sarah Gallichan\Documents\Masters\Results\Master results HIV subset.xlsx", sheet("Sheet1") firstrow clear
3
4 //generate numerical HIV data from string HIV data (rtPCR results)
5 encode HIVStatus, generate(HIV_n)
6
7 //create binary HIV_n data
8 recode HIV_n 2=0
9
10 //label HIV_n data
11 label define L_HIV 0 "Negative" 1 "Positive"
12 label value HIV_n L_HIV
13
14 //determine association between invasive disease and HIV status and provide proportion by column
15 tab Clade_n HIV_n, chi exact col
16
17 //determine association between invasive disease and HIV status and provide proportion by column
18 tab Invasive_n HIV_n, chi exact col
19
20 //further assess association between HIV and invasive disease and clade
21 logit Clade_n HIV_n, or
22 logit Invasive_n HIV_n, or

```

Supplementary Figure 4 Stata Do file containing the commands run for the statistical analysis of the association between HIV status and clade and invasive disease

Supplementary Table 5 Relationship between HIV status (positive or negative) and type of NTS disease (Invasive or gastroenteritis)

		Type of Disease		Total
		Invasive	Gastroenteritis	
HIV status	Positive	148 (74.0)	17 (60.7)	228
	Negative	52 (26.0)	11 (39.3)	

Supplementary Table 6 Association between invasive disease and age group (5 years and above) of the patient from whom the *Salmonella* Enteritidis isolate was collected

Age group (years)	Proportion (%)	Odds Ratio (95% CI)	p-value
5 - 14	12/49 (24.5)	0.32 (0.17 - 0.62)	0.001
15 - 24	16/44 (36.4)	1.76 (0.72 - 4.67)	0.215
25 - 34	58/85 (68.2)	6.62 (2.99 - 14.67)	0.000
35 - 44	60/87 (69.0)	6.85 (3.10 - 15.16)	0.000
45 - 54	44/72 (61.1)	4.85 (2.17 - 10.84)	0.000
55 - 64	26/40 (65.0)	5.72 (2.28 - 14.37)	0.000
65 - ≥70	20/45 (44.4)	2.47 (1.03 - 5.93)	0.044

Supplementary Table 7 Association between HIV positive patients and age group (5 years and above) of the patient from whom the *Salmonella* Enteritidis isolate was collected

Age group (years)	Proportion (%)	Odds Ratio (95% CI)	p-value
5 - 14	7/12 (58.3)	1.40 (0.44 - 4.41)	0.566
15 - 24	14/17 (82.4)	3.33 (0.61 - 18.14)	0.164
25 - 34	44/47 (93.6)	10.48 (2.04 - 53.93)	0.005
35 - 44	38/41 (92.7)	9.05 (1.75 - 46.77)	0.009
45 - 54	21/30 (70.0)	1.67 (0.42 - 6.68)	0.471
55 - 64	14/15 (93.3)	10.00 (0.97 - 102.87)	0.053
65 - ≥70	3/6 (50.0)	0.71 (0.10 - 5.12)	0.738



R14/49 Dr Vanessa Quan

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M1809107

NAME: Dr Vanessa Quan
(Principal Investigator)
DEPARTMENT: National Institute for Communicable Diseases

PROJECT TITLE: GERMS-SA: Laboratory-based Surveillance for Pathogens of Public Health Importance in South Africa

DATE CONSIDERED: Ad Hoc

DECISION: Approved unconditionally

CONDITIONS: Recertification - M140159 (M081117)

SUPERVISOR:

APPROVED BY: 
Doctor CB Penny, Chairperson, HREC (Medical)

DATE OF APPROVAL: 08/11/2018

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the Third Floor, Faculty of Health Sciences, Phillip Tobias Building, 29 Princess of Wales Terrace, Parktown, 2193, University of the Witwatersrand. I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report.** The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in **September** and will therefore be due in the month of **September** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).


Principal Investigator Signature

12/11/2018
Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES



**NATIONAL INSTITUTE FOR
COMMUNICABLE DISEASES**

Division of the National Health Laboratory Service

**DIVISION OF PUBLIC HEALTH SURVEILLANCE &
RESPONSE**

1 Modderfontein Road, Sandringham, 2192

Tel: +27 (0)11 386-6337

Fax: +27 (0)86 758 3326

6 August 2018

Prof Clement Penny
Chairperson
Wits HREC (Medical)
University of the Witwatersrand

Dear Prof Penny

Re: permission for Sarah Gallichan (1726349) to use GERMS-SA (M140159) database fields

As head of the GERMS-SA surveillance programme, I hereby give permission for Sarah Gallichan to have access to the GERMS-SA *Salmonella* (non typhoid) dataset for the years 2012 until 2013 for use in her secondary data analysis for her Masters project. The project titled Snapshot of South African *Salmonella*: Development of a sensitive tool to investigate the distribution of *Salmonella* Enteritidis clades in 2012 and 2013 is for a Masters in Medical Pathology.

She will be granted access to the following variables that are required for her analysis:

- ☐ TCD numbers
- ☐ Year of collection (2012 and 2013)
- ☐ Serotype of *Salmonella* isolated
- ☐ Provincial lab where *Salmonella* isolate was taken (Gauteng, KwaZulu Natal, Western Cape, Mpumalanga)
- ☐ Gender of patient
- ☐ Age of patient
- ☐ HIV status of patient
- ☐ TB status of patient
- ☐ Antimicrobial resistance data.

The data will be available to her on presentation of the ethics clearance certificate from Wits HREC (Medical).

Yours sincerely

Dr Vanessa Quan
Head of GERMS-SA surveillance programme
Division of Public Health Surveillance and Response
011 386 6012
vanessaq@nicd.ac.za



NATIONAL HEALTH LABORATORY SERVICE

Academic Affairs and Research
Modderfontein Road, Sandringham, 2031
Tel: +27 (0)11 386 6142
Fax: +27 (0)11 386 6296
Email: babatyimalopekgokong@nhls.ac.za
Web: www.nhls.ac.za

22 March 2019

Applicant: Ms Sarah Gallichan
Institution: National Institute for Communicable Diseases
Department: Centre for Enteric Diseases - Bacteriology
Email: sarahhg@nicd.ac.za
Tel: 011 555 0368
Cell: 083 291 0805

Re: Approval to access National Health Laboratory Service (NHLS) Data

Your application to undertake a research project "**Snapshot of South African *Salmonella*: Development of a sensitive tool to investigate the distribution of *Salmonella* Enteritidis clades in 2012 and 2013**" using data from the NHLS database has been reviewed. This letter serves to advise that the application has been approved and the required data will be made available to you ***without patients names*** to conduct the proposed study as outlined in the submitted application.

Please note that final approval is granted on your compliance with the NHLS conditions of service and that the study can only be undertaken provided that the following conditions have been met.

- Processes are discussed with the relevant NHLS departments (i.e. Information Management Unit and Operations Office) and are agreed upon.
- Confidentiality is maintained at participant and institutional level and there is no disclosure of personal information or confidential information as described by the NHLS policy.
- A final report of the research study and any published paper resulting from this study are submitted and addressed to the NHLS Academic Affairs and Research office and the NHLS has been acknowledged appropriately.
- NHLS Data cannot be used to track patients as no pre-approval/consent is obtained from Patients.

Please note that this letter constitutes approval by the NHLS Academic Affairs and Research Office. Any data related queries may be directed to NHLS Corporate Data Warehouse, contact number: 011 386 6074 email: zarina.sabat@nhls.ac.za

PP: Babatyi

Dr Babatyi Malope-Kgokong
National Manager: Academic Affairs and Research

Plagiarism Report

MSc Dissertation FINAL

ORIGINALITY REPORT

13%	11%	8%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	bmcinfectdis.biomedcentral.com Internet Source	1%
2	journals.plos.org Internet Source	1%
3	link.springer.com Internet Source	<1%
4	aem.asm.org Internet Source	<1%
5	wwwnc.cdc.gov Internet Source	<1%
6	www.biomedcentral.com Internet Source	<1%
7	orca.cf.ac.uk Internet Source	<1%
8	www.science.gov Internet Source	<1%
9	repository.up.ac.za Internet Source	<1%