The Open University

Open Research Online

The Open University's repository of research publications and other research outputs

Molecular Approaches in the Diagnosis of Diarrhoeal Disease in Children from Rural Gambia

Thesis

How to cite:

Ikumapayi, Usman Nurudeen (2019). Molecular Approaches in the Diagnosis of Diarrhoeal Disease in Children from Rural Gambia. PhD thesis. The Open University.

For guidance on citations see FAQs.

© 2019 The Author

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

Molecular approaches in the diagnosis of diarrhoeal disease in children from rural Gambia

Thesis submitted to Open University, United Kingdom in fulfilment of the requirement of the Doctorate of Philosophy in the field of Sciences

2019

Usman Nurudeen Ikumapayi, PhD

Student ID: C2188593

Affiliate Research Centre: Medical Research Council Gambia at The London School of Hygiene and Tropical Medicine

Antlantic Boulevard – Fajara

P.O. Box 273, Banjul

The Gambia, West Africa

Director of Studies: Professor Martin Antonio PhD, FRCP, FRCPath

Antlantic Boulevard – Fajara

P.O. Box 273, Banjul

The Gambia, West Africa

Supervisor 1: Doctor Anna Roca, BSc, MSc, PhD

Atlantic Boulevard – Fajara

P. O. Box 273, Banjul

The Gambia, West Africa

Supervisor 2: Professor James P. Nataro MD, PhD, MBA

Department of Paediatrics

University of Virginia, Children Hospital

United State of America

Abstract

Enteroaggregative *Escherichia coli* (EAEC) is a genetically diverse enteric pathogen that causes growth faltering among children, acute and chronic diarrhoea among children and adults living in both industrialised and low income countries. The German outbreak of EAEC-Shiga-toxin-*E. coli* in 2011 resulted in over 4000 confirmed cases of diarrhoea with over 54 fatalities in 14 European countries as well as United State of America and Canada. Several studies conducted in sub-Sahara Africa (SSA), Latin America and Asia countries have identified EAEC more frequently than any other bacterial pathogens. In SSA, case fatality due to EAEC is not well documented but the morbidity rate particularly among younger children is huge. Studies conducted in Senegal, Central-Africa-Republic and Tanzania showed that EAEC were endemic among HIV-positive patients with diarrhoea. Few studies from SSA have reported distribution of antimicrobial resistance pattern of EAEC.

The Global Enteric Multisite Study (GEMS), a three-year case-control study conducted in seven African and Asia countries, showed that the prevalence of EAEC was higher among children with no diarrhoea (463/741, 62.5%) compared to children with diarrhoea (278/741, 37.5%).

The aim of this retrospective analytical study nested to GEMS is to explore other molecular approaches that identify infectious EAEC and to show the genetic diversity and antimicrobial resistant pattern of EAEC. Study design of the first approach involves unmatched case-control 428 (157 cases and 271 controls) EAEC isolates that were examined by polymerase chain reaction (PCR) for the presence of 21 common EAEC virulence genes. This investigation implicated plasmid-encoded toxin (*pet*), AAF/1 fimbrial subunit (*aggA*) and hexosyltransferase homolog (*capU*) to be associated with diarrhoea in infants. In addition, two other virulence genes; Shigella exracellular protease A (*sepA*) and EAEC-heat-stable enterotoxin 1 (*EAST1*) were implicated in the EAEC that cause diarrhoea among children under 5 years old.

The second approach utilised qualitative PCR (TaqMan-qPCR) method to assess the use of bacterial load diagnostic tool to diagnose infectious EAEC on selected matched case-control 160 (80 cases and 80 controls) EAEC isolates. Two biomarker genes, *aatA* and *aaiC* were the target in this study and both resulted in higher rate of higher bacterial load in controls (58/80 [72.5%]) compared to cases 48/80 [60%]), p – value 0.096.

The third approach explored bacterial biofilm formation to diagnose infectious EAEC on 400 unmatched cases (150) and controls (250) EAEC isolates. Infectious EAEC produces biofilm to consolidate its colonisation in the host and damage to the tissue. The result of this study showed higher proportion of biofilm-producing EAEC in controls (61%) compared to cases (39%). However, biofilm-producing EAEC isolates that has *aggR* gene combined with one or all of the following virulence genes *aatA*, *Aap*, *Orf3* and *Orf61* revealed strong association with diarrhoea

Investigation into the antimicrobial resistant EAEC on the same 400 unmatched EAEC isolates revealed multi-drug resistant (MDR) EAEC infection as a significant problem among infants in the Gambia. MDR EAEC strains are almost equally distributed among cases and controls, and high (>71%) rate of resistant to Ampicillin, Sulphamethoxazole-trimethoprim and Tetracycline, and moderate (25%) rate of resistant to Chloramphenicol among study children. However, over ninety-four percent of the Gambia EAEC strains are susceptible to Cefotaxime, Ceftazidime, Ceftriaxone, Ciprofloxacin, Gentamicin and Amoxicillin-clavulanic acid.

Additionally, result of whole genome sequencing (WGS) on 50 randomly selected EAEC isolates showed average 94% concordance of resistance genes with phenotypic disc diffusion method.

This thesis provides detailed initial description and exploration of virulence genes associated with EAEC strains circulating in the rural Gambia and has revealed the likely biomarker genes to target in the diagnosis of infectious EAEC that cause diarrhoea in infant.

Acknowledgement

I would like to thank God Almighty for His unique guidance and abundance grace before and during the cause of this study.

I would like to extend my profound gratitude to my director of study and mentor Professor Martin Antonio for the support, supervision, consultations and meetings. More importantly I am grateful for the patience demonstrated throughout the study period.

Also, my profound gratitude goes to my second supervisor Dr. Anna Roca for the relentless support, guidance and mentorship.

My sincere thanks also go to my external supervisor Professor James Peter Nataro of University of Virginia US for the generosity in creating a platform for financial and intellectual support prior and during the cause of the study. Thanks for believing in me.

I would also like to express my gratitude to my first third-party monitor Professor Martin Otta, my current third-party monitor Dr. Momodou Jasseh, my study advisor Dr. Davis Nwakanma and my mentor Dr. Peter Dukes for the meetings, guidance and advice.

I am also grateful to my senior colleagues that include Dr. Debashi Saha, Dr. Jahangir Hossain, Dr. Grant Mackenzie, Dr. Ousman Secka, Dr. Brenda Kwambana Adams, Dr. Adedayo Adetifa, Dr. Leopold Tientcheu and Dr. Nadia Boisen of Staten Serum Institute Denmark for the mentorship.

My profound gratitude also goes to Professor Richard Adebayo Adegbola of GSK for the encouragement, advice and trust.

I am also grateful to Miriam Wathuo and Nuredin Ibrahim Mohammed for providing statistical advice.

My gratitude also go to Dr. Abdul Karim Sesay, Abdoulie Kanteh, Jarra Manneh, Archibald Worwui, and Madikay Senghore for the genomic and bioinformatic support and advice.

I am also grateful to my laboratory colleagues that include Dr. Jainaba M. Roussel, Dr. Majidah Adiamoh, Henry Badji, Dr. Chinelo Ebruke, Catherine Okoi, Rasheed Salaudeen, Sarjo Gibba, Modou Lamin, Momodou Drammeh and Muhammed Wally for the technical and intellectual assistance.

I am also grateful to my aged mother for her prayers.

I would like to thank my dear wife and my humble children for their love, cooperation and understanding during the entire study period.

This study is supported and funded by Bill & Melinda Gates foundation through Professor James Nataro of Centre for Vaccine Development Baltimore currently at Paediatric Children Hospital University of Virginia United State of America and Medical Research Council Unit at London School of Hygiene and Tropical Science (MRCG@LSHTM) The Gambia.

Dedication

This study is dedicated to all those who have in any form contributed to the development and sustenance of Biomedical Science Research in low and middle income countries.

Declaration

The *E. coli* isolates and the deoxyribonucleic acid (DNA) used in this study were isolated from stool samples obtained from children recruited into a three year case-control study (Gate Enteric Multisite study – GEMS) conducted from December 2007 – December 2010 at upper river region (URR), The Gambia. Prior sample collection, taking of consent from guardian of study participants were done by well trained nurses.

I performed all of the following experiments and analyses at the MRCG@LSHTM

I. Culture and isolation and identification of *E. coli*

II. Detection of EAEC using PCR amplification method

III. Detection of target 21 virulent related genes using multiplex-PCR amplification method

IV. Measure of Bacterial load using TaqMan-qPCR assay

V. Detection of biofilm production using Test-Tube (TT), Congo-Red Agar (CRA) and Tissue Culture Plate (TCP) methods

VI. Antimicrobial resistance using disk diffusion (Kirby-Bauer) method

VII. Whole genome (WGS) sequencing was performed with the support of MRCG genomic laboratory team using next generation sequencer (Miseq-illumina) and I performed the genome analysis under the guidance of an experience Bioinformatician

VIII. I also performed all the statistical analysis under the supervision and guidance of an experience statistician.

Peer-reviewed publications from this thesis with me as the first author

1. Identification of Subsets of Enteroaggregative *Escherichia coli* Associated with Diarrheal Disease among Under 5-Year Old Children from Rural Gambia. American Journal of Tropical Medicine and Hygiene, 97(4), 2017, pp. 997-1004

Manuscripts from this thesis that are to be submitted to journals for publication with me as the first author;

1. Antibiotic resistance and genetic diversity of Enteroaggregative *Escherichia coli* among diarrhoeal and non-diarrhoeal children from rural Gambia. Journal to submit to: Journal of Antimicrobial Chemotherapy (JAC 5IF) or Emerging Infectious Disease (EID 7/8 IF)

2. A quantitative approach in the diagnosis of infectious Enteroaggregative *Escherichia coli*. Journal to submit to: Diagnostic Microbiology Infectious Disease or JMM (2IF)

Abbreviations

-	Negative
+	Positive
±	Plus-Minus
-ve	Negative
+ve	Positive
≤	Equal to or Less than
≥	Equal to or Greater than
µg	microgram
%	percentage
95%CI	95% Confidence Intervals
AA	Aggregative Adherence
AAF	Aggregative Adherence Fimbriae
aEPEC	Atypical Enteropathogenic *Escherichia coli*
aEAEC	Atypical Enteroaggregative *Escherichia coli*
ADH	Arginine Dehydrolase
AE	Attaching and Effacing
AggR-	AggR absent
AggR+	AggR present
Amp	Ampicillin
AMC	Amoxacillin-clavulanic acid
API	Analytical profile index
ATP	AdenosineTriphosphate
bp	Base-pair
BF	Biofilm formation
BF-	Biofilm not form
BF+	Biofilm formed
BFP	Bundle-Forming Pilus
BGS	Buffer Glycerol Saline
BHI	Brain Heart Infusion
cAMP	Cyclic Adenosine MonoPhosphate
cGMP	Cyclic Guanosine MonoPhosphate
CART	Classification And Regression Tree
CAZ	Ceftazidime
CB	Carry Blaire
CD	Cell Dendritic
CDS	Coding Sequence
CDEC	Cell Detaching *Escherichia coli*
CFs	Colonisation Factors
CLSI	Clinical laboratory Standard Institute
Cq	Quantification cycle
CRA	Congo Red Agar
CRO	Ceftriaxone
CRR	Central River Region
C_t	Cycle threshold
CVD	Centre for Vaccine Development
CVD432	Gene code *aatA* gene
DAEC	Diffusely Adherent *Escherichia coli*

DAF	Decay-Accelerating Factor
DEC	Diarrhogenic *Escherichia coli*
DNA	Deoxyribonucleic Acid
EAF	EPEC-Adherence Factor
EAEC	Enteroaggregative *Escherichia coli*
ECP	Escherichia coli common pillus
EDTA	Ethylenediaminetetraacetic acid
EHEC	Enterohaemorrhagic *Escherichia coli*
EIEC	Enteroinvassive *Escherichia coli*
EPEC	Enteropathogenic *Escherichi coli*
EPI	Expanded Programme immunisation
ELISA	Enzyme-linked immunosorbent assay
EQA	Externa quality assurance
ESBL	Extended spectrum beta-lactamase
ETEC	Enterotoxigenic *Escherichia coli*
FOX	Cefoxitin
Gb3s	Globotriaosylceramides
GC	Guanylate cyclase
GCP	Good Clinical Practice
GCLP	Good Clinical Laboratory Practice
GEMS	Global Enteric Multisite Study
GM-CSF	Granulocyte Macrophage Colony-Stimulating Factor
H_2O	Water
H_2S	Hydrogen Sulphide
HBL	High Bacteria Load
HC	Haemorragic colitis
HCL	Hydrochloric acid
HCP	Haemorrhagic coli pilus
HEp2	Human Laryngeal Epithelial Cell Line Type 2
HIV	Human immunodeficiency virus
HUS	Haemolytic Uremic Syndrome
ICAM	Inter-cellular adhesion molecule
IL	Inter-leukin
INF	Interferon
iTOL	Interactive tree of life
Km	Kilometre
kDa	Kilo-Dalton
LA	Localised adhesion
LBL	Low bacteria load
LDC	Lysin decarboxylase
LEE	Locus of enterocyte effacement
LoD	Limit of detection
LRR	Lower River Region
LT	Heat-labile Toxin
MAPK	Mitogen-activated protein kinases
MDR	Multidrug resistant
MLST	Multi-Locus Sequence Typing
MM	Master-mix
MRCG	Medical Research Council The Gambia

MSD	Moderate-to-severe diarrhoea
MUAC	Mid upper arm circumference
MVT	Meningococcal vaccine trial
NBR	North Bank Region
Non-MSD	Non-Moderate-to-severe diarrhoea
nm	Nano-meter
ODC	Ornithine Decarboxylase
OR	Odd Ratio
ORT	Oral rehydrationTherapy
ORS	Oral Rehydration Solution
PAMPs	Pathogen associated molecular patterns
PBS	Phosphate buffer saline
PCR	Polymerase Chain Reaction
PERCH	Pneumococcal Etiology Research for child health
PI	Pathogenicity index
PMN	Polymophonuclear leukocyte
PRRs	Pattern recognition receptors
PVT	Pneumococcal vaccine trial
QPCR	Quantitative Polymerase Chain Reaction
SCC	Scientific Co-ordinating committee
SHC	Sentinel health centre
SPATE	Serine protease autotransporter of Enterobacteriacea
SNP	Single nucleotide polymorphism
Stx	Shiga-toxin
ST	Heat stable toxin
SXT	Cotrimoxazole
TCBS	Thiosulphate citrate bile-salt sucrose
TCP	Tissue culture plate
TD	Traveller’s diarrhoea
TBE	Tris-Borate-EDTA
TET	Tetracycline
TLRs	Toll-like receptors
TNF	Tumour necrosis factor
TSB	Tryptone soya broth
TT	Tube test
UK	United Kingdom
URR	Upper River Region
USA	United state of America
UTI	Urinary Tract Infection
UV	Ultra-violet
VFDB	Virulence Factor Data Base
VFS	Virulent factor score
VIDA	Vaccine Impact on diarrhoea in Africa
WBR	West Bank Region
WGS	Whole Genome Sequencing
WHO	World Health Organisation

List of tables

Table 1.1: Distribution of Diarhoegenic *Escherichia coli* pathotypes from Diarrhoeal and Non-Diarrhoeal Children enrolled in GEMS Gambia site 87

Table 1.2: Distribution of Diarhoegenic *Escherichia coli* pathotypes from Diarrhoeal and Non-Diarrhoeal Children enrolled in GEMS Gambia site 88

Table 3.1: Primer sequence for the detection of 21 EAEC Virulence associated genes...........105

Table 3.2: OligoNucleotide sequence of primers and probe used 112

Table 3.3: Concentration of DNA extracts used for whole genome sequencing 127

Table 3.4: Bioinformatic Analysis performed and detail of tools used 137

Table 4.1: Baseline information of study population 144

Table 4.2: Distribution of EAEC virulence genes from case and control children (age 0-59 month) 144

Table 4.3: Distribution of EAEC virulence genes in case and control children in three age strata 147

Table 4.4: Characterization of virulence factor in the EAEC pathotype obtained as sole pathogen from case and control children age 0-11 month 148

Table 5.1: Univariable analysis showing effect of each confounder on the Association between being a case/control and High bacterial load 162

Table 5.2: Multivariable analysis showing effect of all the confounders on the Association between being a case/control and High bacterial load 163

Table 5.3: Analysis showing virulence/pathogenicity index of each virulence factor on high and low bacterial load among MSD and non-MSD children age less than 5 years old (n=93) 165

Table 5.4: Virulent factor score against Bacterial load and other variables among MSD children (n=48) 167

Table 5.5: Virulent factor score against Bacterial load and other variables among Asymptomatic children (n=45) 168

Table 6.1: Evaluating the use of CRA for biofilm screening tests using TCP as gold standard 184

Table 6.2: Evaluating the use of TT for biofilm screening tests using TCP as gold standard.. 184

Table 6.3: Categigorising biofilm producing EAEC base on intensity of biofilm produced....186

Table 6.4: Distribution of virulence genes in Biofilm producing and Non-Biofilm Producing EAEC strains among Cases and Controls.. 187

Table 6.5: Association of virulence factor and the formation of biofilm (BF +) and the non-formation of biofilm (BF -) **among EAEC isolates** from diarrhoeal and non-diarrhoeal children (n=400) 188

Table 6.6: Association of EAEC biofilm producer with virulence genes having AggR in the background 189

Table 6.7: Association of Antimicrobial resistance and the biofilm formation (BF +) and the non-biofilm formation (BF-) from diarrhoeal and non-diarrhoeal children (n=400) 190

Table 7.1a: Comparing antibiotic sensitivity results of disk diffusion and MIC E-test methods.......... 203

Table 7.1: Antimicrobial resistance patterns of EAEC strains from diarrhoeal and non-diarrhoeal children 204

Table 7.2: Antibiotic Resistant Pattern of EAEC strains from two age strata diarrhoeal and non-diarrhoeal children 206

Table 7.3: Distribution of MDR EAEC strains among three age strata of study children 207

Table 7.4: Association between Multidrug resistant EAEC strains and the corresponding virulence factors 210

Table 7.5: Association of Animals as risk factors for Multidrug resistant EAEC strains from children 212

Table 7.6: Multidrug Resistance Patterns among study participant's recruitment zones (villages)........ 212

Table 8.1: Summary of the genome content of the 50 Isolates Examined 228

Table 8.2: Distribution of Antimicrobial resistance genes among diarrhoeal and non-diarrhoeal children 230

Table 8.3: Comparison between whole genome sequencing and phenotypic (disk diffusion) based predictive Antimicrobial Susceptibility Test (AST) 231

Table 8.4: Multilocus Sequencing Typing that results in STs and ST-complex 232

Table 8.5: Distribution of ST among cases ad controls 234

Table 8.6: Distribution of Resistance genes against sequence types (STs) 240

Table 8.7: Distribution of plasmid Associated Resistance Gene in EAEC isolates among cases and controls 242

List of figures

Figure 1.1: Map of The Gambia 30

Figure 1.2: Interaction of Diarrhoea and Malnutrition 38

Figure 1.3: Pathogenic mechanisms of enteropathogenic and enterohaemorrhagic *Escherichia coli* 45

Figure 1.4: Pathogenic mechanisms of enterotoxigenic *E. coli*, enteroagreggative *E. coli* and diffusely adherent *E. coli* 47

Figure 1.5: Stages of pathogenesis of EAEC Numbers in circles show the progression of EAEC pathogenesis 66

Figure 1.6: Showing countries with EAEC outbreaks, case series and cohort studies 78

Figure 1.7: Flow chart showing detail of death in GEMS study 88

Figure 3.1: Flow chart of study samples 95

Figure 3.2: Showing positive and negative indole tests 98

Figure 3.3: PCR reaction cycle pattern used in the amplification of the 21 virulence genes ...106

Figure 3.4: Gel image of group 1 (*sat, sepA, pic, sigA*, *pet* and *astA*) of the 21 putative virulence genes 107

Figure 3.6 Showing quantification progression of gene *aatA* (HEX) 114

Figure 3.7 Showing quantification progression of gene *aaiC* (FAM) 114

Figure 3.7a: Showing highest concentration of standad at Cq 16 and at log starting quantity 6.6. Similarly, lowest concentration at Cq 38 and at log start quantity 1.6 of gene *aatA* (HEX)..... 114

Figure 3.8: Test tube method for screening biofilm – Tube A, B, C and D serves as positive controls whilst tube E and F serves as negative controls 117

Figure 3.9: Congo red Method for screening biofilm – Dark black colonies B indicate biofilm formation whilst pink-red colonies A indicate absence of biofilm 118

Figure 3.10: Tissue culture plate screening method for the presence or absence of biofilm – Well A1 and well A2 are Biofilm positive and negative respectively 120

Figure 3.11: FastQC – Quality Score per read base 133

Figure 3.12: Distribution of Quality score across all sequences 134

Figure 3.13: Whole genome sequencing work flow 136

Figure 3.14: Dichotomous decision tree to determine the phylogenetic group of an EAEC isolate by the presence or absence of *chuA*, *yjaA* and *TspE.C2* 138

Figure 4.1: Association of virulence factors with diarrhoea in children aged 0-59 months 149

Figure 4.2: Association of virulence factors with diarrhoea in children aged 0-11 months 150

Figure 5.1: proportional distribution of stool consistency in EAEC sole infection among MSD with high and low bacterial load 164

Figure 5.2: proportional association of virulent factor score and bacterial load among MSD and non-MSD children 169

Figure 6.1: Schematic representation of Biofilm life-cycle 181

Figure 6.2 (A, B, C and D): Proportional distribution of Biofilm in EAEC strains from cases and controls 185

Figure 6.2a: Bimodal distribution of the mean Absorbance 186

Figure 6.3 (A & B): Comparison of Antibiotic resistance pattern of **biofilm producing- EAEC** between cases and controls (n=219) 191

Figure 6.4 (A & B): Comparison of Antibiotic resistance pattern of **non-biofilm producing EAEC** between cases and controls (n=181) 192

Figure 7.1: Distribution of Multi-Drug-Resistant EAEC strains from case and control 204

Figure 7.2: Antibiotic Multi-resistance patterns of EAEC strains from MSD and non-MSD children 205

Figure 7.3: Proportional Distribution of MDR EAEC strains among different age strata of study children (MSD and non-MSD combined) 207

Figure 7.4: Proportional distribution of MDR-EAEC strains among age strata MSD children only 208

Figure 7.5: Distribution of MDR EAEC among study children 208

Figure 7.6: Distribution of MDR EAEC among MSD children 209

Figure 7.7: Distribution of MDR EAEC among non-MSD children 209

Figure 7.8: Association of class II (pic & sepA) SPATE proteins and other specific gene combination with multi-drug-resistant EAEC strains 211

Figure 8.1: Sketch sequencing approaches for the diagnosis and monitoring of infectious diseases.. 223

Figure 8.2: Core SNP phylogeny showing the evolutionary relationship of taxa................... 235

Figure 8.3: Pairwise core SNP distance histogram showing the distribution of SNPs according to their proximity to the nearest gene 236

Figure 8.4: Core SNPs density revealed 471951 SNPs across the approximate 5.2 million base pair 237

Figure 8.5: Pan Genome heat mapping showing the core genes (common to all isolates) and the accessory genes that are shared between isolates, determined by a pairwise comparison....... 238

Figure 8.6: Molecular phylogenetic analysis by Maximum Likelihood method 239

Figure 8.7a: Phylogenetic tree showing relatedness of EAEC strains 243

Figure 8.7b: Phylogenetic tree showing relatedness of EAEC strains in circular format..........244

Table of Content

Chapter 1: Literature Review 27

1.1 Introduction..27

1.2 The Gambia and Upper River Region Gambia....................................29

1.3 An overview of diarrhoeal disease..32

1.3.1 Definition of diarrhoea 33

1.3.2 Clinical syndrome of diarrhoea 33

1.3.2.1 Acute watery diarrhoea..33

1.3.2.2 Dsentery diarrhoea...34

1.3.2.3 Persistent diarrhoea...34

1.3.3 Transmission and Spread of Diarrhoea 34

1.3.4 Pathogenesis of Microbial agent of Diarrhoea 35

1.3.5 Risk factors...35

1.3.5.1 Behavioural risk factors 35

1.3.5.2 Host factors 35

1.3.5.3 Age factors...36

1.3.5.4 Seasonality 36

1.3.6 Interrelationship between nutritional state 37

1.3.7 Management of acute-diarrhoea...39

1.3.8 Timely recognition 40

1.4 Diarrhoeagenic *Escherichia coli* (DEC) strains at a glance 42

1.5 Review of Enteoaggregative *Escherichia coli* (EAEC) 49
1.5.1 Discovery 49
1.5.2 Classification 51
1.5.3 Epidemiology 53
1.5.3.1 Developing Countries 53
1.5.3.2 Developed Countries 54
1.5.4 Transmission and Reservoir 55
1.5.5 Pathogenesis 58
1.5.5.1 Brief explanation 58
1.5.5.2 Adherence and Colonisation 62
1.5.5.3 Increase Mucus or Biofilm production 65
1.5.5.4 Release of Toxins 67
1.5.5.5 Immunologic / Inflammatory response 70
1.5.5.6 Host genetic / susceptibility factor 71
1.5.5.7 Pathogenicity Islands 72
1.5.5.8 Malnutrition 73
1.5.6 Clinical Manifestation (Symptom) 74
1.5.6.1 Brief description 74
1.5.6.2 Shigatoxin producing EAEC strain 75
1.5.6.3 Outbreak of EAEC 77
1.5.7 Laboratory Diagnosis 79

1.5.7.1 Aggregation adherence assay......................................79

1.5.7.2 Polymerase Chain Reaction (PCR) assay 80

1.5.7.3 Biofilm assay 80

1.5.7.4 Serologic-Serotyping assay..81

1.5.7.5 Newly proposed assay..81

1.5.7.6 Genome sequencing 82

1.5.8 Treatment and Drug Resistance 83

1.5.8.1 Brief explanation...83

1.5.8.2 Complication of Antibiotics in the Treatment of infections caused by Stx-containing EAEC strains 84

1.5.8. 3 Vaccine and other preventive therapy.........................85

1.6 Summary of Gambian site specific GEMS study in which this study is nested 86

Chapter 2: Research question Hypotheses, Aims and Objectives89

2.1 Research question 90

2.1 Hypothesis 90

2.2 Aims...91

2.3 Specific objectives 91

Chapter 3: Materials and Methods 92

3.1 Laboratory settings at MRCG Basse field station and Fajara................92

3.2 Study population:..93

3.3 Case Definition of Moderate-to-Severe Diarrhoea (MSD)...................93

3.4 Definition of Control 93

3.5 Study approval 94

3.6 Good Clinical Laboratory Practice (GCLP) 94

3.7 Sample selection 94

3.8 Detection of EAEC 96

3.9 Bacteriology 96

3.9.1 Gram stain reaction 97

3.9.2 Indole production: 98

3.9.3 Biochemical identification kit (API 20E) 99

3.10 Isolation of Genomic DNA 102

3.11 Molecular detection of EAEC single targeting *aatA* and *aaiC* 102

3.12 PCR amplification to detect 21 EAEC virulence genes 103

3.13 Statistical analyses 108

3.14 Significance of Combinations of EAEC Genes: 108

3.15 Quantisation of Bacterial load for EAEC 109

3.15.1 DNA extraction from cultured EAEC 042 strain 109

3.15.2 Total DNA extraction directly from stool samples 109

3.15.3 QPCR amplification for the 160 total DNA extractions 111

3.15.4 Amplification efficiency and limit of detection (LoD) or Quantitation of bacterial load 113

3.15.5 Statistical Analysis 115

3.16 Biofilm production in EAEC strains 116

3.16.1 Strains 116

3.16.2 Tube screening method 116

3.16.3 Congo red agar screening method 118

3.16.4 Tissue culture plate screening method 119

3.17 Antibiotic resistance Assay..121

3.17.1 Summary 121

3.17.2 Antibiotic resistance of using disk diffusion method...........121

3.18 Whole Genome Sequencing of EAEC strains ....,..............124

3.18.1 Isolation of genomic DNA from EAEC isolates 124

3.18.2 DNA Quality control 127

3.18.3 Library preparation (frgmentation)..129

3.18.4 Size selection of Adaptor - Ligated DNA 131

3.18.5 Library analysis 132

3.18.6 Template preparation and automated sequencing 132

3.18.7 Raw data quality control 133

3.18.8 Sequencing analysis 135

3.18.9 Phylogenetic analysis 135

Chapter 4: Prevalence of Virulence genes among EAEC strains from MSD and non-MSD children..139

4.1 Introduction..139

4.2 Result 143

4.3 Discussion 151

4.4 Conclusion 157

Chapter 5: A quantitative assessment of clinical significance of EAEC in diarrhoea using TaqMan-QPCR 159

5.1 Introduction 159

5.2 Result 161

5.3 Discussion 170

5.4 Conclusion 176

Chapter 6: Biofilm production in EAEC strains from diarrhoeal and non-diarrhoeal children 178

6.1 Introduction 178

6.2 Result 182

6.3 Discussion 193

6.4 Conclusion 197

Chapter 7: Antibiotic resistance patterns of EAEC strains from diarrhoeal and non-diarrhoeal children from rural Gambia.....198

7.1 Introduction 198

7.2 Result 200

7.3 Discussion 213

7.4 Conclusion 217

Chapter 8: Whole Genome Sequencing of Selected EAEC strains from case and control study children..219

8.1 Introduction 219

8.2 Result 224

8.3 Discussion 245

8.4 Conclusion 249

Chapter 9: Discussion of Hypotheses, Aims and Objectives 250

Chapter 10: Concluding remarks 253

10.1 Limitations of the study 253

10.2 Future studies 255

10.3 Closing remarks 257

Chapter 11: Reference 258

Chapter 12: Appendix...274

Appendix A – Study approvals 274

Appendix B – GCP/GCLP Certificate.....................................276

Appendix C – Creating standard curves with genomic DNA......277

Appendix D – Media preparations..280

Appendix E – Peer-reviewed journal from the thesis 282

Chapter 1: Literature Review

1.1 Introduction

This study was **nested to** a three year case-control diarrhoea project tagged Global Enteric Multisite Study (GEMS) jointly sponsored by Bill and Melinda Gates foundation of USA and Medical Research Council Unit, The Gambia, West Africa. The samples used were obtained in a population of children with moderate-to-severe diarrhoea (MSD) cases and non-moderate-to-severe diarrhoea (non-MSD) controls and were 0-59 months of age. Case definition of MSD is a child with diarrhoea (≥3 abnormal loose stools) within the previous 24 hours with onset within the previous 7 days, following at least 7 days without diarrhoea, and accompanied by evidence of clinically significant dehydration (loss of skin turgor, sunken eyes, or a decision by the clinician to administer intravenous fluids), dysentery (blood in the stool), or clinical decision to hospitalize the child (Farag, Nasrin et al. 2012).

The pathogenesis of Enteroaggregative *Escherichia coli* (EAEC) diarrhoea comprises colonisation of the intestinal mucosa, elaboration of enterotoxins and release of proinflamatory cytokines from the infected epithelial cells. Characteristically, EAEC strains enhance mucus secretion from the mucosa, with trapping of the bacterium in a bacterium-mucus biofilm. The clinical presentation of EAEC is characterized by watery, mucoid, secretory diarrhoea with low-grade fever and occasional vomiting.

This retrospective analytical study aimed at providing the first case-control data to evaluate role of Enteroaggregative *E. coli* (EAEC) in diarrhoeal disease and assessment

of molecular approaches to the diagnosis of diarrhoea caused by EAEC in children in The Gambia.

The benefit of this study will influence policy regarding treatment and recommending an uncomplicated, affordable and accessible diagnostic tool that decreases the prevalence of diarrhoea caused by EAEC among children in the Gambia.

1.2 The Gambia and Upper River Region Gambia

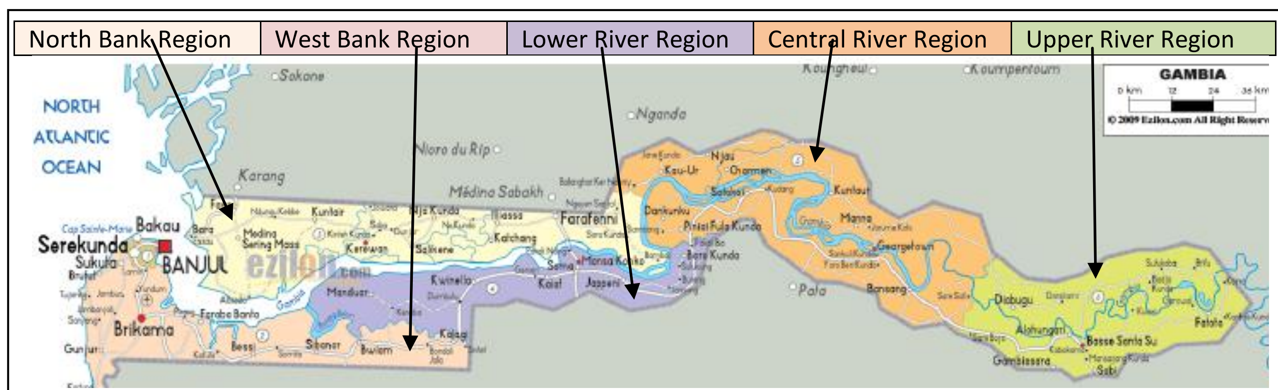
The Gambia is a country in West-Africa that is mostly surrounded by Senegal. It is the smallest country on mainland Africa that has become a centre for tropical medical research, largely due to the over 70 years presence of Medical Research Council Unit The Gambia at London School of Hygiene and Tropical Medicine (MRCG@LSHTM) providing required scientific and clinical facilities to The Gambia and to the west-African regions.

The Gambia is situated 12° north of the equator. The country is bordered to the north, east and south by Senegal. The western side of the country borders the North Atlantic Ocean with 50 miles of the coastline. The country is narrow and its border mirrors the meandering Gambia River which empties into the Atlantic Ocean. The country is less than 30 miles wide with a total area of 11,300 km^2, 1,300 km^2 of Gambia's area is covered by water. Its size is about 10, 500km^2. Currently, the country is divided into five administrative regions that include West Coast, North Bank, Lower River, Central River (CRR) and Upper River regions (URR) with two municipal areas which are Banjul and Kanifing (figure 1.1). The MRC Unit The Gambia operates in all the regions but has field stations in two regions (West Coast – Keneba and Basse-URR) with headquarter at Fajara in Kanifing municipality.

The climate of The Gambia is tropical. From June to October is the rainy season period with intermittent warm weather. The annual average rainfall is 800-1200 mm. From November to May there are cool temperatures and is part of a dry season. The temperature during hot season (February to May) at CRR and URR can go as high as 47°C. The commonest fruit in the Gambia is Mango of various species.

The Gambia population is estimated at 2 million, and about 90% of the population are Muslims and 63 % live in rural villages. The official language is English however; the country has a wide variety of ethnic groups, each preserving its own language and tradition. The Mandinka tribe is largest, followed by the Fula, Wolof, Jola, Sarahule, Serers, Karoninka, Manjago and Bianukas [https://en.wikipedia.org/wiki/The_Gambia]. The Krio known as Aku and the Hausa are the smallest ethnic minorities in The Gambia. About 1% of the Gambia population are non-African descendants that include Europeans and Lebanese origin (Gambia 2015) [Central Statistics Department, The Gambia 2015]

Figure 1.1: Map of The Gambia



(Maps 2015)https://www.ezilon.com/maps/africa/gambia-maps.html

The economy of the Gambia is dominated by farming, fishing and largely tourism. About a third of the population lives below the international poverty line of US$1.25 a day (Gambia 2016) https://en.wikipedia.org/wiki/Economy_of_the_Gambia

Currently, Gambia life expectancy for females is 57 years and for males is 54 years with mortality rate of 39.1 per 1000 for children less than 5 years of age [https://data.unicef.org/topic/child-survival/under-five-mortality/].

The report in 2010 showed maternal mortality rate per 100,000 births as 400, and the under-5 mortality rate per 1000 births, was 106 and the neonatal mortality, as a percentage of under-5 mortality, is 31. It was estimated that the number of midwives per 1,000 live births is five and the lifetime risk of death for pregnant women is one in 49 [https://en.wikipedia.org/wiki/The_Gambia].

In October 2010, a report showed that Gambia had made significant improvements in polio, measles immunisation and the PCV-7 vaccine (Roca, Hill et al. 2011; Scott, Odutola et al. 2014). Additionally, in August 2013, a nationwide coverage rotavirus vaccine was introduced (Unicef 2013)[https://www.unicef.org/gambia/media_8418.html]. A regional representative of the WHO (Thomas Sukwa) commended Gambia "The Gambia Expanded Programme on Immunization (EPI) program is one of the best in the World Health Organisation African Region" [https://en.wikipedia.org/wiki/The_Gambia]

1.3 An overview of diarrhoeal disease

Infectious diarrhoea is one of the principal causes of morbidity and mortality particularly in young children globally. Diarrhoeal illnesses account for 22% of the 10 million annual worldwide deaths of children under 5 years of age (Black, Morris et al. 2003). More than half of these cases occur in Africa and South Asia. In emergency conditions, 90% of deaths are due to diarrhoea (Toole and Waldman 1997). A study conducted in 2000 estimates that diarrhoea accounts for only 13% of all childhood deaths, amounting to 1.4 million deaths per year (Murray 2001). Since then the incidence of diarrhoea reported varies hugely with the seasons and a child's age, the incidence is high in the first two years of life and declines as a child grows older. In 2004, a study recorded a decline in mortality caused by diarrhoea among children under five years from an estimated 5 million deaths to 1.5 million deaths over the last 20 years (Pinto 2009). Despite this decline, WHO recorded in 2017 maintains diarrhoea was the second most common cause of death among children under five years old globally and the disease kills about 525000 children under five (WHO 2017) (http://www.who.int/mediacentre/factsheets/fs330/en/). Globally, there are about 1.7 billion cases of childhood diarrhoeal illness annually and it is regarded as the leading cause of malnutrition in children under five years old hence, risk of life-threatening diarrhoea is mostly associated to this age group (WHO 2017) (http://www.who.int/mediacentre/factsheets/fs330/en/).

1.3.1 Definition of diarrhoea

Diarrhoea is defined in epidemiological studies as the passage of three or more loose or watery stools in 24-hour period. Infectious diarrhoea is caused by a bacterial, viral or parasitic infection mostly in the small intestine. Under normal conditions, there is a balance in absorptive and secretory functions of intestinal water and electrolytes. The inner surface of the small intestine lined with specialised cells known as enterocytes are responsible for secretion and adsorption during the process of digestion (Cuting 1998). When the intestine fails to absorb water or when there is an increase in fluid secreted into the intestine, the amount of water in stool increases and the individual has diarrhoea

1.3.2 Clinical syndrome of diarrhoea

The three known clinical syndromes of diarrhoea are acute watery diarrhoea, dysentery and persistent diarrhoea. Each of these reflects a different pathogenesis and requiring different approaches to treatment.

1.3.2.1 **Acute watery diarrhoea** involves the passage of frequent loose or watery stools without visible blood. This symptom can last up to 14 days but most episodes often last less than seven days. Vomiting may occur and fever may be present. The severity loss of water and salt from the body leads to dehydration and ultimately results into death due to breakdown in vital organs function. Several previous studies have shown that about 60% of stool samples from children with acute diarrhoea yielded enteric pathogens, which include rotavirus and diarrhoeagenic-*E. coli* as the most common (Kang, Ramakrishna et al. 2001; Reither, Ignatius et al. 2007). Others are *Campylobacter* species, *Shigella* species, *Salmonella* species, *Vibrio cholera*, Norovirus and Cryptosporidium species (Kotloff, Nataro et al. 2013). All these play an important role in many different geographic areas.

1.3.2.2 **Dysentery diarrhoea** involves visible blood in the faeces. The effect includes rapid anorexia, weight loss, and damage to the intestinal mucosa by the invasive pathogen such as *Shigella*, Enteroinvasive *E. coli*, *C. jejuni* and occasionally salmonella (WHO-document 1992).

1.3.2.3 **Persistent diarrhoea** is of unusually long duration, at least 14 days. The episode may begin either as watery diarrhoea or as dysentery. Diarrhoeal stool volume may be huge, with risk of dehydration and frequency of severe weight loss. There is no specific microbial cause for persistent diarrhoea. However, Enteroaggregative *E. coli* (EAEC) and Cryptosporidium were fund to play a greater role compared to other diarrhoeagenic agents (WHO-document 1992) (Jensen HB 2016; Molloy FS 2010). It should be noted that persistent diarrhoea must not be confused with chronic diarrhoea, which refers to recurrent or long-lasting diarrhoea due to non-infectious causes, such as sensitivity to gluten or inherited metabolic disorders.

1.3.3 Transmission and Spread of Diarrhoea

The infectious agents that cause diarrhoea are usually transmitted by the faecal-oral route which includes ingestion of faecal contaminated water or food, person-to-person transmission, and direct contact with infected faeces (Baker 2011). A number of behaviours that help spread diarrhoeal pathogens include preparing food with hands that have been contaminated during defecation and not properly washed, allowing an infant to crawl, or a child to play in an area where human or animal faeces are present (Baker 2011).

1.3.4 Pathogenesis of Microbial agents of Diarrhoea

Infection by a diarrhoeal causing pathogen following ingestion is by adhesion and colonisation to a receptor on the enterocyte cells of the small intestine. Secretory diarrhoea therefore occurs when pathogens stimulates infected enterocytes to secrete water and salts into the small intestines. If the cells are hugely affected, the intestine cannot reabsorb all of the fluid secreted and the infected individual develops diarrhoea (Cuting 1998). Invasive diarrhoea occurs when pathogens kills enterocytes, reducing the surface area available for adsorption of water and in turn reducing the ability of the intestine to absorb water and causing diarrhoea (Cuting 1998). However, all of these are induced by different mechanisms of action of different enteric pathogens.

1.3.5 Risk factors

There are four major risk factors that make younger children highly susceptible to diarrhoea. These are behavioural, host immune response, age and seasonality factors.

1.3.5.1 **Behavioural risk factors** – These include failure of the mother to breast-feed exclusively for the first 4-6 months of life or failure to continue breast-feeding for at least 1 year, using infant feeding bottles that easily become contaminated with faecal bacteria, storing cooked food at room temperature, drinking water that is contaminated with faecal bacteria, failing to wash hands before handling food and failing to dispose of faeces hygienically (Barrell and Rowland 1979; Feachem 1984).

1.3.5.2 **Host Immune factors** – These includes undernutrition, recent measles infection and immunosuppression. In undernourished children there is increase in frequency, severity, duration and risk of death from diarrhoea. Also, diarrhoea and dysentery tend to occur more frequently and or severely in children with measles and those children who had measles in

the previous four weeks. This speculated to have resulted from immunological impairment caused by measles (WHO 1992) (http://rehydrate.org/diarrhoea/tmsdd/1med.htm). Diarrhoea due to immunosuppression resulting from viral infection can be brief or can be longer in individuals infected with acquired immunodeficiency syndrome (AIDS) and in this condition diarrhoea can also be caused by unusual microbial agents.

1.3.5.3 **Age factor** – In the case of age factor, most diarrhoeal episodes occur during the first two years of life. Incidence is highest in the group 6-11 months, when weaning often occurs. This pattern reflects the combined effects of declining levels of maternally-acquired antibodies, the lack of active immunity in the infant, the introduction of food that may be contaminated with faecal bacteria and direct contact with human or animal faeces when the infant starts to crawl. Most enteric pathogens stimulate at least partial immunity against repeated infection or illness, which helps to explain the declining incidence of disease in older children and adults (WHO 1992) (http://rehydrate.org/diarrhoea/tmsdd/1med.htm).

1.3.5.4 **Seasonality** - In the case of seasonality factor, distinct seasonal patterns of diarrhoeal occur in many geographical locations. For example, in temperate climates bacterial diarrhoeas tend to occur more frequently during the warm season, whereas viral diarrhoeas, particularly disease caused by rotavirus, peak during the winter. In tropical areas, rotavirus diarrhoeas tends to occur throughout the year, increasing in frequency during the drier, cool months, whereas bacterial diarrhoeas tend to peak during the warmer, rainy season. The incidence of persistent diarrhoea follows the same seasonal pattern as that of acute watery diarrhoea(WHO 1992) (http://rehydrate.org/diarrhoea/tmsdd/1med.htm).

1.3.6 Interrelationship between diarrhoea and malnutrition

The intimate relationship between diarrhoeal illness and undernutrition among children from developing countries is well documented (Lima and Guerrant 1992). Diarrhoeal injury to the gut can impede children growth which can results in extreme forms of chronic malnutrition that predispose young children to diarrhoeal related mortality like in the case of moderate and severe stunting (Guerrant, Oria et al. 2008; Roche, Cabel et al. 2010). Diarrhoea is a recognised cause of malnutrition because demands for nutrients are high during diarrhoea, as during other infectious diseases, consequently, nutrient intake and absorption are often declined. Each episode of diarrhoea often lead to weight loss and growth faltering, and if occurs frequently, there may be too little to "catch up" on growth which means unable to make up for the growth that failed to occur between episodes. Children who experience frequent episodes of acute diarrhoea, or have persistent diarrhoea, are more likely to become malnourished than children who experience fewer or shorter episodes of diarrhoea. In summary, the impact of diarrhoea on nutritional status is proportional to the number of days a child spends with diarrhoea each year. Malnutrition also contributes to the problem of diarrhoea. Malnourished children due to inadequate feeding, severe and more frequent acute and persistent diarrhoeal episodes as well as frequent dysentery stand high risk of dying from this complication (WHO-document 1992). Hence, diarrhoea and malnutrition combine to form a vicious circle (figure 1.2) which, if it is not broken may eventually lead to death (WHO-document 1992).

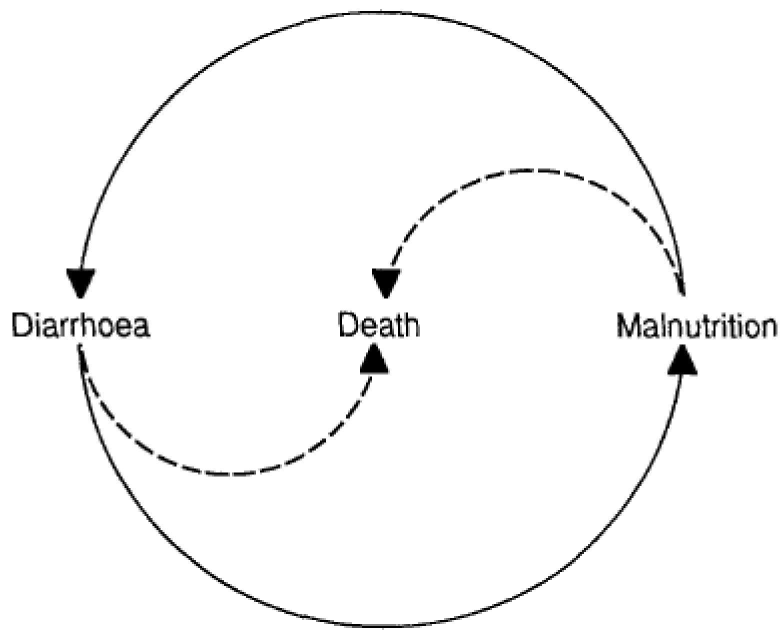


Figure 1.2: Interaction of Diarrhoea and Malnutrition

(Adapted from WHO document of library cataloguing data 1992 of student manual) WHO 91816

1.3.7 Management of acute-diarrhoea

The most common risk with acute diarrhoeal illnesses particularly in children is dehydration and loss of electrolytes and in the developing countries malnutrition has been implicated. So the first step in managing acute diarrhoea is to correct dehydration and electrolyte imbalance (Warren 1983) which can be accomplished with an oral glucose or starch-containing electrolyte solution in the vast majority of cases. Oral Rehydration Therapy (ORT), one of the greatest medical innovations of the 20th century (Santosham, Keenan et al. 1997) has proven to be highly effective. Oral Rehydration Solution was formulated base on the observation that glucose-sodium co-transport was unaffected in cholera and that the recognition of secretory and absorptive processes in the intestine are quite separate. ORT solutions contain specific concentration of sodium, glucose, potassium, chloride and alkali (bicarbonate or citrate) in clean water (Guerrant, Van Gilder et al. 2001). These constituents of ORTs help to restore the electrolyte balance and hydration.

In some cases treatment with an antibiotic may be necessary (Estrada-Garcia, Perez-Martinez et al. 2014). However, antimicrobial therapy must be carefully weighed before commencing treatment. This is due to the fact, that the harmful consequences of non-prudent use of antibiotic increase risk of complications and conditions such as antimicrobial resistant-enteric pathogens, side effects of antibiotic treatment, superinfections when normal flora are eradicated by antibiotics and likelihood induction of disease-producing phage by antibiotics (such as Shiga-toxin phage induced by quinolone antibiotics) which might outweighed the benefit of the antimicrobial therapy (Nguyen 2005).

1.3.8 Timely recognition

Early recognition and treatment of infectious diarrhoea is paramount. So the detection of clinical symptoms, particularly the danger signs such as vomiting, convulsions, lethargy or unconsciousness, lost of appetite and blood in stool are very important. A thorough history must include both clinical and epidemiological features in evaluating patient who present with diarrhoeal illnesses (Nguyen 2005). The clinical information includes the duration and frequency of diarrhoea, stool characteristics (appearance, colour, watery, bloody, mucous, purulent and formed or unformed), quantity of stool produced, presence of dysenteric symptoms (fever, tenesmus, blood and/or pus in stool), symptom of volume deletion (thirst, tachycardia, decreased urination, lethargy, decreased skin turgor), and associated symptoms and their frequency and intensity (nausea, vomiting, abdominal pain, cramp, headache, myalgias, altered sensorium).

The epidemiological information that is very useful to public health professionals in decision making includes identification and sub-typing of the causative agent, prompt notification of pathogens-specific diagnoses and other epidemiological factors that can reduce the rate of transmission and lead to timely detection and control of diarrhoea outbreaks.

Investigation: In this thesis, "**Molecular approaches in the diagnosis of diarrhoeal disease in children from a developing country" (Gambia),** we investigate the role of Enteropatogenic *E. coli* (EPEC), Enterotoxigenic *E. coli* (ETEC) and Enteroaggregative *E. coli* (EAEC) in diarrhoea by analysing the GEMS Gambia site specific data that showed high prevalence of EAEC among MSD and non-MSD children but with no association with diarrhoea. These three diarrhoegenic *E. coli* (DEC) strains are known to be of global health importance, specifically afflicting humans, particularly, children under five years old in developing countries. Additionally, we utilized molecular diagnostic approaches that involve

group multiplex-PCR and PCR to detect EAEC-virulence factors, measured bacterial load for EAEC using qPCR TaqMan assay and screened the isolates for biofilm and biofilm-producing-genes in order to establish EAEC pivotal role in childhood diarrhoeal disease. Furthermore, we investigated antimicrobial resistant patterns of EAEC isolates in order to know the prevalence of the set of antibiotic resistant-EAEC strains that are in circulation in this region, discover the transmission pattern of resistant strains and ultimately establish mechanisms to reduce or eliminate the spread of the antimicrobial resistant agents.

In addition, whole genome sequencing (WGS) regarded as the hallmark of microbial diagnosis was used on some randomly selected EAEC strains in order to better understand and interrogate the genetic components of the EAEC from Gambian children.

1.4 Diarrhoeagenic *Escherichia coli* (DEC) strains at a glance

Diarrhoeagenic *E. coli* (DEC) strains which are sometimes referred to as *E. coli* pathotypes are *E. coli* strains that cause diarrhoea in the infected host when genetic element that encodes for virulence factors is acquired.

E. coli was first discovered in the gut in 1885 by the German bacteriologist-paediatrician Theodore von Escherichia, who called the organism Bacterium coli commune (Mora, Lopez et al. 2012). *E. coli* remain the commonest facultative anaerobe of the member of Enterobacteriaceae family of human colonic flora. The organism typically colonises infant gastrointestinal tract within hours of life, and, thereafter, *E. coli* and the host derived mutual benefit (Nataro and Kaper 1998). Most *E. coli* strains are normal flora of the intestinal microbiota of humans and other animals, however, a good number of the strains have acquired virulence factors that enable them to cause important intestinal and extraintestinal disease that includes diarrhoea, haemolytic uremic syndrome (HUS), haemorrhagic colitis (HC), urinary tract infection (UTI), septicaemia and neonatal meningitis. Diarrhoeagenic *E. coli* pathotypes represent a leading cause of paediatric diarrhoea in developing countries (Nataro and Kaper 1998; Estrada-Garcia, Lopez-Saucedo et al. 2009) and also an emerging cause of diarrhoea in developed countries (Robins-Browne and Hartland 2002; Cohen, Nataro et al. 2005). A wide range of bacteria, viruses and parasites cause diarrhoea (Kotloff, Nataro et al. 2013). Among the bacterial agents that are purported to possess the ability to cause severe and fatal diarrhoeal disease in younger children, DEC is the most implicated and represents a major public health problem in the developing countries (Nataro and Kaper 1998). In addition, DEC constitute a reservoir of latent diarrhoeal infection which may hinder control and future elimination. The contribution of DEC diarrhoea to malnutrition and growth impairment is likely more detrimental than even rotavirus infections (Mondal, Haque et

al. 2009; Okeke 2009). DEC strains have been classified into 7 groups, based on epidemiological, clinical and molecular characteristics: **Enteropathogenic** *E. coli* (EPEC), **Enteroaggregative** *E. coli* (EAEC), **Enterohaemorrhagic** *E. coli* (EHEC), **EnteroInvasive** *E. coli* (EIEC), **Diffusely adherent** *E. coli* (DAEC) and **Cell-detaching** *E. coli* (CDEC) (Nataro and Kaper 1998; Guion, Ochoa et al. 2008). DEC strains that are responsible for about 40% of all diarrhoeal episodes in developing countries include ETEC, EPEC and EAEC (Clarke 2001; O'Ryan, Prado et al. 2005; Dutta, Guin et al. 2013). This explains why GEMS has chosen these three DEC strains as the target microbial agents of diarrhoea. Generally, *E. coli* remain as a commensal provided genetic elements encoding for virulent factor are not acquired (Kaper, Nataro et al. 2004). Identification of DEC requires ability to differentiate these strains from non-pathogenic *E. coli* that constitute normal intestinal flora. Thus, molecular identification and classification of DEC is based on the presence of different chromosomal or plasmid-encoded virulence genes, which are absent in the commensal *E. coli*. ETEC is defined by the elaboration of the heat labile (LT) and/or heat stable (ST) enterotoxins. EPEC is characterised by attaching and effacing (AE) lesions that harbour *eae* gene that encodes structural gene for outer membrane protein Intimin. EPEC is classified into typical and atypical EPEC. Typical-EPEC (tEPEC) strains are those harbouring the chromosomal attaching and effacing (*eae*) and the EPEC adherence factor (EAF) virulence plasmid encoding bundle-forming pili (Kaper, Nataro et al. 2004), whereas atypical-EPEC (aEPEC) strains harbour the *eae* gene only (Nataro, Mai et al. 2006). Enteroaggregative *E. coli*, a well-recognised diarrhoeal pathogen mostly among children in developing world has been traditionally defined as an aggregating cell adherence phenotype. The ability of EAEC to adhere to HEp2 cell line in a stacked brick manner confers the term aggregative adherence (AA). The AA phenotype is associated with specific fimbriae

(AAFs) which are encoded by plasmids (pAAs), and an EAEC molecular marker Center for Vaccine Development 432 (CVD432) has been shown in the pAA (Dutta, Guin et al. 2013). Also, EAEC strains harbouring virulence factors under the control of AggR regulator are designated typical EAEC and strongly associated as a diarrhoeal pathogen (Nataro 2005; Nataro, Mai et al. 2006). Diffusely adherent ***E. coli*** (DAEC) is never associated with any outbreaks and volunteer studies did not result in disease (Nataro, Mai et al. 2006) so its pathogenicity status is unclear. A recently discovered additional putative pathotype, known as Cell-detaching ***E. coli*** (CDEC) harbours the *E. coli* haemolysin gene and also secretes cytotoxic necrotizing factors (CNF1) (Kaper, Nataro et al. 2004). Studies have described pathogenic mechanisms of the six DEC in detail (Nataro and Kaper 1998; Nataro and Martinez 1998; Kaper, Nataro et al. 2004; Croxen and Finlay 2010). The schematic diagram of molecular mechanisms of EPEC, EHEC, ETEC, EAEC and DAEC pathogenicity is illustrated in figures 1.3 and 1.4 below (Croxen and Finlay 2010).

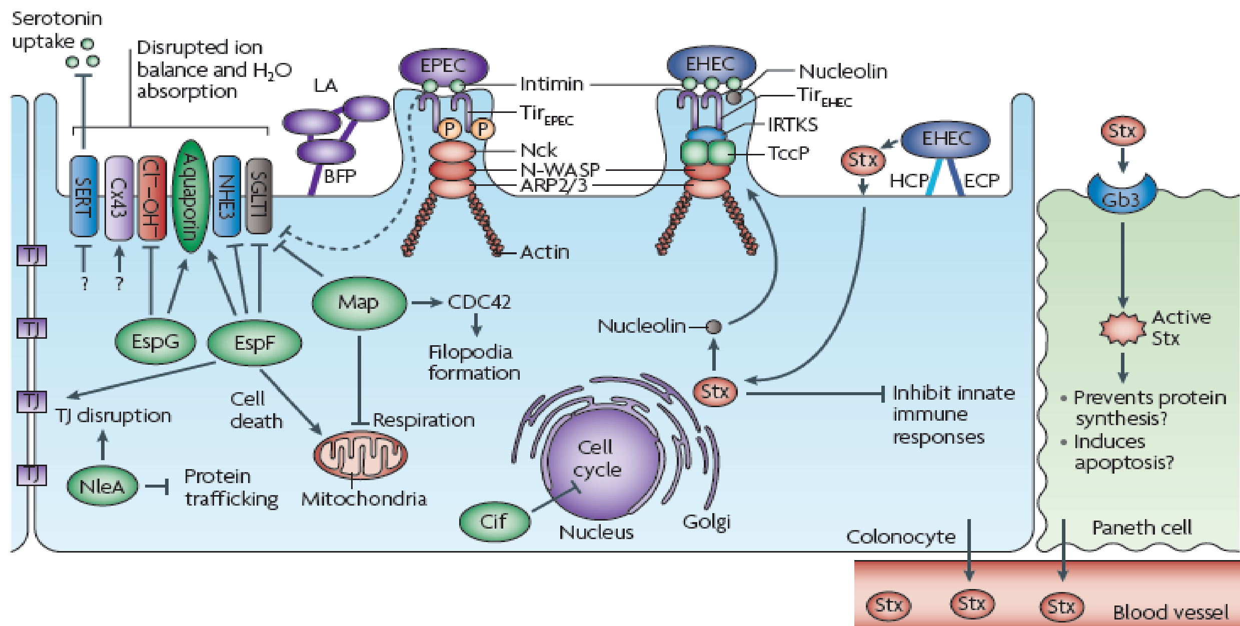


Figure 1.3: Pathogenic mechanisms of enteropathogenic and enterohaemorrhagic *E. coli*. (Adapted from Croxen and Finlay 2010)

EPEC and EHEC are attaching and effacing (A/E) pathogens that efface the microvilli and subvert host cell actin to form pedestals beneath the attachment site. The pedestal formation mechanisms shown for EPEC and EHEC are based on studies of the prototypical strains EPEC E2348/69 and EHEC O157:H7. Effectors secreted by the type III secretion system can affect Cl^{-}–OH^{-} and Na^{+}–H^{+} exchanger activity, mislocalize aquaporins and inhibit sodium-d-glucose co-transporter 1 (SGLT1) (Croxen and Finlay 2010).

EPEC attaches to the small bowel through the bundle-forming pilus (BFP), forming localized adhesions (LA). Intimate attachment is mediated by the interaction between intimin and the translocated intimin receptor (Tir). Tir is phosphorylated by host tyrosine kinases, and phosphorylated Tir recruits Nck, which activates neural Wiskott–Aldrich syndrome protein (N-WASP) and the actinrelated protein 2/3 (ARP2/3) complex to mediate actin rearrangements and pedestal formation. Using the locus of the enterocyte effacement-encoded type III secretion system, a large repertoire of effector proteins is injected into the host cell, subverting host cell pathways) (Croxen and Finlay 2010).

EHEC mechanism of pedestal formation is slightly different from that used by EPEC. Tir is not phosphorylated, and pedestal formation is Nck-independent. The actin rearrangements that are necessary for pedestal formation are mediated by Tir cytoskeleton-coupling protein (TccP; also known as EspFU), which is linked to Tir through the host protein insulin receptor tyrosine kinase substrate (IRTKS; also known as BAIAP2L1) and interacts with N-WASP to activate the ARP2/3 complex. In addition to this intimate attachment, EHEC attaches to the large bowel through the *E. coli* common pilus (ECP) and the haemorrhagic coli pilus (HCP). EHEC injects many of the same effectors as EPEC into the host cell to manipulate host processes(Croxen and Finlay 2010).

Shiga toxin (**Stx)**; also known as **verocytotoxin** is released following phage-mediated lysis in response to stress, further contributing to disease. Globotriaosylceramides (Gb3s) on Paneth cells in the human intestinal mucosa act as receptors for Stx.(Croxen and Finlay 2010).

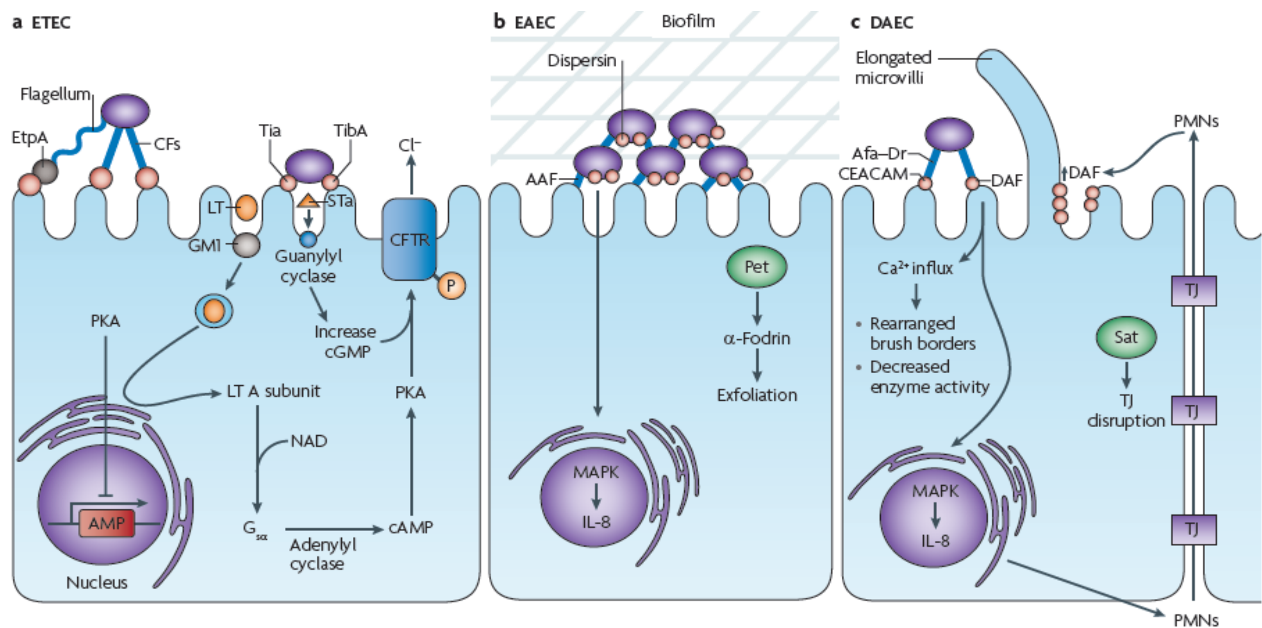


Figure 1.4: Pathogenic mechanisms of enterotoxigenic *E. coli*, enteroagreggative *E. coli* and diffusely adherent *E. coli* (Adapted from Croxen and Finlay 2010)

ETEC becomes anchored to enterocytes of the small bowel through colonization factors (CFs) and an adhesin that is found at the tip of the flagella (EtpA). Tighter adherence is mediated through Tia and TibA. Two toxins, heat-labile enterotoxin (LT) and heat-stable enterotoxin (ST), are secreted and cause diarrhoea through cyclic AMP (cAMP)- and cyclic GMP (cGMP)-mediated activation of cystic fibrosis transmembrane conductance regulator (CFTR) (Croxen and Finlay 2010)

EAEC attaches to enterocytes in both the small and large bowels through aggregative adherence fimbriae (AAF) that stimulate a strong interleukin-8 (IL-8) response, allowing biofilms to form on the surface of cells. Plasmid-encoded toxin (Pet) is a serine protease autotransporter of the Enterobacteriaceae (SPATE) that targets α-fodrin (also known as SPTAN1), which disrupts the actin cytoskeleton and induces exfoliation (Croxen and Finlay 2010).

DAEC forms a diffuse attaching pattern on enterocytes of the small bowel, which is mediated through afimbrial (Afa) and fimbrial adhesins, which are collectively known as Afa–Dr fimbriae. Most Afa–Dr fimbriae bind to complement decay-accelerating factor (DAF); a subset of Afa–Dr fimbriae bind to receptors in the carcinoembryonicantigen- related cell-adhesion molecule (CEACAM) family. The autotransported toxin Sat has been implicated in lesions of tight junctions (TJs) in Afa–Dr-expressing DAEC, as well as in increased permeability. Polymorphonuclear leukocyte (PMN) infiltration increases surface localization of DAF (Croxen and Finlay 2010).

1.5 Review of Enteroaggregative *Escherichia coli* (EAEC)

1.5.1 Discovery

The six categories of *E. coli* (ETEC, EPEC, EIEC, EHEC, EAEC and DAEC) have virulence attributes that help the bacteria to cause diseases by different mechanisms (Weintraub 2007). Three (ETEC, EIEC and EHEC) of the six categories are known to possess specific virulence attributes that include toxins, invasins and colonisation factors. Over a century ago a study confirmed set of *E. coli* serotypes later referred to EPEC that were associated with diarrhoea outbreaks (Ewing 1963). In 1979, this set of serotypes regarded as EPEC was investigated for *in vitro* adhesion assay and were found to bind to the Hep-2 cells in a localized pattern (Cravioto and Arrieta 1979). A few years later studies showed adherent non-EPEC strains that were associated with diarrhoea, the strains were named 'enteroadherent *E. coli*' (Cravioto and Arrieta 1979; Mathewson, Johnson et al. 1985; Mathewson, Oberhelman et al. 1987). Round about the same time, Nataro and colleagues observed in their experiment two different phenotypes among the enteroadherent strains which are diffuse and aggregative adherent strains (Nataro, Kaper et al. 1987). The finding of the aggregative adherent strains was the first description of EAEC by James Nataro. Thus, the aggregative adherence patterns signify a new strain with distinct class of diarrhoeagenic *E. coli* called Enteroadherent-aggregative *E. coli* (Nataro, Kaper et al. 1987) now known as Enteroaggregative *E. coli* (EAEC). In summary, a study comparing adherence patterns of 516 *E. coli* isolates from faecal samples of children with and without diarrhoea. The study used DNA probes to examine three adherence patterns to HEp-2 cells and were distinguished and described as (1.) Localised, (2.) True diffuse and (3.) Aggregative. Localised adherence was attributed to EPEC, True diffuse adherence had no association with diarrhoea whilst Aggregative adherence was associated with diarrhoea in 84 of the 253 probe negative strains from children with diarrhoea. The

aggregative adherence is characterised by a stacked-brick formation of bacterial cells attached to the Hep-2 cells (Nataro and Kaper 1998).

Heterogeneity of EAEC virulence in a volunteer study: Four EAEC reference strains used in the study include EAEC 042, 17-2, 34b and JM221 isolated from a child with diarrhoea in Lima Peru, Chile, India and an adult wit diarrhoea in Mexico respectively (Nataro et al. 1995). Twenty volunteer study participarnts were allocated into 4 groups of 5 each and each group received a different EAEC strain. It is important to mention that the 20 volunteers were screened for serologic evidence of antibodies to te 14-kDa protein encoaded by the 17-2 plasmid by Western immunoblot of which 17 were seronegative and 3 seropositive. The 3 seropositive volunteers were distributed into different groups. Study participants were given a single dose of 10^{10} colony forming unit (cfu) of bacteria.

Of the 20 study participants, only 4 experienced loose stool and the 4 are those inoculated with the dose of EAEC 042 strain and 3 of them met the case definition of diarrhoea. All the 20 study participants secreted their organism by 24 hour after inoculation and 17 continued to shed at 96 hour, at which time antibiotic therapy commenced.

The absence of disease in EAEC 042 infected individual who was seropositive for 14-kDa protein encoded by 17-2 plasmid showed that EAEC 17-2 strain is protective. A possible explanation of the EAEC 042 virulence is dependence on host age and genetic factors. For example, infants who are naturally known to possess poor adaptive immune system will easily develop diahrroea if infected by EAEC 042 strain.

Despite the fact that this experiment was hailed by most scientific community but there was no strong evidence that EAEC 042 cause disease in children and adults who are not immunocompromised. Also, the suitability of the EAEC 042 as a reference strain is doubtful because the challenged experiment showed that the strain lacks *AAF/I* which is an important

variant among the five aggregative adherence fimbriae variants that enhance the the strain to colonise host mucosal epithelial cells or enterocytes (Nataro, Yikang et al. 1994 and Jonson, Struve et al. 2015). Additionally, a reference strain suppose to be consistence in its virulence characteristics and features.

1.5.2 Classification

EAEC is classified into typical and atypical based on the presence and absence of a virulent factor *aggR*. Following the discovery of EAEC, a method to identify the new pathotype was developed using a probe that hybridised with an ATP-binding cassette transporter apparatus that translocates dispersin across the bacterial cell membrane (Nataro, Mai et al. 2006). The majority of the samples that were probe positive carries *aggR* factor but not all diarrhoeagenic strains were positive for *aggR* hence a general classification of EAEC into typical (having *aggR*) and atypical (not having *aggR*) groups was recognised (Harrington, Dudley et al. 2006; Croxen, Law et al. 2013). Further classification can be based on difference adherence patterns of some strains affinity to infect the small bowel and others infect both the small bowel and the colon (Okhuysen and Dupont 2010; Croxen, Law et al. 2013). Another recent classification was the Shiga-toxin-EAEC and non-Shiga-toxin-EAEC. Shiga-toxin-EAEC was implicated in the German outbreak EAEC in 2011 causing haemolytic uremic syndrome (HUS). This outbreak resulted in over 4,000 confirmed cases with over 54 fatalities in 14 European countries, the USA and Canada (Buchholz, Bernard et al. 2011; Karch, Denamur et al. 2012). The identified strain included features of EAEC with capacity to produce Stx 2a (Frank, Werber et al. 2011). Reports from countries such as France (Morabito, Karch et al. 1998), Japan (Iyoda, Tamura et al. 2000), Central Africa Republic (Mossoro, Glaziou et al. 2002) and Northern Ireland (Dallman, Smith et al. 2012) have established EAEC acquiring Shiga toxins among patient with HUS.

Additionally, using serology means to serogroup/serotype EAEC is enormous challenge due majorly to autoagglutination and cross reactivity of many strains that share serotypes differentially adhere to HEp-2 cells, a gold standard method use to detect and to classify *E. coli* pathotypes (Jenkins, Tembo et al. 2006; Estrada-Garcia and Navarro-Garcia 2012). Furthermore, the use of Multilocus Sequense Typing (MLST) investigation in a Nigeria study has not only shown the complexity of EAEC sequence types but also the multiplicity of EAEC pathogenic lineages that strengthens the global heterogeneity of EAEC (Okeke, Wallace-Gadsden et al. 2010; Croxen, Law et al. 2013). Many studies have shown that EAEC was found scattered among the 6 major *E. coli* phylogenetic groups (A, B1, B2, C, D and E) (Czeczulin, Whittam et al. 1999). An epidemiological study also investigated the potential clustering of EAEC strains into different phylogenetic groups found strains in phylogroups A, B1, B2 and D. This emphasise the multiple linages of EAEC origin and enhance phylogenetic diversity of the strains (Okeke, Wallace-Gadsden et al. 2010).

1.5.3 Epidemiology

1.5.3.1 Developing Countries

Many studies from developing countries have shown the association of EAEC with persistent and acute diarrhoea (Bhan, Khoshoo et al. 1989; Lima, Fang et al. 1992; Paul, Tsukamoto et al. 1994; Pai, Kang et al. 1997; Araujo, Tabarelli et al. 2007). Most diarrhoeal investigations in children have shown a significant presence in the prevalence of EAEC compared to the controls (Gonzalez, Diaz et al. 1997; Okeke 2009; Opintan, Newman et al. 2010). A case-control diarrhoeal investigation in Mongolian children showed that EAEC strains with the *agg*R gene were found to be associated with diarrhoea (Sarantuya, Nishi et al. 2004). Similarly, a study performed in Kolkata, India implicated the EAEC strains detected using gold standard technique (HeLa cell assay) and PCR amplification to be associated with diarrhoea in children <36 month of age (Dutta, Pal et al. 1999) and the study further reported more frequent cases of watery diarrhoea than cases of mucoid diarrhoeal children (72% versus 28%) (Hebbelstrup Jensen, Olsen et al. 2014). A study in Vietnam revealed frequency of EAEC associated with diarrhoea in children less than 2 years of age (Nguyen, Le Van et al. 2005). Multiple investigations revealed the association of EAEC with diarrhoea particularly in populations in low-income countries (Moyo, Maselle et al. 2007), prominently in association with persistent diarrhoea (≥14 days). Recent studies in Iran and Egypt demonstrated the implication of EAEC in paediatric diarrhoea (Ali, Ahmed et al. 2014; Bafandeh, Haghi et al. 2015). However, a high rate of carriage of EAEC in children has been reported by several other studies that include case-control study conducted 2003 and 2006 in South India (Rajendran, Ajjampur et al. 2010). Similar report was shown from other case-control studies conducted in north-eastern Brazil (Scalesky 2001), Mali (Boisen, Scheutz et al. 2012) and Gambia (Ikumapayi, Boisen et al. 2017).

1.5.3.2 Developed Countries

An earlier investigation in 1985 has implicated EAEC responsible for 14.9% of diarrhoea in a US student population visiting Guadalajara (Mathewson, Johnson et al. 1985). Again, report from Adachi's diarrhoeal investigation detected EAEC in 33% overseas travellers who are citizen of Americans, Canadians and Europeans from three different regions that include Guadalajara in Mexico, Ocho-Rios in Jamaica and Goa in India (Adachi, Jiang et al. 2001). A Scandinavian case-control study also revealed the association of EAEC with diarrhoea (Bhatnagar, Bhan et al. 1993). Likewise, East London investigation unmasked the recovery of EAEC from children suffering from acute and persistent diarrhoea (Chan, Phillips et al. 1994). Additional evidence of association of EAEC with diarrhoea in Europe and Eastern Europe was demonstrated in another case-control investigation in Germany were 2% of diarrhoeal cases revealed presence of EAEC but none fund among the healthy control (Huppertz, Rutkowski et al. 1997). Similar evidences EAEC diarrhoea were shown in cohort/case series (Presterl, Nadrchal et al. 1999; Knutton, Shaw et al. 2001) and a one year cohort investigation in Denmark where EAEC was detected in 25 (14%) among Danish children attending day-care facilities and diarrhoea was reported in 6 (24%) of the EAEC infected children (Hebbelstrup Jensen, Stensvold et al. 2016). The more confirmatory evidences were the outbreaks that occurred in Serbian neonatal ward in which some children died (Cobeljic, Miljkovic-Selimovic et al. 1996) and Japan in which school children who ate contaminated lunch were infected and developed severe diarrhoea and EAEC was detected in 10% of cases (Itoh, Nagano et al. 1997). Other outbreaks involving children and adults have been reported in United Kingdom (Smith, Cheasty et al. 1997; Spencer, Smith et al. 1999) as well as France (Boudailliez, Berquin et al. 1997).

1.5.4 Transmission and Reservoir

Transmission of EAEC is recognised as faecal-oral route (Jiang, Greenberg et al. 2002) and mostly occurs through either drinking contaminated water or consumption of contaminated food such as salads and other food deserts as reported in England case control study (Tompkins, Hudson et al. 1999) and investigation of contaminated vegetables from a popular restaurant at Guadalajara, Mexico (Koo, Jiang et al. 2008). The France outbreak of EAEC was speculated to be person-to-person transmission (Boudailliez, Berquin et al. 1997). A more current and major outbreak of Shiga-toxin (Stx) producing EAEC in northern Germany in May 2011 was epidemiologically associated with the consumption of fenugreek sprout as the most likely source of infection (Buchholz, Bernard et al. 2011). It was reported that the seeds were imported as a lot in late 2009 from Egypt, and it is still unknown if the point of contamination occurred at the site where seeds were produced, during transportation, or at the importer (Buchholz, Bernard et al. 2011). A community wide outbreak in the India village was epidemiologically associated with the consumption of water from open well (Pai, Kang et al. 1997). A study investigated the growth of EAEC in drinking water revealed that viability of EAEC strains can be up to 60 days at normal storage temperature and the strain survival in mineral water compared to spring water (Vasudevan, Annamalai et al. 2003). In Australia, EAEC was identified in water samples obtained from surface water of source of drinking water by PCR targeting *aggR* gene. A study in Bangladesh investigated water during both winter and summer detected EAEC in 17% and 4% of the water samples tested in the two seasons respectively (Akter, Islam et al. 2013).

Food handling has been speculated in the transmission of EAEC. For example, study conducted in Sao-Paulo Brazil showed presence of EAEC in 3% of milk samples investigated from infant feeding bottles that were handled by mothers (Morais, Morais et al. 1998). Food handlers, particularly those working in tourist hotels have identified primary carriers of

EAEC. An investigation of two successive foodborne outbreaks of gastroenteritis that occurred in Italy farm-holiday resort implicated a pecorino cheese prepared with unpasteurised milk as the source of the outbreaks (Scavia, Staffolani et al. 2008) suggesting the source as animal origin. The Burkina-Faso investigation on faecal samples from cattle, chicken and pigs yielded EAEC positive at 7%, 6% and 32% respectively, reiterate animals as possible reservoirs. Also, a study conducted in Brazil detected EAEC in 7.4% of dogs with diarrhoea and 3.9% of dogs without diarrhoea (Puno-Sarmiento, Medeiros et al. 2013). Again, a Gambia epidemiological study associated animal that includes Cow, fowl and ruminant with EAEC high bacterial load among diarrhoeal children (Ikumapayi 2016). However, other thorough studies have reported absence of EAEC in animals. For example, in Great Britain, 1,227 *E. coli* isolates from 401 cows, 406 sheep and 400 pigs were investigated for EAEC, interestingly, no EAEC strain was detected (Cassar, Ottaway et al. 2004). Similarly, in a French study, EAEC was not detected in wastewater or effluents in an investigation that wanted to establish source of EAEC in rivers (Bibbal, Kerouredan et al. 2014) and the target genes for this particular study are *aggR*, *aap* and *aatA*. A very recent study conducted in Spanish 46 farms (20 cattle farms, 17 sheep farms and 9 goat farms) located in Eastern and southern Spain investigated to help establish whether ruminants are a potential source of EAEC transmission to humans (Orden A. Jose 2017). Interestingly again, the result of the investigation showed EAEC negative by revealing absence of the target (*aggR*, *aap* and *aatA*) genes for *E. coli* isolates from faecal cultured samples from the 920 ruminants (Orden A. Jose 2017). Despite overwhelming evidences that exonerate animals such as cattle, sheep and goat from being reservoirs of typical EAEC pathogenic strains to humans, it is not absolute that EAEC cannot transiently colonise ruminants. Therefore, there is likelihood that ruminants can be a reservoir of EAEC and a potential source of transmission to humans. Further investigations that elucidate risk factors and the reservoir for

EAEC is needed. However, poor sanitation and crowded living conditions increase the propensity for EAEC to spread.

1.5.5 Pathogenesis

1.5.5.1 Brief explanation

Generally, understanding the complexity of interactions between host and bacteria is crucial for unmasking pathogenesis of infectious disease. In the case of enteric disease Philipson's review provides detail explanation on host-pathogen interactions (Philipson, Bassaganya-Riera et al. 2013). The intestinal epithelium is continuously exposed to trillions of microorganisms and confronts the challenge to peacefully coexist with harmless bacteria, at the same time responding to pathogens (Vossenkamper A. 2011; Philipson, Bassaganya-Riera et al. 2013). The ability for a host to resist colonisation or make infection impossible is determined by well structured cellular and molecular interactions between the host and pathogen at the mucosal interface. A single layer of epithelial cells which is the epithelial barrier provides the first line of defence against pathogenic microorganisms. The epithelial barrier integrity is formed by **"tight-junctions"** between cells and protective mucus-gel that coats the cells (Gouyer, Gottrand et al. 2011). In the Philipson's review we learnt that if an enteric pathogen passes through the mucus layers, group of established evolutionarily conserved pathogen-associated molecular patterns (PAMPs) expressed on the microbial surfaces are recognized by a set of receptors called pattern recognition receptors (PRRs) which expressed on epithelial cell surfaces such as toll-like receptors (TLRs). TLRs activate potent innate responses by triggering signalling pathways that regulate gene transcription, such as NFκB and mitogen-activated protein kinases (MAPK) and activate the production of a large repertoire of pro-inflammatory mediators to orchestrate the influx of leukocytes (Eddy and Storey 2007). More specifically, secretion of Interleukin 8 (IL-8) and Chemokine C-X-C motif Ligand1 (CXCL1) by enterocytes generates a chemotactic gradient promoting the recruitment of neutrophils to facilitate clearance of bacteria through phagocytosis (Eckmann and Kagnoff 2005). Epithelial cells also secrete Chemokine C-C motif Ligand 20 (CCL20) in

response to enteric pathogens to enhance infiltration of cells expressing Chemokine Receptor 6 (CCR6). Dendritic cells expressing CCR6 are brought to the underlying lamina propria to hasten antigen presentation and activation of the adaptive immune system (Wells, Rossi et al. 2011). Th17 cells are CCR6+ and implicated as primary contributors to defence against extracellular bacterial infections. In addition to the secretion of cytokines to mediate cellular trafficking, epithelial cells produce potent antimicrobial proteins such as β-defensins, cathelicidins and calprotectin in response to stimulation from enteric pathogens or proinflammatory cytokines for further defence against infection (Eckmann and Kagnoff 2005). Importantly, a great amount of attention has recently shifted away from the host response and toward understanding the protective barricade created by commensal microbiota during infection (Littman and Pamer 2011). The combined efforts of innate and adaptive immune responses with the beneficial influence of the gastrointestinal microbiome generally contribute to successful eradication of disease in healthy individuals.

Multidrug resistance efflux pumps including the AcrAB-TolC system have been reported to be associated with the colonization and persistence of bacteria in the host and to have roles in bacterial pathogenicity (Piddock 2006). One study suggested that EAEC strains possessing CVD432 and EAEC heat-stable enterotoxin-1 (EAST-1) virulence markers are most commonly associated with chronic diarrhoea in children (Pereira, Ferraz et al. 2007); whereas another study suggested that EAEC strains possessing *aggR*, *aap*, and *astA* that encodes for EAST-1 protein are most commonly associated with acute diarrhoea in adults (Huang, Mohamed et al. 2007).

The pathogenesis of EAEC is complex. It may be interesting to know that despite numerous outbreaks and several high level pathogenesis investigations conducted the understanding of EAEC and its pathogenesis is still not definite, partly due to the paucity of suitable animal models and heterogeneity of its virulence factors. Pathogenic bacteria such as EAEC have

developed strategic mechanisms to conceal recognition and consequently enhance survivability during interaction with its host; these strategies are mostly driven by genetically encoded virulence factors. EAEC strains harbour a 60- to 65-MDa virulence plasmid (pAA) that encodes many of the known virulence factors including the aggregative adherence fimbriae (AAF), Pet toxin, the transcriptional regulator *AggR* and the secretory protein dispersin. A key virulence factor harboured by pAA is the transcriptional activator *AggR* which is considered the master regulator of virulence due to its capability to activate a large cluster of virulence genes in EAEC permitting adherence while also promoting the production of cytotoxins and enterotoxins (Aslani, Alikhani et al. 2011). A study has proven that *Agg*R activates the expression of at least 44 genes in the EAEC prototype strains 042 (Morin, Santiago et al. 2013). Additionally, in order to mediate secretion of protein, EAEC possess a type VI secretion system (T6SS) that is chromosomally encoded on the pathogenicity island *pheU* and transcriptionally regulated by *AggR*. Two gene clusters known as Sci-1 and Sci-2 are present on *pheU* are responsible for encoding T6S machines (Dudley, Thomson et al. 2006). Also, the identification EET2 gene cluster in the EAEC 042 genome sequence showed evidence for T3SS prevalence (Ren, Chaudhuri et al. 2004). As it is speculated that these secretion systems play a key role in EAEC virulence due to expulsion of toxic proteins and association with biofilm formation (Aschtgen, Bernard et al. 2008) yet their fundamental roles in mechanisms of pathogenesis is unknown (Philipson, Bassaganya-Riera et al. 2013). Therefore, heterogeneity among EAEC strains remains a major factor that complicates our understanding of pathogenic mechanisms underlying infection; an accomplice too is the limited rigorous studies to show definite immunoregulatory responses by the host that potentiate EAEC clearance. Nonetheless, many studies have suggested that infection can be summarised in three general stages that are; (1) adherence and colonisation,

(2) increase mucus or biofilm production and (3) toxin release and host response (Estrada-Garcia and Navarro-Garcia 2012).

1.5.5.2 Adherence and Colonisation

As custom for pathogenic enteric bacteria, attachment to the intestinal mucosa is the first step in colonisation and production of disease by EAEC. The defining feature of pathogenic EAEC strains is their ability to produce the staked-brick-pattern adherence termed aggregative adherence (AA) that consolidate firmness and abundance adherence of these bacteria to the intestinal mucosa. The plasmid-borne *aggR* gene is an important gene for the pathogenesis and adherence properties of EAEC, where strains possessing the *aggR* gene are known as "typical EAEC strains" (Morin, Santiago et al. 2013). The best studied virulence factor is *aggR*, a well-recognised transcriptional activator that promotes the expression of both chromosomal and plasmid-encoded virulence factors, including AAF and dispersin (aap). The adhesion is facilitated by fimbriae termed aggregative adherence fimbriae or factor (AAF). Although three fimbriae (AAFs) encoded by the pAA plasmid are responsible for EAEC adherence which are agg*A* (AAF/I), aaf*A* (AAF/II), agg-3 (AAF/III). Each EAEC isolate carries only one AAF subtype at a time. agg*A* is responsible for aggregative phenotype and human erythrocyte haemagglutination of EAEC (Nataro 2005), aaf*A* allows EAEC to adhere to the intestinal mucosa (Czeczulin, Balepur et al. 1997), aag-3 function as an adhesion (Bernier, Gounon et al. 2002). Earlier studies have showed stacked-brick aggregative adherence with Hep-2 cells, likewise biopsies from paediatric intestinal mucosa cultured with EAEC strains 17-2 and 221 demonstrate EAEC ability to adhere to jejuna, ileal and colonic mucosa (Hicks, Candy et al. 1996). Furthermore, three membrane-associated proteins (MAP), of 18, 20 and 80kDa, are believed to play an important role in EAEC adherence to and haemagglutination of animal cells (Monteiro-Neto, Bando et al. 2003). A study conducted in Sao Paulo Brazil characterised OMP profiles of EAEC strain from children with diarrhoea observed heterogeneity in OMP profiles (Monteiro-Neto, Bando et al. 2003). Hence, the binding of EAEC fimbriae to components of the extracellular matrix

proteins of intestinal epithelial cells, such as laminin, type-IV collagen, cytokeratin 8, and fibronectin is an initial step in adherence to intestinal mucosa (Farfan, Inman et al. 2008; Izquierdo, Navarro-Garcia et al. 2014).

The characteristics of four recognised AAFs vary between EAEC strains both in morphology and genetic code nonetheless all mediate key function of bacterial attachment to epithelial cells. The five major variants of AAF, with distinct structure of pilin subunits agg*A* (AAF/I), aaF*A* (AAF/II), agg-3 (AAF/III) and AAF/IV have been found in prototype strains EAEC17-2, 042, 55989 and C1010-00 respectively and the four strains develop the aggregative adherence (AA) phenotype (Bernier, Gounon et al. 2002; Harrington, Dudley et al. 2006; Boisen, Struve et al. 2008). The fimbriae often splay out from the bacteria due to the surface protein dispersin encoded by *aap* (Sheikh, Czeczulin et al. 2002; Harrington, Dudley et al. 2006). Dispersin is known to induce changes in the electrostatic surface of the lipopolysaccharide layer of the bacteria which is a demonstration of key role for the adherence properties of EAEC (Harrington, Dudley et al. 2006).
In a more clear term, AafA, the major pilin protein of AAF fimbria, is directly linked to diminished transepithelial resistance (Strauman, Harper et al. 2010). The expression of AAF/I, AAF/II and AAF/IV is sufficient for the induction of polymorphonuclear cell transmigration *in vitro* (Philipson, Bassaganya-Riera et al. 2013). AAFs are highly hydrophobic thus enhancing agglutination in an aqueous environment. In order to propagate the spreading of EAEC for effective attachment and colonisation EAEC secretes a low molecular weight protein Dispersin (*aap*), a positively charged hydrophobic surface protein that maintains electrostatic interactions with the outer lipopolysaccharide layer of the bacteria preventing the positively charged AAF from clinging to bacterial membrane (Sheikh, Czeczulin et al. 2002; Mortensen, Fowlkes et al. 2011). Dispersin is responsible for mediating an antiaggregation phenotype by inducing changes in the outer membrane

polarisation of the bacterial cell which requires an ABC transporter system encoded by att (Velarde, Varney et al. 2007). Also, another transporter located in the outer membrane protein (OMP) called TolC, which is encoded by the aatA has been associated with the secretion of a yet to be characterised factor that contribute to aggregation (Imuta, Nishi et al. 2008). In a volunteer challenge study, dispersin was shown to be highly immunogenic, suggesting that it is a potential vaccine candidate (Nataro, Deng et al. 1995). Factually, AAF fimbriae collapse in the absence of dispersin and lack functionality critical for adherence (Harrington, Sheikh et al. 2009). In addition to these AAF variants, some EAEC strains can encode alternative fimbrial structures, such as type IV pili in EAEC strain C1096 (Cobeljic, Miljkovic-Selimovic et al. 1996; Dudley, Abe et al. 2006). However, some EAEC strains lack AAFs, and their aggregative adherence has been linked to the *hral* gene on the genome (also known as *hek*) or to possession of alternate adhesions such as *Hda*A (Bhargava, Johnson et al. 2009) regulated by aggR regulon but distantly related to the Dr family of adhesions. Undeniably, other unravels AAF and adherence factors exist that need unravelling. Other molecules associated with EAEC colonisation include Pic – a member of serine protease autotransporter (SPATE) that encodes on the chromosome of EAEC strain 042 and thought to play a role in EAEC colonisation and growth. Pic possesses hemagglutinin and mucinolytic activity that enable it to penetrate the intestinal mucus layer and enhance the use of nutrients from mucin for possible development of EAEC (Harrington, Sheikh et al. 2009; Philipson, Bassaganya-Riera et al. 2013). The importance of Pic protein in the pathogenicity of EAEC cannot be underestimated as the protein has been detected as an important virulent factor in other enteropathogens such *S. flexneri* and uropathogenic *E. coli*. Additionally, human neutrophils challenged with purified Pic protein resulted in impaired chemotaxis and transmigration but increased activation of the neutrophil oxidative burst while activated T cells experience Pic-induced apoptosis (Ruiz-Perez, Wahid et al. 2011).

1.5.5.3 Increase Mucus or Biofilm production

Following the successful adherence of epithelial cells by EAEC, the epithelial cells are stimulated to produce a thick mucus layer above the enterocytes forming a biofilm (figure 1.5). Formation of biofilm is an important pathogenicity trait of EAEC and the formation is mainly in the colon although report has implicated small intestine too (Hicks, Candy et al. 1996). The formation of biofilm play a key role in persistent infection by allowing bacteria to evade the local immune system and by restricting the transport of antibacterial factors that include antibiotics (Tokuda, Nishi et al. 2010). Colonisation is consolidated by EAEC encased themselves with biofilm and recruit cells forming micro-colonies that are interspersed within fluid-filled channels (Mohamed, Huang et al. 2007; Garnett and Matthews 2012). The biofilm then protects the colonies restricting antimicrobial penetration. Animal and *in vitro* culture studies demonstrate that EAEC survives within the mucus layer, elucidating our understanding why individuals infected, especially children in developing countries with pre-existent malnutrition, may develop mucoid stools, malnutrition, and persistent colonization with prolonged diarrhoea. However, the mucus layer and biofilm possibly do not explain malnutrition in affected children. This is because, for the biofilm to impair nutrient absorption, it would have to cover most of the small intestinal mucosa, but there is no evidence that this actually occurs. It is more likely that inflammatory responses or altered intestinal microbiota are primarily responsible (Kaur, Chakraborti et al. 2010). Many studies have shown the importance of expression of AAF for biofilm formation by EAEC (Sheikh, Hicks et al. 2001; Sheikh, Czeczulin et al. 2002; Boisen, Struve et al. 2008). Other factors involved in the formation of biofilm include *shf* gene shown to be important for solid biofilm production in EAEC strain 042 (Fujiyama, Nishi et al. 2008). The ShF gene codes for the 32.8-kDa Shf protein has been localized in one of the three open reading frames between *aafC* and *aatA*, and has also been implicated in biofilm formation (Fujiyama, Nishi et al.

2008), and it is predicted to be similar to IcaB, a mediator of biofilm formation in *Staphylococcus epidermidis* (Heilmann, Schweitzer et al. 1996). Also important are the implication of *yafK* that codes for a 28-kDa protein and *fis* gene, which codes for a DNA-binding protein involved in growth regulation likely due to their involvement in the regulation of AAF gene expression (Sheikh, Hicks et al. 2001). Other registered genes associated with biofilm formation are plasmid-borne *aatA* gene (Mohamed, Huang et al. 2007) encoding the dispersin transporter; the *set1* gene (Wani, Hussain et al. 2012); and the *aggR* gene (Mendez-Arancibia, Vargas et al. 2008). EilA, a HilA-like regulator, and *air*, encoding the predicted OMP in EAEC chromosome, are associated with biofilm formation (Sheikh, Dudley et al. 2006). Interestingly, loss of biofilm formation and diffuse adherence pattern was observed in EAEC at pH 4.0 whereas at pH 7.4, typical aggregative adherence pattern was observed (Figure 1.4) (Kaur, Chakraborti et al. 2010).

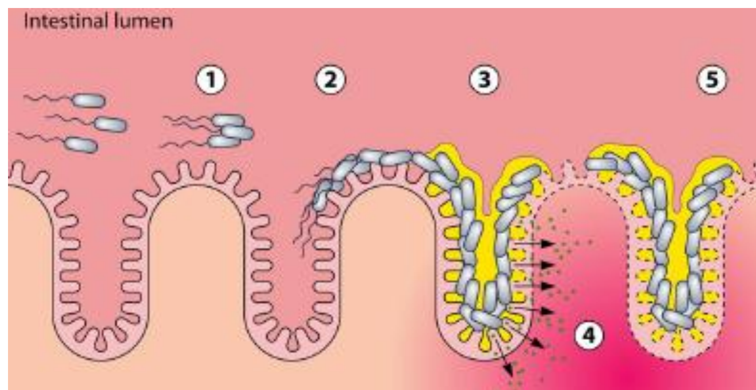


Figure 1.5: Stages of pathogenesis of EAEC Numbers circle guide to show the progression of EAEC pathogenesis

(1) Agglutination of planktonic EAEC bacteria.
(2) Adherence to the intestinal epithelium and colonisation of the gut.
(3) Formation of biofilm.
(4) Release of bacterial toxins, inducing damage to the epithelium and increased secretion.
(5) Establishment of biofilm. (Hebbelstrup Jensen, Olsen et al. 2014).
(Adapted from Betina Hebblstrup Jensen 2014. Epidemiology and Clinical Manifestations of Enteroaggregative *Escherichia coli*. Clinical Microbiology Reviews P.614-630).

1.5.5.4 Release of Toxins

Once the biofilm formation is established, more damage to the intestinal epithelium is required to consolidate the release of bacterial toxins (Hebbelstrup Jensen, Olsen et al. 2014). The secretion of toxins is thought to play an important role in secretory diarrhoea which is a clinical manifestation of EAEC infection (Harrington, Dudley et al. 2006; Arenas-Hernandez, Martinez-Laguna et al. 2012). The putative enterotoxins and cytotoxins that EAEC secretes in this EAEC stage of pathogenesis elicit a host inflammatory response that resulted in mucosal toxicity causing morphological changes in the structure of the mucosa characterised by microvillus vesiculation, enlarged crypt openings and increased epithelial cell extrusion (Harrington, Dudley et al. 2006). Three main enterotoxins have been identified which are EAEC heat-stable enterotoxin-1 (EAST1), plasmid-encoded enterotoxin (Pet) and Shigella-enterotoxin-1(ShET1).

EAST1 is a 4.1 kDa toxin first detected in EAEC strain 17–2 that has now been associated with other diarrhoeagenic strains of *E. coli* providing evidence for its relationship to enteropathogenic induced diarrhoea (Menard and Dubreuil 2002). The role of EAST1 in the molecular pathogenesis is not clearly understood, although it has been hypothesized that the toxin enhances the initial phase of watery diarrhoea seen in many patients (Savarino, McVeigh et al. 1996). EAST1 binds to the extracellular domain of guanylate cyclase (GC) on the apical membrane of enterocytes and then induces high production levels of cGMP inside cells inhibiting the Na/Cl transport system. This significantly reduces the absorption of electrolytes and water from the intestine at the villus tips resulting in elevated secretion of water in crypt cells (Telli, Guiral et al. 2010).

Pet a serine protease autotransporter enterotoxin generates high toxicity in human epithelial cells resulting in structural damage to the cell. After internalization via receptor-mediated endocytosis, Pet is delivered to the cytoplasm by means of retrograde trafficking

accompanied by cleavage of spectrin, also known as the actin-binding protein fodrin (Fig. 1.4) within microvilli cytoskeleton leading to cell elongation, exfoliation, rounding and ultimately the release of cells from the substratum (Croxen and Finlay 2010; Navarro-Garcia 2010).

ShET1 enterotoxin encoded by the *set* gene was first identified in *S. flexneri* and may be associated with increased fluid secretion (Fasano, Noriega et al. 1995). The toxin-induced damage observed in the intestinal epithelium, associated with EAEC infection, other complication includes haemorrhagic necrosis and shortening of villi, enlarged crypt openings, and formation of crypt abscesses (Nataro, Hicks et al. 1996; Navarro-Garcia, Sears et al. 1999). ShET1 appears to induce intestinal secretion via cAMP and cGMP however much of the biochemistry and mechanism of action surrounding this toxin remain elusive (Navarro-Garcia and Elias 2011).

The SPATEs constitute a large family of extracellular proteases secreted by *Enterobacteriaceae* via the type-V secretion system (Dudley, Thomson et al. 2006). The SPATE genes can be either chromosomal or plasmid borne, organized into 2 phylogenetically different classes: class I SPATEs are cytotoxic to epithelial cells and include proteins encoded by the *pet*, *sigA*, and *sat* genes, whereas non-cytotoxic class II SPATEs have more diverse effects and include proteins encoded by the *pic* and *Shigella* extracellular protease (*sepA*) genes (Boisen, Scheutz et al. 2012). Pet (plasmid-encoded toxin) cleaves spectrin in the epithelial cytoskeleton (Navarro-Garcia, Sears et al. 1999) resulting in the deformation and exfoliation of the cell, and is associated with mucoid stools (Eslava, Navarro-Garcia et al. 1998). Sat (secreted autotransporter toxin) originally discovered in uropathogenic and diffusely adhering *E. coli,* has been described as the most commonly detected SPATE among EAEC strains it cleaves the intracellular protein spectrin and cause cytoskeletal damage to tight junctions between intestinal epithelial cells (Guignot, Chaplais et al. 2007). Also, sat has

been shown to cause loosening of cellular tight junctions in kidney cells and vacuolation in both kidney cells and bladder cells (Guyer, Radulovic et al. 2002). SigA, a SPATE largely associated with *S. flexneri* pathogenesis, is capable of inducing fodrin degradation causing catastrophic morphological changes in cells (Al-Hasani, Navarro-Garcia et al. 2009). Pic (protein involved in intestinal colonization) is a mucinase that interferes with the integrity of the mucus membrane and induces serum resistance and hemagglutination (Henderson, Czeczulin et al. 1999; Navarro-Garcia 2010). *SepA* is the SPATE most strongly associated with severe diarrheal illness (Boisen, Ruiz-Perez et al. 2009) but only moderately prevalent in EAEC strains and its key role in EAEC pathogenesis remain largely uncharacterised (Philipson, Bassaganya-Riera et al. 2013).

1.5.5.5 Immunologic / Inflammatory response

Almost simultaneous to adherence, EAEC induces a host inflammatory response. The initial inflammatory response to EAEC infection is dependent on the host innate immune system and the type of EAEC strain causing the infection. EAEC carrying "virulence" genes are not always associated with disease; however, virulence factors such as flagellin, *AggR*, AAF fimbria and dispersin are associated with increased levels of faecal cytokines and inflammatory markers, such as interleukin (IL)-1ra, IL-1*β*, IL-8, interferon (INF)-*γ*, and inflammatory markers that include lactoferrin, faecal leukocytes, and occult blood (Greenberg, Jiang et al. 2002). IL-8 is an important proinflammatory chemokine involved in EAEC pathogenesis and is responsible for recruiting neutrophils to the epithelial mucosa without mucosal injury, and facilitates intestinal fluid secretion (Kucharzik, Hudson et al. 2005). Other than IL-8 production by EAEC strain, EAEC strain 042 has been shown to induce production of other proinflammatory cytokines that include IL-1β, IL-6, IL-12, and TNF, as well as IL-10 a regulatory cytokine (Estrada-Garcia, Perez-Martinez et al. 2014). In addition, *in vitro* studies have shown that EAEC induces the activation of mitogen-activated protein kinases (MAPK) on intestinal cells, that in turn activates the transcriptional factor NF-κB leading to the secretion of IL-8 and potentially other cytokines (Goyal, Konar et al. 2010; Khan, Konar et al. 2010). EAEC also activates the production of eicosanoid-based PMNs, including neutrophils, chemoattractant, which in turn, leads to the recruitment and transmigration of neutrophils to the gut mucosa, causing intestinal damage that may promote EAEC colonization (Boll, Struve et al. 2012). Both cytokine production and PMN transmigration contribute to EAEC pathogenesis and are a hallmark of inflammatory infectious diarrhoea.

1.5.5.6 Host genetic / susceptibility factor

Clinical manifestations of EAEC diarrhoea vary from individual to individual, depending upon the genetic composition of the host (Kaur, Chakraborti et al. 2010). Thus, genetics plays an important role in determining the host's susceptibility to diarrheal illness. Many studies of genetic susceptibility to infection with enteric pathogens have been carried out in adult subjects who developed traveller's diarrhoea (Flores and Okhuysen 2009; Mohamed, DuPont et al. 2011). IL-8 a proinflammatory chemokine that functions as a neutrophil chemoattractant by involve in the recruitment and the transmigration of neutrophils into the intestinal mucosa and then disrupts epithelial tight-junctions, ultimately induce colitis. This mechanism of action is common among diarrhoea-inducing pathogens (Philipson, Bassaganya-Riera et al. 2013). The presence of an AA genotype at the−251 position in the IL-8 promoter region homozygous for a single nucleotide polymorphism (SNP) produces higher levels of faecal IL-8 and more frequently develops symptomatic EAEC diarrhoea than those heterozygous for the gene after exposure to EAEC (Jiang, Okhuysen et al. 2003). In addition to IL-8, intestinal epithelial cells infected with EAEC 042, the prototype strain, upregulate the expression of IL-6, tumour necrosis factor (TNF)-α, growth-related gene product (GRO)-α, GRO γ, intercellular adhesion molecule (ICAM)-1, granulocyte macrophage colony-stimulating factor (GM-CSF), and IL-1α. These cellular responses are primarily mediated by flagellin (fliC), a major bacterial surface protein of EAEC (Harrington, Strauman et al. 2005), which causes IL-8 release by binding to Toll-like receptor 5 (TLR5). TLR5 signals through P38 mitogen-activating protein kinase (MAPK) and nuclear factor-kappa B (NF-κB) induce transcription of pro inflammatory cytokines from monocytic cells (Kaur, Chakraborti et al. 2010).

Polymorphisms in the promoter genes of lactoferrin – an iron-binding antimicrobial glycoprotein, osteoprotegerin, and Cluster of Differentiation 14 (CD14) are important

elements of the intestinal immune system, have been associated with increased susceptibility to diarrhoea in US travellers to Mexico. Thus, small variations in the genome can determine the susceptibility to a particular pathogen and/or influence disease severity. As the contribution of variations in a single gene on disease susceptibility or severity is at most modest, work is still needed to identify other host genetic factors that are important in determining susceptibility to EAEC infection.

1.5.5.7 Pathogenicity Islands

Different pathogenicity islands have been identified in EAEC strains. Study of a genomic island at the tRNA *pheU* locus, encodes the *aaiC*-associated type VI secretion system, which is regulated by the *aggR* gene (Dudley, Thomson et al. 2006; Hebbelstrup Jensen, Olsen et al. 2014). The *Shigella* species *she* pathogenicity island found in some EAEC strains encodes the SPATEs Pic and ShET1 enterotoxin, thereby conferring toxic and mucinolytic activities (Henderson, Czeczulin et al. 1999). Furthermore, two pathogenicity islands associated with extraintestinal *E. coli* strains, the *Yersinia* high-pathogenicity island, encoding the yersiniabactin siderophore, and the *hly* pathogenicity island, encoding hemolysin and P-fimbriae, have also been found in EAEC isolates (Schubert, Rakin et al. 1998).

1.5.5.8 Malnutrition

From public health perspective, the most significant outcome of EAEC infection and EAEC persistent diarrhoea is on malnourished children living in developing countries, because it has been linked with growth shortfalls and decreased intellectual development of these children (Lima and Guerrant 1992; Steiner, Lima et al. 1998). EAEC persistence in human intestine subclinically induces chronic inflammation in the absence of diarrheal disease (Steiner, Lima et al. 1998; Opintan, Newman et al. 2010). The strains exert a complex pathogen-host immune interaction where the host inflammatory response to EAEC infection is dependent on the host innate immune system and the EAEC strain. Pathogenically, EAEC infection is characterized by release of cytokines from the intestinal mucosa and lactoferrin (Steiner, Lima et al. 1998; Greenberg, Jiang et al. 2002). These observed inflammatory potentials of the EAEC strains resulted in damaging the intestinal epithelium and reducing its absorptive function, leading to nutrient depletion and malnutrition. In turn, malnutrition further facilitates the infection and perpetuates the cycle of infection (Guerrant, Oria et al. 2008). Growth retardation due to EAEC infection was observed in a mouse model (Roche, Cabel et al. 2010). The growth impairment was found to be dependent on the dose of bacteria used for challenge. It was observed that malnourished EAEC-inoculated mice had reduced growth velocity and increased shedding of EAEC in stools compared to nourished mice (Philipson, Bassaganya-Riera et al. 2013).

1.5.6 Clinical Manifestation (Symptoms)

1.5.6.1 Brief description

The clinical features of EAEC illness have been described in volunteer studies, outbreaks and sporadic cases. The characteristic clinical picture includes watery secretory diarrhoea, often with mucus, with or without blood, low-grade fever, abdominal pain, nausea and vomiting (Nataro, Deng et al. 1995; Adachi, Ericsson et al. 2002). Although bloody diarrhoea is not a distinctive feature of EAEC illness, a study reported that one-third of affected infants less than two years of age had grossly bloody stools (Cravioto, Tello et al. 1991). In a recent study we conducted among children that were hospitalized for diarrhoea EAEC was identified in~7%. In those 20 patients in whom EAEC was the only etiological agent identified, 55% had mucus in faeces, 50% had more than six stool movements per day and 10% had bloody stools (Estrada-Garcia, Perez-Martinez et al. 2014). Overall, EAEC diarrhoeal episodes have been frequently associated with the presence of mucus, PMNs and lactoferrin in stools (Adachi, Ericsson et al. 2002; Cennimo, Abbas et al. 2009; Opintan, Newman et al. 2010). The site of colonization is believed to include the colon and the terminal ileum (Hicks, Candy et al. 1996; Andrade, Freymuller et al. 2011). The incubation time ranges from 8 h to 52 h (Huang, Koo et al. 2004; Scavia, Staffolani et al. 2008). A study by Steiner et al. in 1998 found that children in developing countries who were diagnosed with EAEC infection suffered from growth retardation regardless of the presence of diarrhoea (Steiner, Lima et al. 1998). Bloody diarrhoea has been reported only rarely and involves mostly small children (Sarantuya, Nishi et al. 2004; Denno, Shaikh et al. 2012). However, the German O104:H4 EAEC Shiga toxin-expressing outbreak strain caused haemorrhagic colitis and haemolytic uremic syndrome (HUS), leading to considerable morbidity and casualties (Rasko, Webster et al. 2011; Scheutz, Nielsen et al. 2011). The outbreak strain contained the EAEC genes *aggR*, *aggA*, *set1*, *pic*, and *aap* and a prophage encoding the *stx*2 gene (Bielaszewska, Mellmann et

al. 2011). Urinary tract infections (UTIs) associated with EAEC (Guyer, Radulovic et al. 2002; Olesen, Scheutz et al. 2012) and one case of urosepsis in an immunosuppressed female (Herzog, Engeler Dusel et al. 2014) have also been described recently. An outbreak of UTIs associated with EAEC in 1991 was reported in a Danish study (Olesen, Scheutz et al. 2012), where the UTI outbreak strain contained the following combination of EAEC genes: *sat*, *pic*, *aatA*, *aggR*, *aap*, *aaiC*, and *aggA*.

1.5.6.2 Shigatoxin producing EAEC strain

Genetic variability in both host and EAEC strains can significantly impact the susceptibility and outcome of EAEC infection. For example, the capacity for specific EAEC strains to produce Stx2 and cause HUS-induced mortality demonstrates enhanced virulence (Boisen, Melton-Celsa et al. 2015). Likewise, host age dictates disease severity which explains why infected children are more susceptible to persistent EAEC diarrhoea compared with healthy adults.

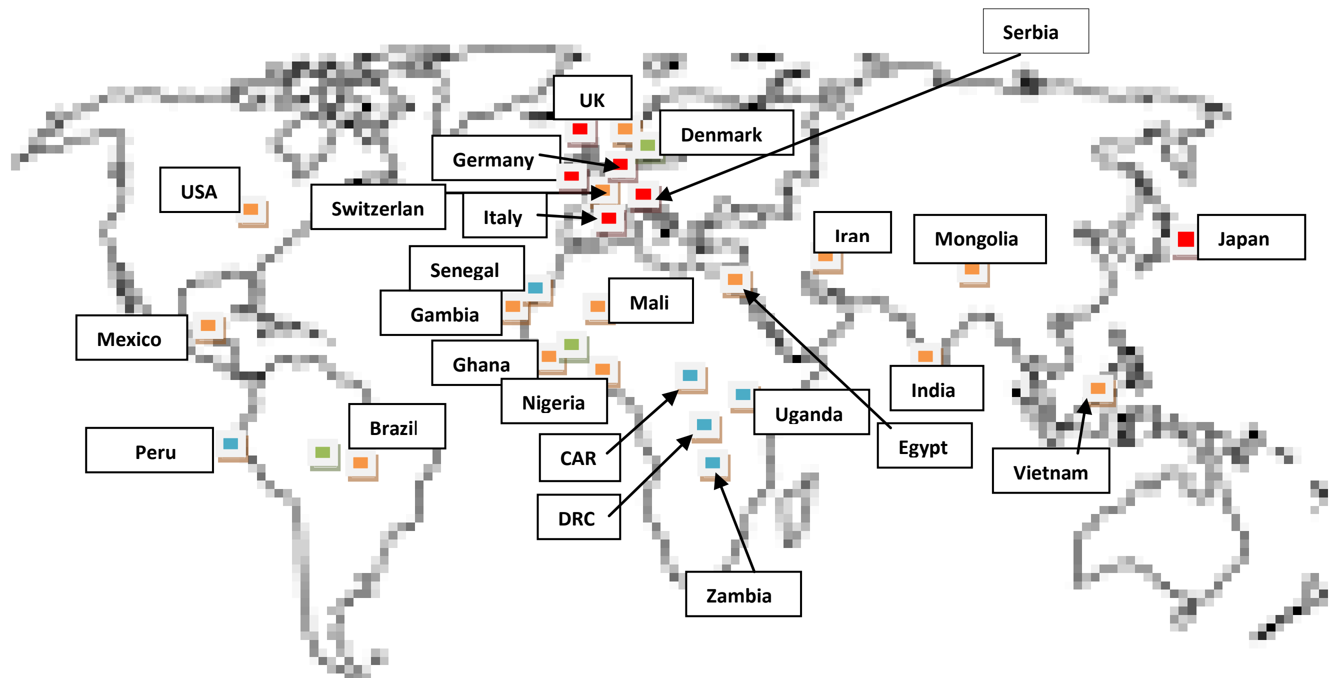
The German outbreak in 2011 was caused by EAEC strain that has adopted the ability to produce Shiga-toxin (Stx2) strain serotype O104:H4. The strain is chromosomally encoded cytotoxic verotoxic that targets globotriaosylceramide (Gb3) receptors located on host intestinal and kidney cells. Death from infection with Stx2-producing EAEC strains is strongly linked to the development of haemolytic uremic syndrome (HUS), a life-threatening disease induced by Stx2 shortly after the onset of diarrhoea. Stx2 undergoes retrograde transport to induce endothelial cell apoptosis causing significant gastrointestinal damage (Philipson, Bassaganya-Riera et al. 2013). Additionally, Stx2 is able to enter systemic circulation and induce glomerular occlusion as blood is filtered through the capillary arrangement in the kidney. The resulting haemolytic anaemia and acute renal failure are

complications that most commonly affect children and contribute to increased mortality rates (Lingwood, Binnington et al. 2010; Pacheco and Sperandio 2012). Interestingly, in the German outbreak, death occurred in patients who had not developed HUS; these cases most commonly occurred in elderly females (Frank, Werber et al. 2011). Acquisition of a Stx2 bacteriophage is the leading factor for hypervirulence, phenomenon that may have occurred in mammalian intestines or an environment where both human and ruminant faeces were present (Laing, Zhang et al. 2012; Philipson, Bassaganya-Riera et al. 2013). Survivability and Shiga toxin production alone are not likely the sole causes of HUS in EAEC infected patients. EAEC O104:H4 adherence to the intestinal mucosa is mediated by AAF/I and potentially more aggressive than EHEC LEE mediated adherence (Philipson, Bassaganya-Riera et al. 2013). Additionally, EAEC infections induce proinflammatory responses and epithelial barrier disruption possibly enhancing systemic dissemination of shiga-toxin and HUS induction providing an explanation for the strain's hypervirulent activity. In addition to Stx2 gaining systemic accessibility, severe epithelial damage induced by the toxin could have allowed bacterial components to enter peripheral blood exaggerating inflammation systemically leading to death by sepsis in non-HUS patients (Philipson, Bassaganya-Riera et al. 2013).

1.5.6.3 Outbreak of EAEC

Prior the German outbreak of Stx2 producing EAEC strain, there were other reported outbreaks of gastroenteritis linked to EAEC. For example, the largest reported outbreak of EAEC, occurred in 1993 in Gifu prefecture, Tajimi, Japan where 2,697 children developed food poisoning symptoms following consumption of school lunches (Itoh, Nagano et al. 1997). Twelve of the 30 faecal samples collected from 30 children with severe protracted diarrhoea tested positive for EAEC by the Hep-2 cell assay and the *astA* gene by PCR, interestingly the strains were untypable:H10. Another outbreak was the Serbian nursery outbreak in 1995, where EAEC was detected by the HEp-2 cell assay and were belonging to serotype O4 in 12 of 19 babies who had fever, diarrhoea and weight loss (Cobeljic, Miljkovic-Selimovic et al. 1996). Another outbreak implicating EAEC occurred in a police institute in Japan in 2005, staff experienced gastroenteritis following consumption of food suspected of being contaminated with EAEC. Investigation revealed four staff member and one food handler tested positive for EAEC in stool samples and the identified strains have an identical serotype O126:H27 (Harada, Hiroi et al. 2007). Another EAEC outbreak took place in 2008 in Italy at a farm holiday resort, where the guests developed gastroenteritis after having consumed unpasteurized cheese (Scavia, Staffolani et al. 2008). EAEC was isolated in stool samples from six restaurant guests and one staff member; the strains were identified by the HEp-2 cell assay and the EAEC strains identified belong to serotype O92:H33, and were tested positive for the virulence genes *aggR*, *aat*, *aap*, and *set1A* by PCR. Testing for norovirus, but not rotavirus, was performed in the study. The number and devastating effect of EAEC outbreaks reported showed a considerable potential for food-borne transmission of EAEC. So, lack of community, national and international surveillance of EAEC can lead to missed cases of diarrhoea outbreaks caused by EAEC.

Figure 1.6: Showing countries with EAEC outbreaks, case series and cohort studies



Outbreaks, Case series, Malnourishment & HIV-infected person

http://besttabletfor.me/outline-of-the-world-map.html/outline-world-map-throughout-of-the

1.5.7 Laboratory Diagnosis

1.5.7.1 Aggregation adherence assay

The gold standard method to identify EAEC is to sub-culture confirmed five colonies of *E. coli* per patient in static Luria Bertani broth at 37°C overnight and then infect semiconfluent Hep-2 cells for three hours and look for the aggregation adherence (AA) pattern. A positive EAEC strain aggregates to produce a hallmark "stacked-brick" appearance, where the bacilli are elongated and sometimes line up in a single layer on the surface of the cell (Nataro, Kaper et al. 1987; Cobeljic, Miljkovic-Selimovic et al. 1996; Dudley, Thomson et al. 2006). However, this method neither distinguishes between pathogenic and non-pathogenic strains nor differentiates typical-EAEC from atypical EAEC and is unsuitable for EAEC outbreaks and limited to research settings, as it requires specialised equipment and labour intensive and experienced personnel (Croxen, Law et al. 2013). A study has shown majority of the HEp-2 positive strains also been positive for antiaggregation protein transporter gene by PCR (Schmidt, Knop et al. 1995). Although in another study, 10% of the EAEC strains verified by HEp-2 assay were negative in the PCR assay which clearly shows difficulty of providing a genotypic definition for EAEC and design specific molecular biological assays for detection (Jenkins, Chart et al. 2006; Weintraub 2007). A cryptic DNA fragment sequence known as "CVD432," or aggregative adherence (AA), from the pAA has been used as an EAEC molecular marker in epidemiological studies and comprises the locus *aat* that encodes an ABC transporter system (Baudry, Savarino et al. 1990; Okeke, Lamikanra et al. 2000). A transcription activator known as "*AggR*," the gene of which lies on pAAs, has been described as the major EAEC virulence regulator for diverse virulence genes (Nataro 2005). Multiple PCR-based assays have been developed to identify the *aggR* gene, and detection of additional virulence genes, such as *aap*, *astA*, and *set1A*, significantly increases the detection of strains associated with causing diarrhoea in U.S. and European patients (Vila, Gene et al. 1998;

Cennimo, Abbas et al. 2009). However, strains that do not carry the *aggR* gene have also been isolated from gastrointestinal outbreaks (Cobeljic, Miljkovic-Selimovic et al. 1996).

1.5.7.2 Polymerase Chain Reaction (PCR) assay to detect virulence (Marker) genes

Recently, amplification by multiplex PCR of either the plasmid-carried gene *aatA* (CVD432) or the chromosomally carried *aaiC* locus is considered sufficient to confirm EAEC in the recently released GEMS, an initiative to comprehensively identify major enteric pathogens rapidly at sites where the diarrheal burden is high (Panchalingam, Antonio et al. 2012).

1.5.7.3 Biofilm assay

Biofilm formation is found useful in screening diagnostic tool when a large number of strains are examined in clinical and epidemiologic studies. All EAEC strains in a study demonstrated an OD570 > 0.2 in the assay, and the incidence of EAEC among the strains with an OD570 > 0.2 was 89.2% (Iwanaga, Song et al. 2002; Wakimoto, Nishi et al. 2004). Furthermore, the test may be available without a spectrophotometer, since a biofilm demonstrating an OD570 > 0.2 is clearly visible. In addition, this assay may contribute to demonstrating of the true incidence of EAEC with and without *AggR* among clinically isolated *E. coli* strains. Of the 28 PCR-positive (AggR and EAST) strains screened for biofilm, 25 (89.2%) demonstrated positive results by microtiter plate method (Kaur, Chakraborti et al. 2010).

1.5.7.4 Serologic-Serotyping assay

Serologic screening was adopted in a study where sera from children (control group) living in an endemic area show no antibody response to Pet but sera from children with diarrhoea caused by EAEC showed high titres of antibody against this toxin (Bellini, Elias et al. 2005). In addition, rabbit anti-Pet sera recognized 50% of the EAEC strains recovered from stools after culture supernatant concentration by immunoblotting (Bellini, Elias et al. 2005). The emergence of EAEC infection in Brazil (Zamboni, Fabbricotti et al. 2004) and the detection complexity of Pet expressing EAEC isolates led to the development of a methodology for Pet detection directly from supernatants of bacterial isolates using a slot blot immunoassay (Taddei, Fasano et al. 2005).

Serotyping - of EAEC is a problem due to their aggregative phenotype, many of the strains auto-agglutinate and is often described in the literature as nontypable or as O-rough. EAEC from German children demonstrated 14 typable isolates and all belonged to different serotypes (Huppertz, Rutkowski et al. 1997). In another study in UK, 97 EAEC strains were serotyped to 40 different O-types. In one of the studies, 93 out of 143 EAEC strains could be serotyped and belonged to as many as 47 different serotypes (Jenkins, Tembo et al. 2006). Serotyping is no longer reliable in the diagnosis of diarrhoeagenic *E. coli* infections.

1.5.7. 5 Newly proposed assay

These diagnostic tests include an enzyme-linked immunosorbent assay (ELISA) for quantitative detection of secretory immunoglobulin A to EAEC (Sutjita, Bouckenooghe et al. 2000) and cytokine response patterns to enteropathogens in which a specific pattern may become a distinguishing pathogen signature (Greenberg, Jiang et al. 2002). More studies and better diagnostic tools are needed to allow for a better understanding of the true epidemiology of EAEC in children.

1.5.7.6 Genome sequencing

The first complete genomic sequence of *E. coli* 042, the prototypical member of the EAEC was performed and then published in 2010 (Chaudhuri, Sebaihia et al. 2010). The study showed the genome of EAEC 042 consisting of a circular chromosome of 5,241,977 bp, one plasmid pAA of 113,346 bp and other major features of the EAEC 042 genome that are serve as template for future diagnostic and intervention strategies for the EAEC pathotypes.

In the detection of Stx2 EAEC strain_whole genome-phylogenesis confirmed strain O104:H4 as an EAEC strain. Alignment of an EAEC O104:H4 isolate TY2482 against the prototype EAEC strain 55989 chromosome ultimately revealed the presence of the large conjugative plasmid pAA which resembled the AAF gene-coding cluster from strain 55989 (Philipson, Bassaganya-Riera et al. 2013). Interestingly, pAA TY2482 encoded for AAF/I rather than the more common AAF/III. The isolate lacked the locus of enterocyte effacement (LEE; responsible for bacterial adherence), intimin adherence factor and a type-III secretion system normally identified in enterohaemmorhagic *E. coli* (EHEC) strains (Rohde, Qin et al. 2011). Since EAEC virulence factors are encoded on plasmids, bacteriophages and genetic pathogenicity islands, the traits are easily transferred to new emerging strains (Brzuszkiewicz, Thurmer et al. 2011; Philipson, Bassaganya-Riera et al. 2013). The phenomenon of genome sequence of TY2492 reveals the ability for Shiga toxin-producing *E. coli* to produce various adhesion mechanisms portraying the ability for pathotypes to overlap and evolve into more virulent strains. Therefore, rapid responses in sequencing efforts during the EAEC O104:H4 outbreak suggests that genomic epidemiology will become a standard molecular strategy to elucidate infectious disease outbreaks (Grad, Lipsitch et al. 2012).

1.5.8 Treatment and Drug Resistance

1.5.8.1 Brief explanation

In children, treatment of bacterial gastroenteritis including EAEC is primarily supportive and directed toward maintaining hydration and electrolyte balance. Antibiotic therapy is rarely indicated and should be deferred until culture results are available. Oral rehydration therapy (ORT) is the preferred treatment for fluid and electrolyte losses caused by diarrhoea in children with mild-to-moderate dehydration. Intravenous hydration is often administered for severe dehydration or when vomiting prevents ORT. Antimicrobial therapy should be used in cases of severe diarrheal disease to reduce the duration of illness, particularly because of its association with persistent diarrhoea in children.

In the case of traveller's diarrhoea (TD) antibiotics are usually recommended but experts in travel medicine discourage the use of absorbable antimicrobial agents for TD prophylaxis, rifaximin, a poorly absorbed antibiotic, has been proposed for prevention of TD (de la Cabada Bauche and Dupont 2011; Estrada-Garcia, Perez-Martinez et al. 2014). However, EAEC infections are often successfully treated with ciprofloxacin and other fluoroquinolones as well as azithromycin, rifaximin, amoxicillin/clavulanic acid and nalidixic acid (Glandt, Adachi et al. 1999; Infante, Ericsson et al. 2004), but there are multiple antibiotic-resistant strains (Okeke and Nataro 2001; Mortensen, Fowlkes et al. 2011). For example, in southern India, EAEC is increasingly resistant to quinolones (Raju and Ballal 2009). In adult patients in the United States, EAEC is susceptible to rifaximin or a single dose of azithromycin with or without loperamide (DuPont 2007; Ericsson, DuPont et al. 2007).

The progressive increase in antibiotic resistance among EAEC strains in developing countries is cause for concern (Estrada-Garcia, Cerna et al. 2005). Several investigators have suggested that lactoferrin may protect infants from gastrointestinal infections, including EAEC, and

might be an alternative treatment for antibiotic resistant EAEC strains (Ochoa and Cleary 2009).

1.5.8.2 Complication of Antibiotics in the Treatment of infections caused by Stx-containing EAEC strains

The detailed explanation provided in Croxen review (Croxen, Law et al. 2013) is as follows. Prior to the 2011 outbreak in northern Germany, there was no standardized treatment for Stx-containing EAEC. During the 2011 German outbreak, 3 children with Stx-associated HUS showed rapid clinical improvement with eculizumab, but result from a subsequent nonrandomised trial with 298 patients were unclear and unreliable (Menne, Kielstein et al. 2012; Hauswaldt, Nitschke et al. 2013). Patients who had no clinical improvement during plasmapheresis and/or were suffering from severe neurological complications were preferentially selected for the trial, leading to a selection bias that complicates the results (Hauswaldt, Nitschke et al. 2013). As eculizumab disrupts the complement cascade, clinicians at the time were required to treat with a prophylactic antibiotic to prevent meningitis (Croxen, Law et al. 2013). In general, antibiotics are normally not recommended for STEC, as they increased the risk for development of HUS by stimulating Stx production. Because of this risk, clinicians treating stx-expressing EAEC strain O104:H4 selected azithromycin, which *in vitro* represses the expression of stx (Bielaszewska, Idelevich et al. 2012). Monitoring of STEC shedding in patients receiving azithromycin showed that these patients were rapidly decolonized (Nitschke, Sayk et al. 2012). Because of this, long-term (>28 days) carriers of STEC O104:H4 were treated with azithromycin, and after a 3-day course of treatment, all 15 were negative for shedding as well as HUS-related symptoms (Nitschke, Sayk et al. 2012). Further studies have since shown that sub-inhibitory concentrations of ciprofloxacin increase Stx production in STEC O104:H4 but that

meropenem, rifaximin, tigecycline, and azithromycin do not (Hauswaldt, Nitschke et al. 2013). Stx production by STEC O157:H7 responds differently to these same antibiotics (Hauswaldt, Nitschke et al. 2013). Nonetheless, the use of azithromycin to eliminate Shiga toxin-containing strains such as O104:H4 from patients is still considered a controversial treatment; if used early in treatment, it is still unclear if it plays a role in the development of HUS and or if used later in treatment, it may actually increase the risk of sudden cardiac death (Seifert and Tarr 2012).

1.5.8.3 Vaccine and other preventive therapy

As EAEC proteins are antigenic, it remains possible that a vaccine could be developed, but as of yet, there is none. However, report from a vaccine study that uses ETEC heat-labile toxin, showed a decreased in the rate of infection and severity of disease caused by ETEC, and despite the presence of EAEC in the placebo groups, the vaccine-treated group had no EAEC detected, suggesting that the vaccine may also exert protection against EAEC (Frech, Dupont et al. 2008). In addition, an *in vitro* study demonstrates treatment with lactoferrin inhibits EAEC enteroadhesion and biofilm formation, however, it is not yet clear whether lactoferrin as a nonantibiotic approach is effective for the treatment and prophylaxis of EAEC but it is a potential though untested nonantibiotic treatment for the prevention of EAEC (Ochoa and Cleary 2009). In addition, IL-8 genotypes may define populations likely to benefit from therapeutic intervention such as prophylactic antibiotics and vaccines (Kaur, Chakraborti et al. 2010).

1.6 **Summary of Gambian site specific GEMS study in which this study is nested**

The ultimate goal of the Global Enteric Multicentre Study (GEMS) was to conduct an investigation that provides data needed to guide development and implementation of enteric vaccines and other public health interventions that can decrease Paediatric morbidity and mortality from diarrhoeal disease. One of the strategies adopted to achieve the goal was to use common standard techniques across the selected seven sites (Mali, The Gambia, Kenya, Mozambique, Bangladesh, India and Pakistan) to identify microbiologic aetiology of diarrhoeal disease. Another important strategy was to characterize the phenotype and genotype distribution of major enteric pathogens. The overall results from this large study were detailed and helpful in developing policies required to achieve the study objectives (Kotloff, Nataro et al. 2013).

The Gambia site specific results were comprehensive and revealing. Following standard and rigorous laboratory methods that included culture, immunologic and molecular techniques (Panchalingam, Antonio et al. 2012) five enteric pathogens were showed to be significantly associated with moderate-to-severe diarrhoea (MSD) among children less than 5 years old in The Gambia, these pathogens include Rotavirus, Shigella, Norovirus, Cryptosporidium and Enterotoxigenic *E. coli* (ETEC-ST) (Kotloff, Nataro et al. 2013). Additional other targeted pathogens were detected but were not significantly associated with moderate-to-severe diarrhoea include Aeromonas, Campylobacter, Salmonella, Typical EPEC, ETEC-LT, EAEC, EHEC, *V. parahaemolyticus*, Giardia, *Entamoeba histolytica*, Adenovirus, Astrovirus and Sapovirus.

The Gambia GEMS results in relation to the three Commonest DEC (EPEC, ETEC and EAEC) showed that only ETEC was significantly associated with MSD whilst EPEC and EAEC were found not to be significantly associated with diarrhoea (table 1.1 & 1.2).

Surprisingly, EAEC, the most frequently detected bacterial pathogens in GEMS study was not significantly associated with diarrhoea among Gambian children. This finding contradicts report from previous studies that have implicated EAEC as a cause of diarrhoea mostly in children from developing countries (Okeke 2009). Furthermore, of the 42 deaths recorded among GEMS study participants, four death occured among diarrhoeal children with sole EAEC infection (figure 1.7). Therefore, we were obligated to do further investigations/characterisations on the EAEC strains in order to reveal virulence factors harboured by the EAEC responsible for diarrhoeal disease among children from rural Gambia using molecular approaches. Also, we explored quantitation of bacterial load and expression of biofilm and characterisation of biofilm producing genes.

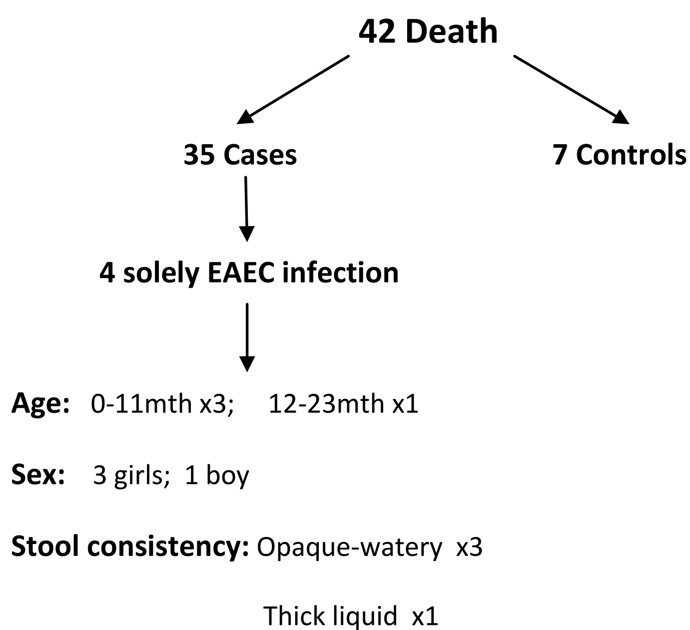
Table 1.1: Distribution of Diarrhoeagenic *Escherichia coli* pathotypes from Diarrhoeal and Non-Diarrhoeal Children enrolled in GEMS Gambia site

DEC Pathotype	Case (N = 519) No. (%)	Control (N = 746) No. (%)	Total (N = 1265) No. (%)	OR (95% CI)	P - value
EAEC only	241 (46.44)	385 (51.61)	626 (49.49)	0.81(0.64-1.02)	0.07
ETEC only	141 (27.17)	139 (18.63)	280 (22.13)	1.63(1.24-2.15)	0.0003
EPEC only	87 (16.76)	124 (16.62)	211 (16.68)	1.01(0.74-1.38)	0.94
EAEC+ETEC	21 (4.05)	36 (4.83)	57 (4.51)	0.83(0.46-1.48)	0.51
EAEC+EPEC	14 (2.70)	37 (4.96)	51 (4.03)	0.53(0.26-1.02)	0.04
ETEC+EPEC	13 (2.50)	20 (2.68)	33 (2.61)	0.93(0.42-1.99)	0.84
EAEC+ETEC+EPEC	2 (0.39)	5 (0.67)	7 (0.55)	0.57(0.05-3.52)	0.50

Table 1.2: Distribution of Diarrhoegenic *E. coli* pathotypes from Diarrhoeal and Non-Diarrhoeal Children enrolled in GEMS Gambia site

DEC Pathotype	Case (N = 571) No. (%)	Control (N = 849) No. (%)	Total (N = 1420) No. (%)	OR (95% CI)	P - value
EAEC	278 (48.69)	463 (54.53)	741 (52.18)	0.79(0.63-0.98)	0.03
ETEC	177 (31.00)	200 (23.56)	377 (26.55)	1.46(1.14-1.86)	0.001
EPEC	116 (20.32)	186 (21.91)	302 (21.27)	0.90(0.69-1.19)	0.47

Figure 1.7: Flow chart showing detail of death in GEMS study



Chapter 2: Research question, Hypotheses, Aims and Objectives

The motivation to perform this study of further characterisation of Gambian EAEC strains from the GEMS case-control study was based on the GEMS result which, against to initial hypotheses, showed that EAEC - the most frequently isolated bacteria in the study - was not associated with diarrhoea among Gambian children <5 years of age. Such results contradict previous studies which have showed EAEC as a cause of persistent diarrhoea and acute gastroenteritis among children from developing countries including The Gambia. Therefore, I planned to conduct further investigation to characterise the EAEC strains isolated from diarrhoea and non-diarrhoea study children to reveal virulence factors harboured by the EAEC responsible for diarrhoeal disease.

2.1 Research questions

1. What are the specific virulence factors harbours by EAEC that are responsible for diarrhoeal disease among children from rural Gambia?
2. Can bacterial load assay confirm EAEC as a true cause of diarrhoea among children from rural Gambian?
3. What role does biofilm play in the diarrhoea caused by EAEC among Gambian children?
4. Can antimicrobial resistant EAEC be associated with diarrhoea among children?

2.2 **Hypotheses**

1. Combination of virulence genes found in EAEC cause moderate-to-diarrhoea among rural Gambian children
2. High bacterial-load in EAEC associated with diarrhoea among rural Gambian children
3. Biofilm producing EAEC cause diarrhoea among Gambian children
4. Presence of *aggR* gene combine with one or more of it regulatory genes in EAEC producing-biofilm can contribute to diarrhoeal illness among Gambia children
5. Multiple antimicrobial resistant EAEC strains may be associated with diarrhoea

2.3 **Aims**

- The aim of this study is to determine appropriate diagnostic tool that can be used to detect EAEC that cause diarrhoea disease among children and to characterise and evaluate the role of EAEC in diarrhoeal illness in children from rural The Gambia.

2.4 Specific objectives

- Determine the prevalence of EAEC known virulence genes among Gambian children with MSD and non-MSD from whom EAEC was isolated.
- Determine a TaqMan-QPCR cut off threshold cycle to compare EAEC bacterial load MSD and non-MSD as a tool to diagnose EAEC diarrhoea
- Evaluate phenotypic biofilm screening assays that include test-tube (TT), congo-red-agar (CRA) and tissue culture plate (TCP) to establish the appropriate screening method that identify infectious EAEC among diarrhoeal children.
- Determine biofilm producing genes in EAEC strains from MSD and non-MSD children
- Determine the association of antimicrobial resistant EAEC strains with diarrhoea.
- Determine the presence or absence of antimicrobial resistance genes in the EAEC isolates and compare with phenotypic resistance result.
- Access the role of sequence types in children diarrhoea caused by EAEC

Chapter 3: Materials and Methods

3.1 Laboratory settings at MRCG Basse field station and Fajara

The MRCG Basse laboratory at Upper River Region (URR) of Gambia is well known for its reputation in the diagnosis of enteric pathogens. The laboratory serves as a training centre for laboratory technicians working in the government health centres within URR and Central River Region (CRR). The laboratory has served many experimental and observational studies for over 30 years. These studies include Pneumococcal Vaccine Trial (PVT), Meningococcal Vaccine Trial (MVT), Pneumo Aetiology of Child Health (PARCH), Global Enteric Multicentre Study (GEMS), Pneumo case-control study (PCS), Malaria In Pregnancy (MIP), Entomology Study (ES), Pneumococcal Surveillance Programme (PSP), the ongoing Vaccine Impact on Diarrhoea in Africa (VIDA) and others. In addition, the laboratory subscribed to external quality control assurance (EQA) oneworld Canada and GCLP accredited.

The Fajara laboratory where the bacterial load (Taqman-QPCR) aspect of this study was conducted is internationally recognised as an excellent research centre in West-Africa. It houses multiple research laboratories and a World Health Organisation reference laboratory; molecular laboratories are fortified with state of the art equipment.

3.2 Study population:

The study participants were children less than 5 years of age who have moderate-to-severe diarrhoea (MSD) and non-diarrhoeal children belonging to a censused population of 35,000 children of URR. The enrolled MSD children are those seeking care at referral Health Centre Basse and five sentinel health centres (SHCs), while, the enrolled 1-3 matched controls are from the community for each index case. Having determined the eligibility criteria for cases and controls, enrolment was subject to obtaining an informed verbal and documented consent from parent or guardian of children.

3.3 Case Definition of Moderate-to-Severe Diarrhoea (MSD)

This is described by Farag *et al* as "a child with diarrhoea (≥3 abnormally loose stools) within the previous 24 hours with onset within the previous 7 days, following at least 7 days without diarrhoea, and accompanied by evidence of clinically significant dehydration (loss of skin turgor, sunken eyes, or a decision by the clinician to administer intravenous fluids), dysentery (blood in the stool), or a clinical decision to hospitalize the child" (Farag, Nasrin et al. 2012).

3.4 Definition of Control

A child without diarrhoea within 14 days of presentation of the index case; and of the same sex, within the same age strata and from the same village or neighbourhood (Farag, Nasrin et al. 2012).

3.5 Study approval

Approval for this study was given by MRCG Scientific Coordinating Committee (SCC) followed by ethical approval by Gambia-Government/MRCG joint ethics committee (Appendix A – SCC and Ethics letter).

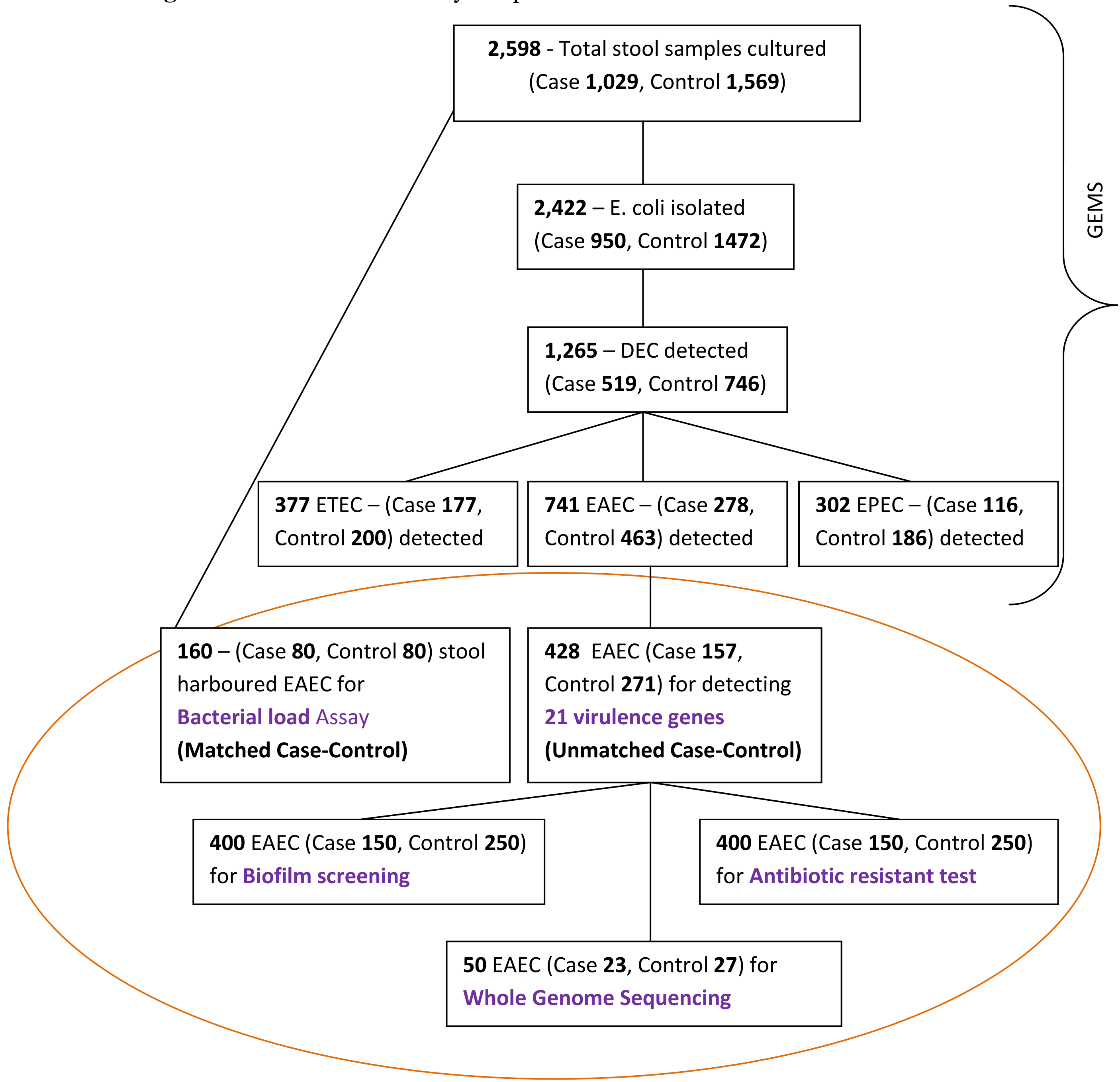
3.6 Good Clinical Laboratory Practice (GCLP)

This study was conducted according to the internationally recognised GCLP guideline – DAID, MRC Good Research Practice (MGRP) and conformed to the International Conference on Harmonisation Good Clinical Practice (ICH-GCP). To ensure principles of GCLP are complied with MRCG conducted series of training on GCLP/GCP that were attended by the laboratory staff involved in this study (Appendix B – GCP/GCLP certificates).

3.7 Sample selection

Being a retrospective study, 428 EAEC strains were randomly selected from a total of 741 EAEC samples. A flow chart showing how the sampling was achieved is detailed below (figure 3.1). The selected EAEC strains were sub-cultured and amplified for re-identification and detection.

Figure 3.1: Flow chart of study samples



3.8 Detection of EAEC

In GEMS study, the detection of EAEC was by culturing faecal sample on selective media and isolating suspect colonies, identifying the colonies as *E. coli* and detecting them as Enteroaggregative *E. coli* (EAEC) by polymerase chain reaction (PCR) an amplification method.

3.9 Bacteriology

Stool samples received in transport media Cary Blair (CB) and Buffered Glycerol Saline (BGS) were aseptically streaked on to the multipurpose and selective solid media that includes McConkey, Xylose lysine deoxycholate, Ryan, Thiosulfate Citrate Bile Salt Sucrose (TCBS)/Alkaline-peptone-water (APW) and Campylobacter Blood agar for the isolation of *E. coli*, Shigella /Salmonella species, Aeromonas species, Vibrio species and Campylobacter species respectively. Also, other opportunistic bacterial flora that cause diarrhoea were isolated and identified. Following overnight incubation at 37°C cultured agar plates were examined for aforementioned bacterial pathogens except Campylobacter agar plate that is examined after 48 hours incubation at 42°C. Specifically from MacConkey agar plate, three suspected colonies of *E. coli* (often lactose fermenting) were purified and identified as *E. coli* by performing gram stain reaction, detecting release of indole using Kovac’s reagent (figure 3.2) and conducting other enzymatic and fermentation test using biochemical reagent kit Analytical Profile Index (API) 20 E (BioMeriux 09567D).

3.9.1 Gram stain reaction

Principle: This stain classified bacteria as Gram-positive or Gram-negative depending on whether the bacteria retain the stain crystal violet (Gram-positive) or are decolourised and take up the red counter stain (Gram-negative).

The four steps involved are **crystal violet**, the primary stain, followed by **Lugol's iodine**, which acts as a mordant by forming a crystal violet iodine complex, then **alcohol**, which decolourises and lastly, **neutral red** or **safranin**, the counter stain.

Procedure

Air or heat dried smear

Cover with crystal violet stain for 30-60 seconds

Rapidly wash off the stain with clean water

Tip off all the water, and cover the smear with Lugol'siodine for 30-60 seconds

Wash off the iodine with clean water

Decolourise rapidly with acetone-alcohol

Wash immediately with clean water

Cover the smear with neutral red stain for 2 minutes

Wipe clean the back of the slide, and place in a draining rack to air dry

Examine microscopically with X100 objective lens

Interpretation of results:

Gram-negative bacteria……………………………………….Pale to dark red

Gram-positive bacteria………………………………………...Dark purple

3.9.2 Indole production:

Principle: The ability of a bacterial agent to split amino acid tryptophan to form indole compound. The enzyme tryptophanase hydrolysed amino-acid tryptophan to produce three possible end products that include indole, pyruvate and ammonium. The production of indole is detected by Kovac's reagent which contains 4-(p)-dimethylaminobenzaldehyde that reacts with the indole to produce a red colour in the surface layer (meniscus) of broth medium (figure 3.2)

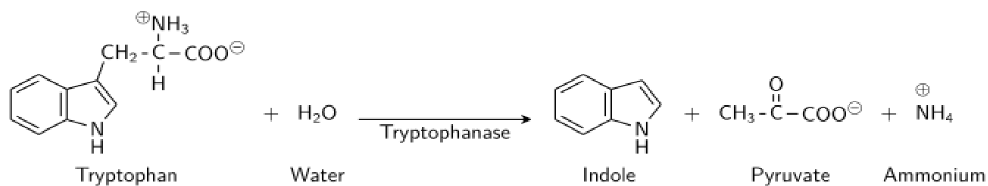
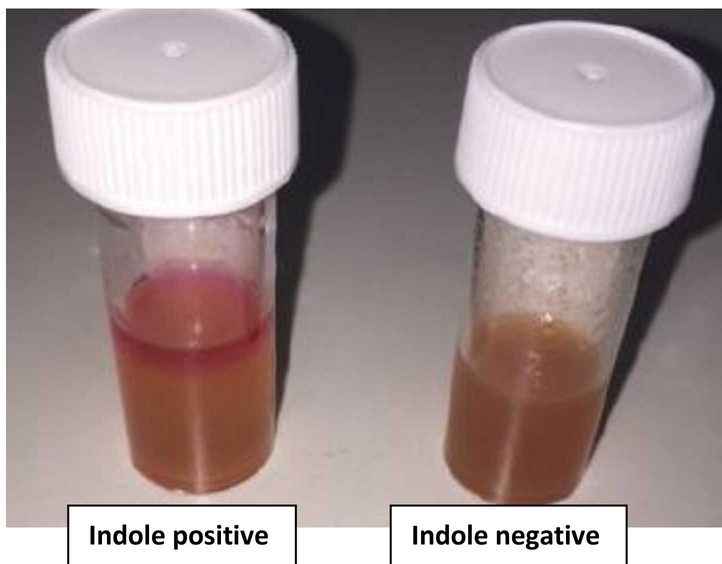


Figure 3.2: Showing positive and negative indole tests



3.9.3 Biochemical identification kit – Analytical Profile Index 20 Enteric (API 20E)

Summary: The standardised identification system for Enterobacteriaceae and other non-fastidious Gram-negative bacilli that include *E. coli* is the Analytical Profile Index (API) 20 E (bioMerieux, catalogue number 20120). The system has 21 miniaturised biochemical tests and a database was used for the biochemical identification of *E. coli* isolates.

Principle: The API 20 E strip consists of 20 microtubes containing dehydrated substrates. These tests are inoculated with a test bacterial suspension that reconstitutes the media. During incubation, metabolism produces colour change, which are either spontaneous or revealed by the addition of reagents.

The reactions are read according to the reading table and the identification is obtained by referring to the Analytical Profile Index or using the identification software.

Procedure

Prepare an incubation box (tray and lid) and distribute about 5 ml of distilled water into the honey-combed wells of the tray to create humid atmosphere

Record the strain reference on the elongated flap of the tray

Remove the strip from its packaging

Place the strip in the incubation box and cover with the lid until ready for inoculation

Prepare a sterile tube (bijou bottle) to contain 5ml sterile saline or distilled water without additives

Using a sterile pipette, remove a single well-isolated test colony from a cultured plate of 18-24 hours old

Carefully emulsified to achieve a homogenous bacterial suspension, which is used immediately to inoculate the strip

Using a sterile pipette, fill both tube and cupule of the tests citrate, VogePokers and gelatine with bacterial suspension

Fill only the tube (and not the cupule) of the other tests

Create anaerobiosis in the tests Arginine Dehydrolase (ADH), Lysine Decarboxylase (LDC), Ornithine Decarboxylase (ODC), Hydrogen sulphide (H_2S) and Urease by over laying with mineral oil

Close the incubation box

Incubate at 36^oC (+/- 2^oC) for 18-24 hours.

After the incubation period, and adding the require reagent to the required tests, read the strip by referring to the Reading Table.

Interpretation of results

Obtain the identification with the numerical profile as follow;

On the result sheet, the tests are separated into groups of 3 and a value 1,2, and 4 is indicated for each. By adding together the values corresponding to positive reactions within each group, a 7-digit number is obtained for the 20 tests of the API 20 E strip. The oxidase reaction constitutes the 21st test and has a value of 4 if it is positive.

Identification is then performed using the database and or by looking up the numerical profiles in the list of profiles.

3.10 Isolation of Genomic DNA

A confirmed and pure colony of *E. coli* was purified on a non-inhibitory Nutrient agar medium. Following 24 hrs incubation at 37^oC, the colonies were harvested and suspended with 500 µl RNAse and DNAse free molecular graded H_2O in 1.8 ml cryo-tube and boil for 20 minutes in a water bath and swiftly cool on ice. The heat-treated bacterial suspension is centrifuge at 3000 rpm for 10 minutes and 5µl of the supanantant was used as DNA template in the PCR.

3.11 Molecular detection of EAEC target *aatA* and *aaiC* genes

In this study we performed monoplex PCR on each isolates that has initially showed presence of EAEC in GEMS *E. coli* multiplex PCR protocol. The target sought for EAEC are the EAEC plasmid-encoded gene *aat*A (primer CVD432F – sequence 5'-CTGGCGAAAGACTGTATCAT-3'and primer CVD432R – sequence 5'-CAATGTATAGAAATCCGCTGTT-3') and the EAEC chromosomally encoded *aai*C (primer AAIC F – sequence 5'-ATTGTCCTCAGGCATTTCAC-3' and primer AAIC R - sequence 5'-ACGACACCCCTGATAAACAA-3', these two loci are known virulence determinants.

PCR assay was performed in a final reaction volume of 20µl, which contained 2.5µl buffer (2mM MgCl2), 10.65µl H_2O, 2µl dNTP (1.25mM), 0.4µl each primer (20pmol/µL) forward

and reverse for marker gene *aat*A and *aai*C, 0.25µL *Taq* polymerase and 3µl of boiled bacterial lysate as the DNA template. PCR were performed using thermocycler (TECHNE Flexigen, Model FFG02FSD, Serial 11733 -1) to achieve preheat at 96°C for 4 minutes, denaturation at 95°C for 20 seconds, annealing at 57°C for 20 seconds, elongation at 72°C for 1 minute and run for 35 cycles and final extension at 72°C for 7 minutes. *E. coli* strain 042 was used as control for EAEC (*aaiC* and *aatA*) strain. Distilled water was used as negative control. PCR products were analysed on a 2% agarose gel containing ethidium bromide (0.1 µl ml-1 in 1 x TBE buffer) and visualise on the 2% (w/v) agarose gel under ultral-violet (UV) radiation. The gel image was captured digitally with a gel documentation system.

3.12 PCR amplification to detect 21 EAEC virulence genes

EAEC colonies were investigated for the presence of the twenty-one putative virulence genes using four multiplex PCR as previously described [28]. The 21 genes were grouped into four. On each group multiplex-PCR was performed. On group 1 (*sat, sepA, pic, sigA*, *pet* and *astA*), multiplex-PCR master mix was achieved using Qiagen kit (Catalogue number 206143) following the manufacturer's instructions. Multiplex-PCR assay was performed in a final reaction volume of 25µl that consists, 12.5 µL mastemix (MM), 2.5µl Q-solution, 6µl primer(MM), 2.5µl H_2O and 1.5µl DNA template. PCR reaction cycles were as follows: 15 minutes preheating at 95°C at the start, 50 seconds denaturation at 94°C, annealing for 1.5 minutes and extension at 72°C for 1.5 minutes with 35 cycles returning to step 2. The final extension was 10 minutes at 72°C (figure 3.3).

On group 2 (*aatA, aggR, aaiC, aaP* and ORF3), group 3 (*aafC, agg3/4C*, *agg3A, aafA, aggA, agg4A*) and group 4 (*air, capU, ailA* and ORF61) Fementers kit (Catalogue # K0171) was used for the PCR master mix (2X) following the manufacturer's instructions. Multiplex-PCR assay was achieved in a final reaction volume of 25µl that compose of 12.5µl (MM), 1µl (25mM Magnesium chloride ($MgCl_2$), 5µl primer (MM), 5µl of H_2O and 1.5µl DNA template. PCR reaction cycles were achieved as follows: 2 minutes preheating at 95°C at the start, 50 seconds denaturation at 94°C, annealing at 57°C (58°C for Group 3&4) for 1.5 minutes and extension at 72°C for 1.5 minutes with 35 cycles returning to step 2. The final extension was 10 minutes at 72°C.

Amplifications were performed using Thermocycler (TECHNE Flexigen, Model FFG02FSD, Serial 11733 -1, manufactured in USA) Amplified PCR products were analysed on a 2% agarose gel containing ethidium bromide (0.1 m ml-1 in 1 x TBE buffer and visualise on the 2% (w/v) agarose gel under ultral-violet (UV) radiation. The gel images were captured digitally with a gel documentation system.

The *E. coli* strains used as controls for detection of the target genes are; C1010-00 (*sat*, *sepA*, *agg3/4C* & *agg4A*) , JM221 (*sat* & *aggA*), 042 (*pic*, *pet*, *astA*, *aatA*, *aggR*, *aaiC*, *aap*, *ORF3*, *aafC*, *aaFA*, *air*, *capU* & *eilA*), 55989 (*sigA*, *agg3A/4C*, *agg3A*), 63 (*sigA*, *agg3/4C* & *agg4A*) and 17-2 (*aggA*) (Boisen, Scheutz et al. 2012). GIBCO distilled water (DNase/RNase free, Catalogue no. 10977-035) was used as negative control.

Gel electrophoresis for virulence genes

Amplified PCR products were analysed on a 2% agarose gel containing ethidium bromide (0.1 m ml-1 in 1 x TBE buffer and visualise on the 2% (w/v) agarose gel under ultral-violet (UV) radiation. The gel images were captured digitally (figure 3.4) with a gel documentation system.

Table 3.1: Primer sequence for the detection of 21 EAEC Virulence associated genes

Target Gene	Function/Description of encoded proteins (Plasmid (P) / Chromosome (C)	Primer Sequence (5' – 3') Primer Forward Primer Reverse	Control Strains	PCR Product base –pair (bp)	Reference (GenBank Accession No.)
sat	Secreted autotransporter toxin (C)	TCAGAAGCTCAGCGAATCATTG CCATTATCACCAGTAAAACGCACC	C1010 JM221	930	**AE014075**
sepA	Shigella extracellular proteins A (P)	GCAGTGGAAATATGATGCGGC TTGTTCAGATCGGAGAAGAACG	C1010	794	Z48219
pic	Protein involve in Intestinal Colonisation (C)	ACTGGATCTTAAGGCTCAGGAT GACTTAATGTCACTGTTCAGCG	042	572	AF097644
sigA	IgA protease-like homolog (C)	CCGACTTCTCACTTTCTCCCG CCATCCAGCTGCATAGTGTTTG	63, 55989	430	NC_004337
pet	plasmid-encoded enterotoxin (P)	GGCACAGAATAAAGGGGTGTTT CCTCTTGTTTCCACGACATAC	042	302	AF056581
astA	(EAST1) EAEC heat-stable enterotoxin 1 (P)	ATGCCATCAACACAGTATAT GCGAGTGACGGCTTTGTAGT	042	110	L11241
aatA	Dispersin transporter protein (P)	CTGGCRAAAGACTGTATCAT CAGCTAATAATGTATAGAAATCCGC TGT	042	642	AY351860
aggR	AAF/I and AAF/II transcriptional activator (P)	GCAATCAGATTAARCAGCGATACA CATTCTTGATTGCATAAGGATCTGG	042	426	Z18751
aaiC	AaiC, secreted protein (C)	TGGTGACTACTTTGATGGACATTGT GACACTCTCTTCTGGGGTAAACGA	042	313	....
aap	Anti-aggregation protein (dispersin) (P)	GGACCCGTCCCAATGTATAA CCATTCGGTTAGAGCACGAT	042	250	Z32523
ORF3	Cryptic protein (P)	CAGCAACCATCGCATTTCTA CGCATCTTTCAATACCTCCA	042	121	....
aafC	Aggregative adherence fimbriae C - Usher, AAF/II assembly unit (P)	ACAGCCTGCGGTCAAAAGC GCTTACGGGTACGAGTTTTACGG	042	491	AF114828
agg3/4C	Usher, AAF/III-IV assembly unit (P)	TTCTCAGTTAACTGGACACGCAAT TTAATTGGTTACGCAATCGCAAT TCTGACCAAATGTTATACCTTCAYT ATG	C1010 55989, 63	409	AF411067 AB255435 EU637023
agg3A	AAF/III fimbrial subunit (P)	CCAGTTATTACAGGGTAACAAGGG AA TTGGTCTGGAATAACAACTTGAACG	55989	370	AF411067
aafA	Aggregative adherence fimbriae A – AAF/II fimbrial subunit (P)	CTACTTTATTATCAAGTGGAGCCGC TA GGAGAGGCCAGAGTGAATCCTG	042	289	AF012835
aggA	AAF/I fimbrial subunit (P)	TCTATCTRGGGGGGCTAACGCT ACCTGTTCCCCATAACCAGACC	JM221, 17-2	220	Y18149 AY344586
agg4A	AAF/IV fimbrial subunit (P)	TGAGTTGTGGGGCTAYCTGGA CACCATAAGCCGCCAAATAAGC	C1010, 63	164	EU637023
air	Enteroaggregative immunoglobulin repeat protein (C)	TTATCCTGGTCTGTCTCAAT GGTTAAATCGCTGGTTTCTT	042	600	....
capU	Hexosyltransferase homolog (P)	CAGGCTGTTGCTCAAATGAA GTTCGACATCCTTCCTGCTC	042	395	AF134403
eilA	Salmonella HilA homolog (C)	AGGTCTGGAGCGCGAGTGTT GTAAAACGGTATCCACGACC	042	248	....
ORF61	Plasmid encoded haemolysin (P)	AGCTCTGGAAACTGGCCTCT AACCGTCCTGATTTCTGCTT	042	108	....

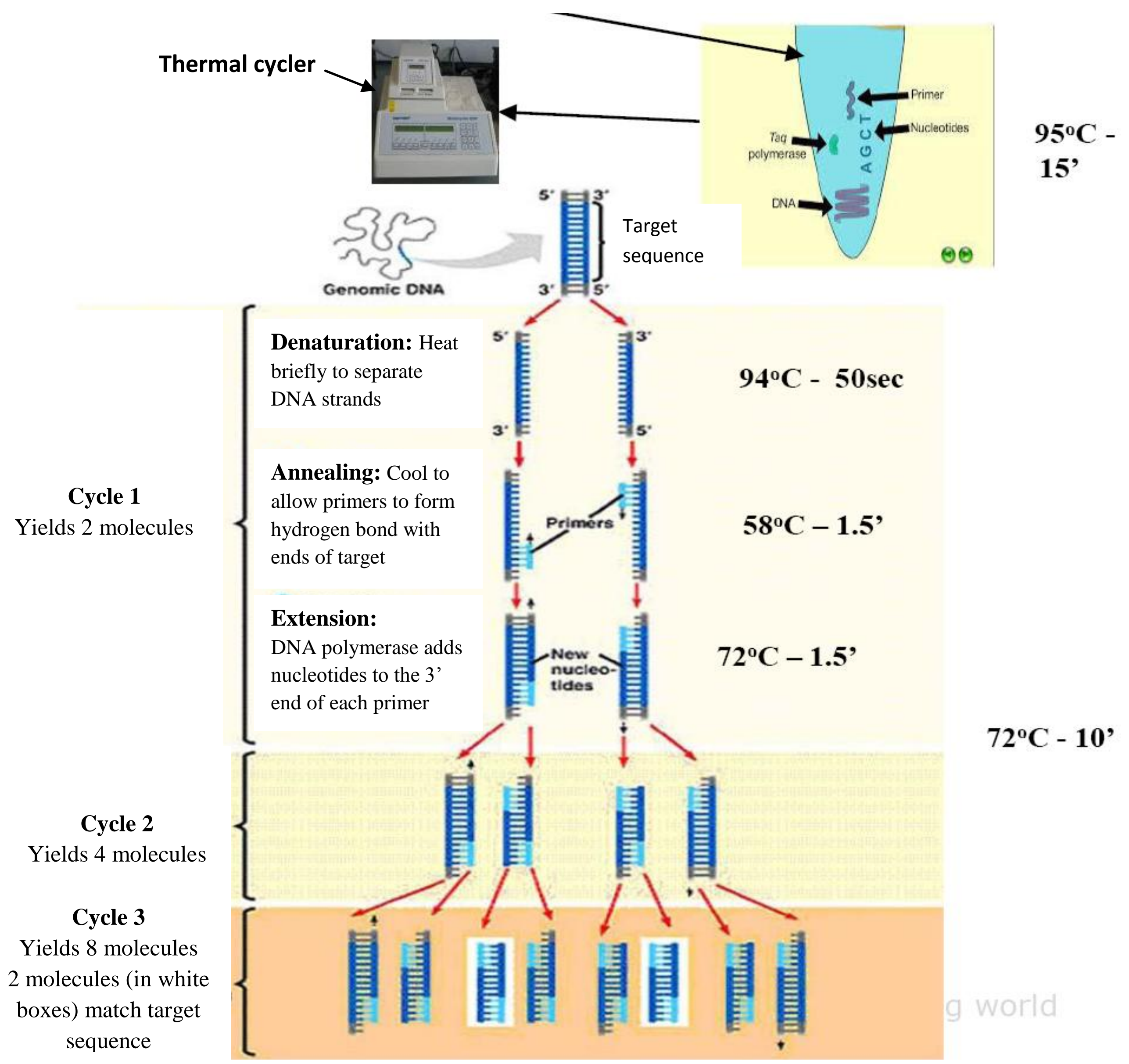
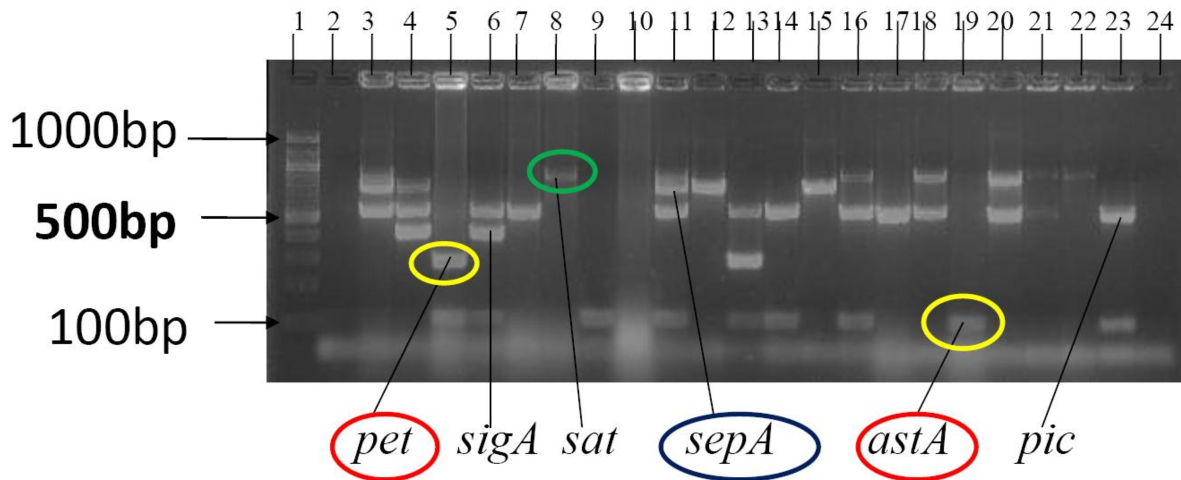


Figure 3.3: PCR reaction cycle pattern used in the amplification of the 21 virulence genes

Figure 3.4: Gel image of group 1 (*sat, sepA, pic, sigA*, *pet* and *astA*) of the 21 putative virulence genes



Lane	Test	Base pair	Genes		Lane	Test	Base pair	Genes
1.	**Ladder**	100-1200bp	**100 - 1200**		**13.**	103434	110, 302, 572	***astA, pet, pic***
2.	**Negative**	No band	No gene to amplify		**14.**	103448	110, 572	***astA, pic***
3.	**C1010**	572, 794, 930	*pic, sepA, sat*		**15.**	103449	794	***sepA***
4.	**63**	430, 572, 794	*sigA, pic, sepA*		**16.**	103451	110, 572, 930	***astA, pic, sat***
5.	**042**	110, 302	*astA, pet*		**17.**	103454	572	***pic***
6.	**55989**	110, 430, 572	*astA, sigA, pic*		**18.**	103490	572, 930	***pic, sat***
7.	**JM221**	572	*pic*		**19.**	103494	110	***astA***
8.	103396	930	*sat*		**20.**	103506	572, 930	***pic, sat***
9.	103399	110	*astA*		**21.**	103509	572, 930	***pic, sat***
10.	103400	No band	Nil		**22.**	103526	930	***sat***
11.	103412	110, 572, 794	*astA, pic, sepA*		**23.**	103527	110, 572	***astA, pic***
12.	103412	794	*sepA*		**24.**	103529	No band	

3.13 Statistical analyses

Bivariate analysis was applied to compare prevalence of virulence factors between cases and controls in different age group using STATA 12 reporting Odds ratios (OR) with 95% confidence intervals (CI). A two-sided p-value < 0.05 was considered statistically significant

Additionally, we used Classification and Regression Tree (CART) pro-Version 6.0 (Salford systems) software to input 21 factors of interest as binary (present/absent) independent variables. Case-control status was input as the binary dependent outcome variable.

3.14 Significance of Combinations of EAEC Genes:

We generated a virulence factor score (VFS), representing the collective number of virulence loci present in each strain. To consider the combinations factors, we employed CART analysis (figure 4.1 and figure 4.2), which builds a model in stepwise fashion to yield the combination of factors most strongly associated with the queried outcome. Each branch of a CART output tree ends in a terminal "node"; each observation falls into exactly 1 terminal node; and each terminal node is uniquely defined by a set of rules, such as having or not having a certain factor.

We considered all genotypic and phenotypic assays performed and considered the association with case status versus control status (figure 4.1 and figure 4.2).

3.15 Quantification of Bacterial load for EAEC

3.15.1 DNA extraction from cultured EAEC 042 strain

Genomic DNA was extracted from a pure culture of EAEC 042, grown overnight at 37°C , and extracted with QIAmp mini kit (Qiagen London, UK). The DNA concentration quantified using nanodrop was performed in duplicate (56 ng and 54 ng) and the average yielded 55ng, the EAEC 042 genome size is 5,355323 base pairs (b)p (5.3Kb). The mass of the genome was calculated by inserting the bacterial genome size value in the formula M=[n][1.096e-21g/bp], where n=genome size, M=mass and e-21 = x10^{-21} (AppliedBiosystem 2003). The obtained mass 5.5e-15g was converted to picogram (pg) resulting 0.0055pg that was used to obtain the final concentration (C_2) of 13750 pg/µl of DNA, the genomic DNA concentration (55ng or 55000pg/ µl) was C_1 and the final volume (V_2) was 100ul. So the resultant V_1(25µl) stock EAEC 042 genomic DNA was made up to 100ul by 75µl of diluent (nuclease free water) and ten-fold serial dilutions were prepared to determine the detection limit for the standard control. Please see detail calculation in Appendix C.

3.15.2 Total DNA extraction directly from stool samples

The extraction of DNA from stool was achieved with the use of QIAamp DNA stool Mini Kit (catalogue no. 51504, Qiagen, Valencia, CA, USA). Isolating DNA requires that, the faecal sample undergoes a lysate preparation process and includes mechanical disruption by bead beating. In brief, 370 mg of 0.1 mm glass beads, acid-washed 212-300 µM (50-70 U.S. sieve) (Sigma G1277-500G) was dispensed in 180-220 mg of each stool sample; 1400ul lysis buffer (previously incubated at 70°C for 15 minutes to ensured precipitate fully dissolved) was added along with 1µl of Phocine Herpesvirus (PhHV) as extrinsic control to evaluate DNA

extraction and amplification efficiency. Bead beat at maximum speed for 3 minutes using BioSpec Mini-Beadbeater (BioSpec 693, Mini-Beadbeater-8 USA) for vigorous homogenization of samples. The cell slurry or the mixture was incubated at 100°C for 5 minutes. Following votexing and centrifuging, 1.2 ml of the supernatant was collected into a separate 2 ml microcentrifuge tube and 1 inibitEX Tablet (supplied in the QIAamp DNA stool Mini Kit) added to the supernatant to absorbed inhibitors. Following 1 minute incubation at room temperature and 3 minute centrifugation supernatant was pipetted into a new 2 ml microcentrifuge tube and pellet discarded. 400µl of the supernatant is added to the 30ul proteinase K, followed with the addition of AL buffer and thorough votexing the mixture was incubated at 70°C for 10 minutes. Following brief centrifuging 400µl ethanol (96-100%) was added to the lysate. Six-hundred microlitre (600µl) of the lysate was applied to the QIAamp spin column and processed and elute to obtained total DNA. DNA is then stored at -80 for testing. Each day extractions are performed a blank is included through the complete protocol and later assayed to rule out contamination during the extraction process.

3.15.2a In a order to develope a well designed primers and probes which are a prerequisite for successful qPCR we adhered to the rules that guide primers and probe design. We first designed the probes by following the guidelines probe design which are;

Use of a well recognised probe design software ‘Primer-Blast’, choosing the probe length for *aaiC* – 22 bases with GC content 31.8% and probe length for *aatA* – 24 bases with GC content 41.7%. Consideration during the process was to ensured that the melting temperature (Tm) 62°C is 5°C higher than the melting temperature (Tm) of the primers and that the 5’ end is couple to a T and not 5’ end G this will help avoid the quenching of many flurophores that include FAM.

Also, we designed primers by following the guidelines for primer design which are;

Use of a well recognised software tool called ‘Primer-Blast’, choosing the primers length and GC% (*aaiC* forward – 20 bases, GC – 45% & *aaiC* reverse 20 bases, GC 45%) and (*aatA* forword 20 bases, GC 45% & *aatA* reverse 22 bases, GC 36.4%), and the Tm of 56°C-70°C was selected, and finally the designed primers and probes were sent to ‘microbiom international’ to develop and prepare and then sent back to me for use.

3.15.3 qPCR amplification for the 160 total DNA extractions

TaqMan-based real-time polymerase chain reaction (qPCR) assay for quantification of TaqMan-qPCR was performed on 160 DNA samples using the Bio-Rad CFX Manager software, Russian Edition that works with the UK English setting of the Windows operating system to provide a localized environment. Individual qPCR reactions consisted of a total volume of 25 µl/reaction that includes 12.5 µl of Bio-Rad iQTM multiplex Powermix (UK), 10.575 µl nuclease free water, 0.1 µl of primer forward and reverse for target genes *aaiC* and *aatA*, 0.2 ul of primer forward and reverse for the PhHV, 0.05 µl, 0.025 µl, and 0.05µl were used for *aaiC*, *aatA* and PhHV probe respectively and 1 µl of sample DNA was added. Also, positive and negative reactions are set up in the process. Detail of primer sequence and probe used for the detection and quantitation of *aaiC*, *aatA* and *PhHV* genes are shown in table 3.2 The primers and probe from metabion international AG (Lena-Christ-Strasse-44/I, D-82152 Martinsried/Deutschland) were diluted and used following the manufacturer guidelines. Hot start of 95°C for 3 minutes was used as initial denaturation and Taq activation. PCR amplification and target detection were performed for 40 cycles of Denaturation at 96°C for 15 seconds, annealing, at 60°C for 30 seconds extension at 72°C for 40 seconds (Table 3.2a). Gene copies were determined by absolute quantification using standard curve fitted for 96-

well plate. The standard curve was constructed from EAEC 042, R^2 values of a linear model fit to the standard curves of cycle threshold (C_t) versus log dilution of DNA in the standard ranged from 1.00 to 0.988 (figure 3.6 & 3.7).

Table 3.2: Oligonucleotide sequence of primers and probe used

Gene	Reaction	Nucleotide sequence (5'-3')	GenBank Sequence #	Reference
aaiC F aaiC R	Forward Reverse	5'-ATT GTC CTC AGG CAT TTC AC-3'→ 5'-ACG ACA CCC CTG ATA AAC AA-3'←	FN554766.1	Boisen 2008
aatA F aatA R	Forward Reverse	5'-CTG GCG AAA GAC TGT ATC AT-3'→ 5'-TTT TGC TTC ATA AGC CGA TAG A-3'←	AY351860	Boisen 2008
PhHV-gB F PhHV-gB R	Forward Reverse	5'-GGGCGAATCACAGATTGAATC-3'→ 5'-GCGGTTCCAAACGTACCAA-3'←	S81228.1	
aaiC Probe		5'-TAG TGC ATA CTC ATC ATT TAA G-3'→	FN554766.1	
aatA Probe		5'-TGG TTC TCA TCT ATT ACA GAC AGC-3'→	AY351860	
PhHV Probe **Quasar 670**		5'-TATGTGTCCGCCACCATCT-3'→	S81228.1	

Key: → Direction of synthesis (forword), ← Direction of synthesis (reverse)

Table 3.2a: Showing amplification run information for the TaqMan assay

40 cycles	**Denaturation**	**96°C for 15 seconds**
	Annealing	**60°C for 30 seconds**
	Extension	**72°C for 40 seconds**

3.15.4 Amplification efficiency and limit of detection (LoD) or Quantitation of bacterial load

Standard curves were constructed using known quantities of genomic DNA (serial dilution 10^7, 10^6, 10^5, 10^4, 10^3, 10^2, 10^1 and 10^0) extracted from EAEC 042 strain. The standard curve was performed in duplicate, and analyses were achieved by plotting the threshold cycles (C_T) against the corresponding log input quantity DNA determining the detection limit of the assay. For comparison of PCR amplification efficiencies and detection sensitivities among sample assays, slopes of the standard curves were calculated by a linear regression analysis with Bio-Rad detection system (figure 3.6 and figure 3.7). Quantification cycles (Cqs) are the PCR cycle values at which the fluorescence from amplification exceeds background that acts as an inverse metric of quantity of nucleic acid. By using the standard curves constructed in this study, the highest concentration of bacteria load is detected at Log Starting Quantity 6.6 at Cq 16, and the lowest concentration of the bacteria is detected at Log Starting Quantity 1.6 at Cq 38 (figure 3.7a). In order to determine the low and high bacterial load, a cut-off was obtained by constructing a slope from the point where Cq 16 intercept with the 1.6 log starting quantity to the point it intercept the slope of the standard curve. At the later intercept point Cq 32 is obtained (figure 3.7a). Therefore, detection of a Cq ≥33 is considered low bacterial load (LBL) while detection of a Cq ≥16≤32 is considered high bacterial load (HBL). It is important to know that our objective is not to identify presence or absence of a pathogen which obviously uses limit of detection or lowest concentration of the two-fold dilution of the standard but to identify values for HBL and LBL

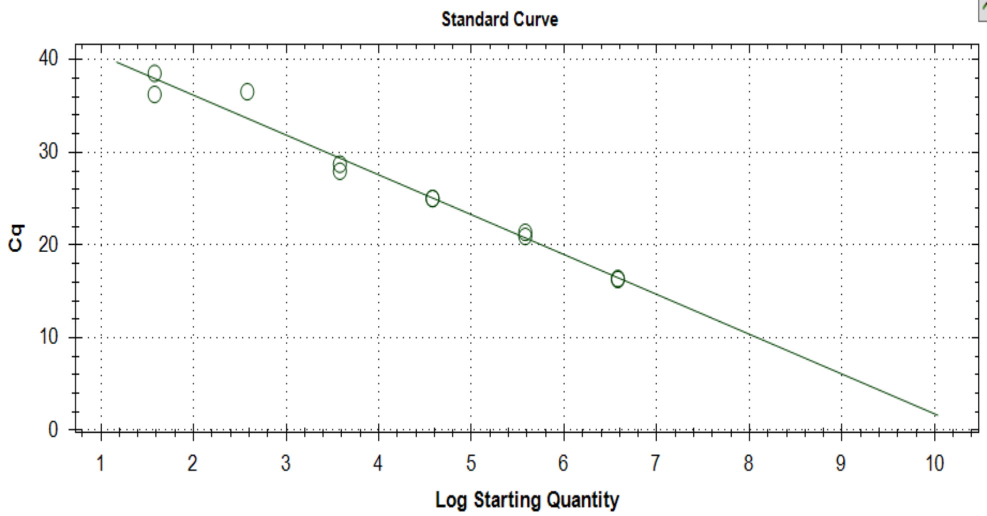


Figure 3.6: Showing quantification progression of gene *aatA* (HEX)

HEX (*aat*A): E=70.8%, R^2=0.977, Slope= -4.301, y-int=44.789

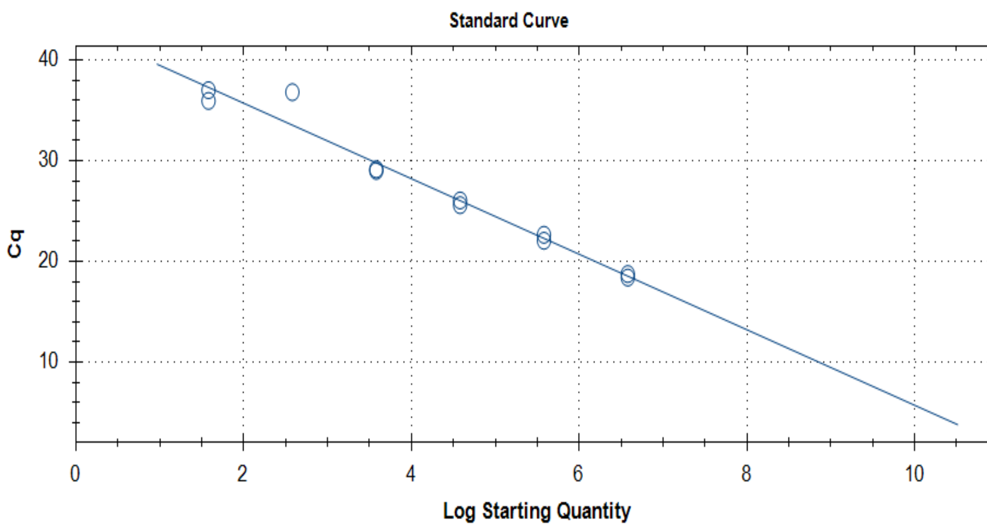


Figure 3.7: Showing quantification progression of gene *aaiC* (FAM)

FAM (*aai*C): E=84.9% R^2=0.970, Slope=-3.748 y-int=43.195

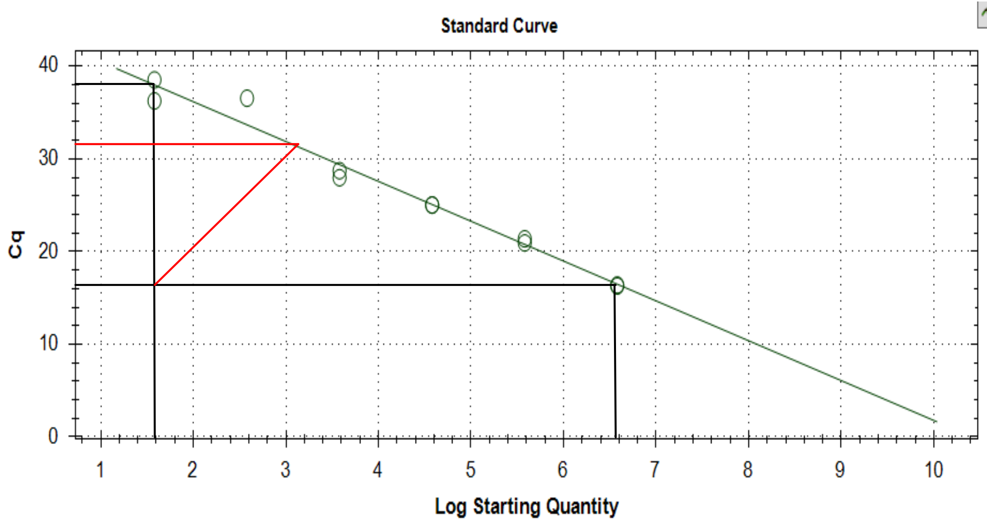


Figure 3.7a: Showing highest concentration of standad at Cq 16 and at log starting quantity 6.6. Similarly, lowest concentration at Cq 38 and at log start quantity 1.6 of gene *aatA* (HEX)

3.15.5 Statistical Analysis

Statistical analysis was performed using stata software version 12 (StataCorp). To determine the accuracy of qPCR in detecting disease, we used receiver operating characteristic analysis. Also, sensitivity and specificity of the qPCR was determined using the detection of EAEC as independent measure of standard. The distribution of Ct values from EAEC-high bacterial load and EAEC-low load of cases and controls were compared using student's t-test.

Conditional logistic regression was used to determine the association between EAEC bacterial load and diarrhoea. A crude model showing the relationship of EAEC bacterial load and diarrhoea was fitted. Then, this model was adjusted for potential confounders one at a time and the association between EAEC bacterial load and diarrhoea noted in the presence of each potential confounder. A forward selection procedure was used to build the final multivariable model starting with the variable with the lowest p-value. Variables with a p-value >0.2 were not added to this model.

3.16 Biofilm production in EAEC strains

3.16.1 Strains

A total of 400 confirmed EAEC strains from symptomatic and non-symptomatic children were studied for production of biofilm or slime layer. The target virulent factor(s) to detect EAEC using amplification technique are either both *aatA* (CVD32) and aaiC or one of the two. In this study, we used three recognised phenotypic biofilm screening methods that include tube test (tt), Congo red agar (CRA) and tissue culture plate (TCP) screening methods. The tube (figure 3.8) and CRA (figure 3.9) are qualitative methods whilst TCP (figure 3.10) is quantitative method. Known biofilm producing *E. coli* strain (EAEC) 042 and non-biofilm producing *E. coli* strain HB101 were included as controls in the three methods.

3.16.2 Tube screening method

Principle: The visual observation of adherence of 24 hours enriched cultured bacteria to smooth surface of a glass tube following stained with simple stain such as Safranin predicts biofilm/slime production (Christensen, Simpson et al. 1985).

Procedure

10ml TSB in Borosilica tube

Inoculate the medium with the 24 hrs fresh isolate of the test bacteria

Incubate at 35°C or 37°C for 48 hrs

Discard the supernatant

Stain the borosilica tube with 0.1% Safrannin or 0.1% crystal violet solution

Wash with distilled/H_2O three times

Air dry and Examine for slime production

Presence of adherence of slime to the inner wall of tube ----------- Positive

Stained ring at the liquid-air interface -------------------------Negative

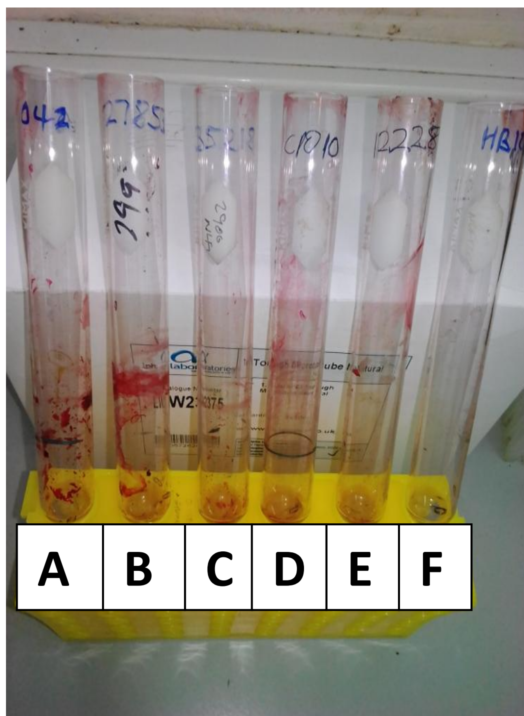


Figure 3.8: Test tube method for screening biofilm – Tube A, B, C and D serves as positive controls whilst tube E and F serves as negative controls

3.16.3 **Congo red agar screening method**

Principle: The enhancement of exopolysaccharide production by bacteria using enriched medium such as brain-heart-infusion-broth with 5% sucrose supplement that detects glucan production and Congo red (CR) is added to stain presence of exopolysaccharide which is slime produced by aquatic Gram negative bacilli (Freeman, Falkiner et al. 1989). Although the exact mechanism of the phenomenon is unclear however an explanation given in an unpublished report revealed that the black colouration seen in Congo red agar (CRA) plate is due to the presence of curlie fibres which contributes to the formation of exopolysaccharide. The curli fibre is distinguishable due to its binding characteristics with CR in a cultured medium.

Procedure

Inoculate the CRA plate with a 24 hrs fresh bacteria growth from either BHI or TSB

Incubate the cultured CRA plate at 37°C for 24 hrs to 48 hrs

Examine the plate for the biofilm production

Darkening or blackening in and around the growth bacteria ---------------------Positive

Absence of darkening or blackening in and around the growth bacteria -------Negative

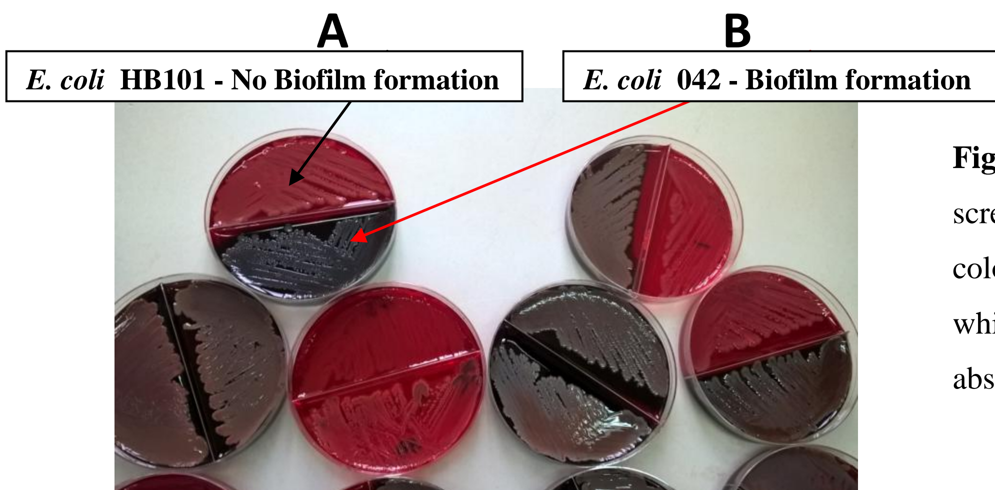


Figure 3.9: Congo red Method for screening biofilm – Dack black colonies B indicate biofilm formation whilst pink-red colonies A indicate absence of biofilm

3.16.4 Tissue culture plate screening method

Principle: Quantitative biofilm assay that involves incubation of bacterial overnight in rich-glucose medium Tryptic soy broth medium in a polystyrene microtitre plate. The plate was stained with crystal violet following washing, and the biofilm quantified using an enzyme-linked immunosorbent assay plate reader (Christensen, Simpson et al. 1985).

Procedure

Inoculate 10ml Tryptic soy broth of 1% glucose with a 24 hour fresh isolates of the test bacteria

Incubate at 37°C for 24 hours

Dilute cultures in 1:100 with fresh medium of TSB-1% glucose (2µl culture + 198µl TSB)

Mix well in 96-well plates

Set up +ve and –ve controls

Incubate at 37°C for 24 hours

Remove each well content by gentle tapping

Wash well with 0.2ml of Phosphate Buffer Saline (PBS) 4 times

Add 10% formaldehyde (25% v/v formalin) to fix the biofilm for 30 minutes. Or, Air dry at 60°C for 45 minutes to fix the biofilm)

Apply 0.1% crystal violet to stain biofilm for 5 minutes

Remove excess stain using distilled water

Air dry the plate

Add 200 µl of 95% ethanol to the well to solubilise

Quantify at 570nm single-wavelength mode (lambda) using ELISA plate reading machine

Background staining is assessed in control wells inoculated with sterile TSB

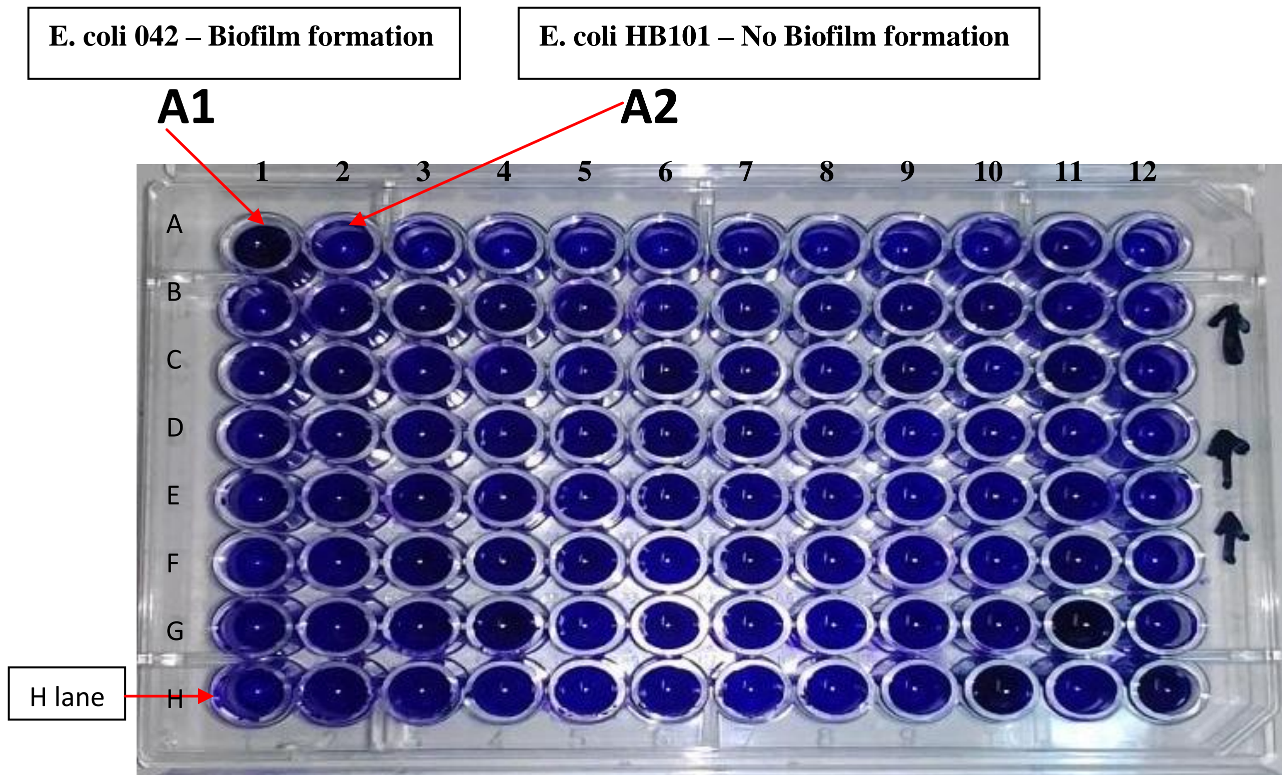


Figure 3.10: Tissue culture plate screening method for the presence or absence of biofilm – Well A1 is Biofilm positive and well A2 is Biofilm negative. Also, well H1-to- H9 and H11 are Biofilm negative while well H10 and H12 are biofilm positive.

3.17 Antibiotic resistance Assay

3.17.1 Summary

Antimicrobial susceptibility testing has two purposes. First, it is used clinically to predict through an *in vitro* assessment the likelihood of successfully treating a patient's infection with a particular antimicrobial agent. Second, it can provide a quantitative measurement of susceptibility, which can be used to monitor the emergence and prevalence of antimicrobial resistance. Thus, changes in the selective ecological pressure of antibiotic can also change the resistant patterns of the epidemic (or endemic) strains. Currently, the three most popular susceptibility-testing methods are disk diffusion, agar dilution/broth microdilution and E-test techniques.

3.17.2 Antibiotic resistance using disk diffusion method

Principle: Discs of blotting paper were impregnated with a known volume and appropriate concentration of different antimicrobial agents, these are placed on a plate of sensitivity testing agar carpet inoculated with the test organism (EAEC). The antimicrobial agent diffuses from the disc into the medium. Following overnight incubation at 37°C, the culture is examined for areas of no growth around the discs (inhibition zones). Bacterial strains sensitive to the antimicrobial agents are inhibited at a distance from the disc whereas resistant strains grow up to the edge of the disc. The width of the inhibition zones is measured in millimetres and gives an indication of the sensitivity of the EAEC to antimicrobial agents being tested. In this method, two techniques are often used. Stokes comparative disc diffusion technique involves a control organism that is inoculated on the sample plate. While the Kirby-Bauer used in this study involves the control strain *E. coli* ATCC 25922) inoculated on

a separate plate. The Kirby-Bauer allows categorization of bacterial isolates as susceptible, resistant, or intermediate to ten commercially acquired antimicrobial agents which include; Ampicillin 10µg, Sulphamethoxazole-trimethoprim 25µg, Tetracycline 30µg, Ceftazidime 30µg, Ciprofloxacin 5µg, Ceftriaxone 30µg, Cefoxitin 30µg , Chloramphenicol 30µg, Gentamicin 30µg and Amoxicillin-clavulanic acid 30µg (Oxoid, Hampshire, UK) were used. The Clinical Laboratory Standard Institute (CLSI 2012) guidelines were followed for the antibiotic resistance assay.

Procedure

Viable four-hundred EAEC isolates stored at -70°C were recovered for the antimicrobial drug resistance investigation by sub cultured on MacConkey agar medium (oxoid 333M) and incubate at 37°C overnight.

Touch the growth single colony of EAEC with sterile straight wire loop. Transfer growth into 2ml sterile distilled water. The suspension is matched and adjusted with the density of the 0.5 McFarland standard (BioMeurieux SA, France).

Within 15 minutes after adjusting the turbidity of the inoculum suspension, a sterile non-toxic swab is dipped into the adjusted suspension of EAEC. The swab is rotated several times and pressed firmly on inside wall of the tube above the fluid level to remove excess inoculum from the swab.

The dried surface of a Mueller-Hinton agar (Oxoids, Basingstoke, England, Unipath catalogue number CM337) plate (88 mm in dm) is inoculated by streaking the swab over the entire sterile agar surface. The streaking procedure is repeated two more times, rotating the

plate approximately 60°C each time to ensure an even distribution of inoculum. This will enhance zone of inhibition to be uniformly circular for a confluent lawn of growth.

The appropriate concentration of impregnated discs of different antimicrobial agents is placed on the surface of the inoculated/streaked Mueller-Hinton agar plate with sterile forceps. The discs were gently pressed down to ensure complete contact with the agar surface without movement.

The plates were inverted and incubated aerobically at 36°C-37°C for 18-24 hours (using LEEC Compact Incubator, England).

The plates were examined for susceptibility assay after 18-24 hours incubation

Interpretation of result:

Zones of inhibition of the control strains (ATCC 25922 *Escherichia coli*) and the test EAEC strains are measured in millimetre in diameters with a ruler.

The zone of inhibition for the control strain and the test strains were confirmed to fall within the acceptable zone range following CLSI 2016 guideline.

3.18 Whole Genome Sequencing of Enteroaggregative *Escherichia coli* (EAEC) strains

3.18.1 Isolation of genomic DNA from EAEC isolates

Preparation of EAEC isolates

Base on the antibiotic sensitivity pattern, fifty-one EAEC Isolates were randomly selected retrieved from the -70°C, sub-cultured on MacConkey agar medium and purified on Nutrient agar medium. Following overnight incubation at 37°C a pure colony was picked and inoculated in sterile enrichment medium tryptic soya broth (TSB). Also, this culture was incubated at 37°C overnight.

Preparation of Reagents and materials

- Equilibrate the sample to room temperature (15–25°C).
- Heat 2 water baths or heating blocks: one to 56°C and one to 70°C.
- Equilibrate Buffer AE or distilled water to room temperature for elution.
- Ensure that Buffers AW1 and AW2 have been prepared according to the instructions in the manual.
- If a precipitate has formed in Buffer ATL or Buffer AL, dissolve by incubating at 56°C.

Procedure

1. Add 1ml of an overnight liquid culture Tryptone Soya Broth (TSB) to a 1.5ml microcentrifuge tube.
2. Centrifuge at 13,000–16,000 × *g* for 2 minutes to pellet the cells. Remove the supernatant.
3. Add 180µl of Buffer ATL Solution. Gently pipet until the cells are resuspended.

4. Add 20 µl proteinase K, mix by vortexing, and incubate at 56°C for 1 hour or until the tissue is completely lysed. (Vortex occasionally during incubation to disperse the sample, or place in a shaking water bath or on a rocking platform).

5. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from the inside of the lid.

6. Add 200 µl Buffer AL to the sample, mix by pulse-vortexing for 15 s, and incubate at 70°C for 10 min. Briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid.

7. Add 200 µl ethanol (96–100%) to the sample, and mix by pulse-vortexing for 15 s. After mixing, briefly centrifuge the 1.5 ml microcentrifuge tube to remove drops from inside the lid

8. Carefully apply the mixture from step 7 (including the precipitate) to the QIAamp Mini spin column (in a 2 ml collection tube) without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the tube containing the filtrate

9. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW1 without wetting the rim. Close the cap, and centrifuge at 6000 x *g* (8000 rpm) for 1 min. Place the QIAamp Mini spin column in a clean 2 ml collection tube (provided), and discard the collection tube containing the filtrate

10. Carefully open the QIAamp Mini spin column and add 500 µl Buffer AW2 without wetting the rim. Close the cap and centrifuge at full speed (20,000 x *g*; 14,000 rpm) for 3 min

11. Place the QIAamp Mini spin column in a new 2 ml collection tube (not provided) and discard the old collection tube with the filtrate. Centrifuge at full speed for 1 min. This step helps to eliminate the chance of possible Buffer AW2 carryover

12. Place the QIAamp Mini spin column in a clean 1.5 ml microcentrifuge tube (not provided), and discard the collection tube containing the filtrate. Carefully open the QIAamp Mini spin column and add 200 µl Buffer AE or distilled water. Incubate at room temperature for 5 min to increase DNA yield, and then centrifuge at 6000 x *g* (8000 rpm) for 1 min

Keep the first 200 µl of extracted DNA for further experiments and label the tube with ID# and "_1DNA" code.

Repeat step 12: Keep the second 200 µl of extracted DNA for further experiments and label the tube with ID# and "_2_DNA" code.

Repeat step 12: Keep the third 200 µl of extracted DNA for further experiments and label the tube with ID# and "_3_DNA" code. (A third elution step with a further 200 µl Buffer AE will increase yields by up to 15%).

Store the extracted DNA microtubes in duplicate at -70°C.

3.18.2 DNA Quality Control

The minimum DNA concentration required to run whole genome sequencing (WGS) is 20 ng/µl. All samples used for the WGS are well above the 20 ng/µl (table 3.3). The instrument used for quantitation was Qubit Fluorometric Quantitation (Thermo Fisher scientific) following manufacturer's instructions.

Table 3.3: Concentration of DNA extracts used for Whole Genome Sequencing

Sample Number	Sample Name	Conc. (ng/µl)	Elution Vol (µl)
1	102191_1_DNA	25.8	200
2	102106_1_DNA	27	200
3	100801_1_DNA	42	200
4	100722_1_DNA	68	200
5	100503_1_DNA	41.6	200
6	100427_1_DNA	41.2	200
7	100415_1_DNA	61.4	200
8	100382_1_DNA	89.6	200
9	100125_1_DNA	120	200
10	100020_1_DNA	61.4	200
11	103069_1_DNA	98.8	200
12	103047_1_DNA	77.8	200
13	102742_1_DNA	36.4	200
14	102296_1_DNA	57	200
15	100119_1_DNA	50.8	200
16	103530_1_DNA	30.8	200
17	103275_1_DNA	55.2	200
18	100404_1_DNA	23.8	200
19	103076_1_DNA	46	200
20	100096_1_DNA	69.4	200
21	100191_1_DNA	47.6	200
22	103016_1_DNA	102	200
23	100590_1_DNA	41.4	200
24	103709_1_DNA	42.6	200
25	102602_1_DNA	79.8	200
26	102871_1_DNA	64.6	200
27	100715_1_DNA	37.4	200
28	102274_1_DNA	22	200
29	103070_1_DNA	26	200
30	103446_1_DNA	62.4	200
31	100127_1_DNA	87.4	200

32	100138_1_DNA	116	200
33	100182_1_DNA	108	200
34	100342_1_DNA	106	200
35.	100569_1_DNA	24.1	200
36	100796_1_DNA	98.2	200
37	102031_1_DNA	71.6	200
38	102098_1_DNA	118	200
39	102375_1_DNA	77.2	200
40	102425_1_DNA	100	200
41.	102469_1_DNA	40	200
42	102705_1_DNA	65.0	200
43	102806_1_DNA	63.2	200
44.	102820_1_DNA	40.3	200
45	102906_1_DNA	40.2	200
46	102951_1_DNA	82	200
47	103276_1_DNA	40.6	200
48	103278_1_DNA	78.4	200
49.	103400_1_DNA	40.7	200
50.	103691_1_DNA	28.6	200
51	103693_1_DNA	99.0	200
52	ENC 1	Too low	
53	ENC 2	Too low	
54	ENC 3	Too low	

3.18.3 Library preparation (Fragmentation)

Library preparation usually includes shearing the DNA either mechanically or enzymatically, adding adaptors and barcodes/indexes and amplification, was done using the NEBNext® Ultra™ II DNA Library Prep Kit Illumina® (E7645).

Genomic DNA was fragmented at 400 base pairs (bp) via sonication (Covaris ™ M220 Focussed-ultrasonicator ™ Instrument) and tagged for multiplexing with NEBNext® adaptors.

Reagents in the NEB #E7645 kit required for Library Preparation

Package 1: Store at –20°C.

- (green) NEBNext Ultra II End Prep Enzyme Mix
- (green) NEBNext Ultra II End Prep Reaction Buffer
- (red) NEBNext Ultra II Ligation Master Mix
- (red) NEBNext Ligation Enhancer
- (blue) NEBNext Ultra II Q5 Master Mix

Required Materials

- 80% Ethanol (freshly prepared)
- Nuclease-free Water
- 0.1X TE (1 mM Tris-HCL, pH 8.0, mM EDTA)
- DNA LoBind Tubes
- Magnetic rack / stand
- PCR Machine

Procedure

NEBNEXT End Prep

Add to sterile nuclease-free tube;

NEBNext Ultra II End Prep Enzyme Mix --------------- 3 µl

NEBNext Ultra II End Prep Reaction Buffer --------------- 7 µl

Fragmented DNA -- 50 µl

Mix well and do a quick spin to collect all liquid from the sides of the tube

Place the End Prep Reaction Mixture in a heated lid thermocycler set to ≥75°C, and run the following programme; (1) 30 minutes at 20°C, (2) 30 minutes at 65°C and (3) Hold at 4°C

Adaptor Ligation

Add the components below to the 60 µl End Prep Reaction Mixture;

NEBNEXT Ultra II Ligation Master Mix -------------------------- 30 µl

NEBNext Ligation Enhancer --------------------------------------- 1 µl

NEBNext Adaptor for illumine ------------------------------------ 2.5 µl

Incubate the Ligation Mixture at 20°C for 15 minutes

Add 3 µl Enzyme to the ligation mixture

Mix well and incubate at 37°C for 15 minutes with the heated lid set to ≥ 47°C

3.18.4 Size Selection of Adaptor-ligated DNA

Do **size selection** if starting material is ≥50 ng. Do **clean up** if starting material is ≤ 50 ng.

Size selection enriches for molecules that were shred to the desired size and have an adaptor ligated to each end. Size selection is accomplished using magnetic beads.

There are two round of selection. First round removes DNA fragment larger than the desired size. Second round removes DNA fragment smaller than desired size. These are accomplished using specific ratios of the beads solution to total volume. Volume of beads required varies depending on the desired fragments.

Library Amplification by PCR

This step increases the amount of library and also select for molecules that have an adaptor ligated to each end (for multiplex library – idecies or barcodes can be introduced at this step if the NEBNext Adaptor and primer are used)

Clean up PCR Reaction

Introduce magnetic-beads of about 45 µl to the PCR reaction and mix well. DNA library bound to the beads. Incubate at room temperature for five minutes. Place the tube/plate on an appropriate magnetic stand to separate the beads from the supernatant. Wash beads with 200 µl 80% ethanol, wait 30 seconds and then remove the ethanol. Air dry the beads (do not over dry). Elute the library from the beads using 0.1 X TE buffer. Mix well. Put the sample in the magnetic field until the sample clear. Remove about 30µl of the supernatant containing the library to a new tube. This library can be stored at -20°C

3.18.5 Library Analysis

Confirm the library size distribution by diluting 1 μl of library 5 folds with 10mM Tris-HCl or 0.1 X TE buffer. Run it on Bioanalyser using the high sensitive chip.

Library Quantitation - KAPA

Final library is quantitated using qPCR-based methods, such as NEBNext Library QUANT-Kit or electrophoretic methods like the Bioanalyzer

3.18.6 Template Preparation and Automated Sequencing

Template preparation was accomplished by bridge amplification and or emulsion PCR. Automated Whole genome sequencing was achieved using the Illumina MiSeq platform with 2 x 250 bp reads. The MiSeq illumina is based on sequencing by synthesis of the complementary strand and fluorescence-based detection of reversibly blocked terminator nucleotides. The template includes multiple instruments with varying throughput range from 0.3 to 15 Gb and read length 1 x 36 to 2 x 300 bp with scalability. MiSeq instrument (figure 3.13) is an attractive choice for diagnostic and public health laboratories as it offers low to mid sample throughput, affordable pricing and user friendly.

Data Analysis in the Sequencer

Sequencing from pool liberary are separated base on the unique indicies introduced during the samples prepararion. Local clustring that involves reads with similar sequence base call are localy cluster together. Forward reads and reverse reads appear to create contiguous

sequence are align back to the reference genome for variant identification. The pair-end information is use to resolve to make alignment.

3.18.7 Raw Data Quality Control

Short Illumina reads were quality controlled using FastQC (v0.11.5;). This is done to ascertain the quality of the sequence products prior proceeding to analysis of the data. Therefore, FastQC checks for sample contamination, Reads quality, problematic reads, number of reads mapped, percentage genome covered, quality scores and filters supporting reads and sample present (figure 3.11 and figure 3.12).

Figure 3.11: FastQC: Quality Score Per read base

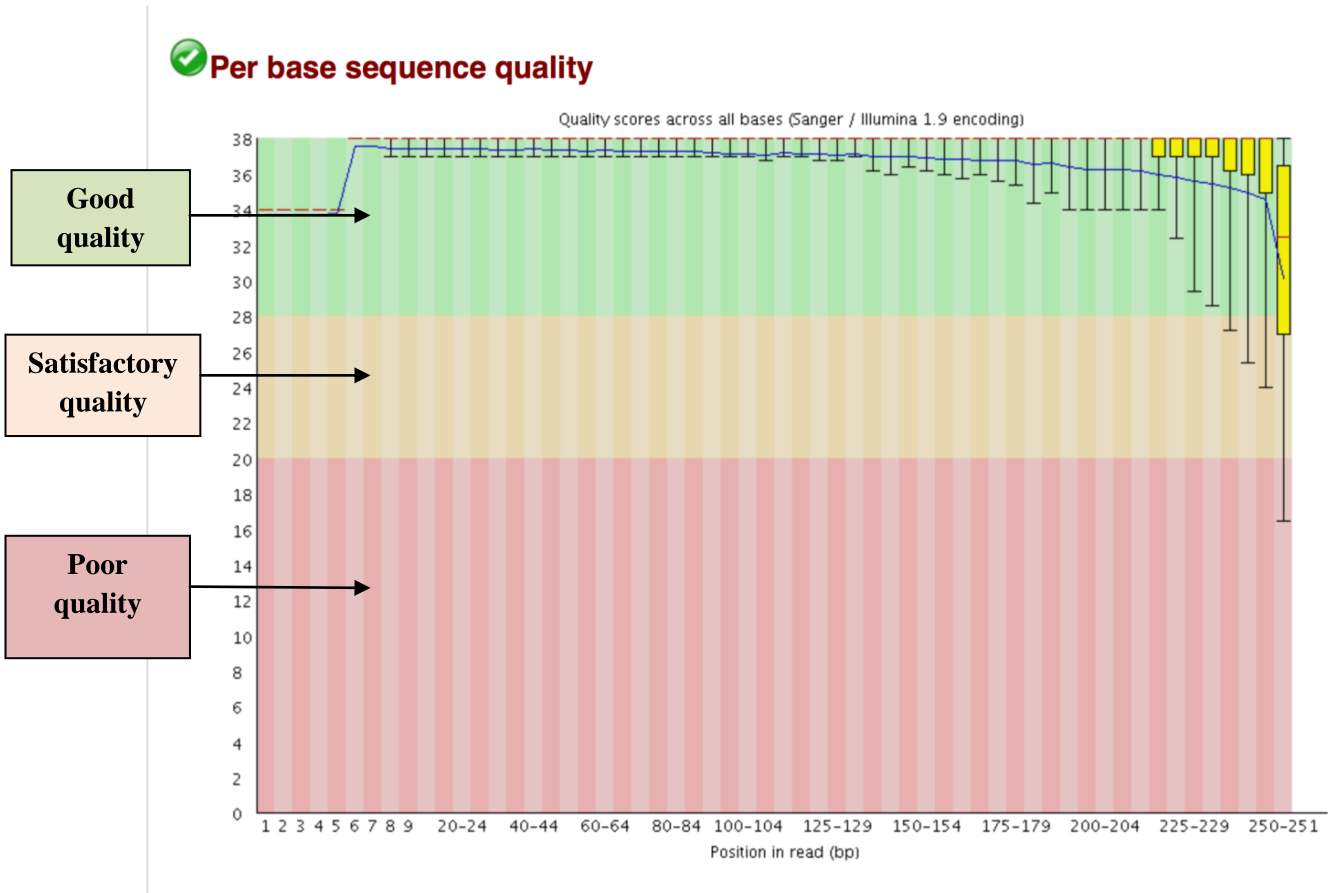
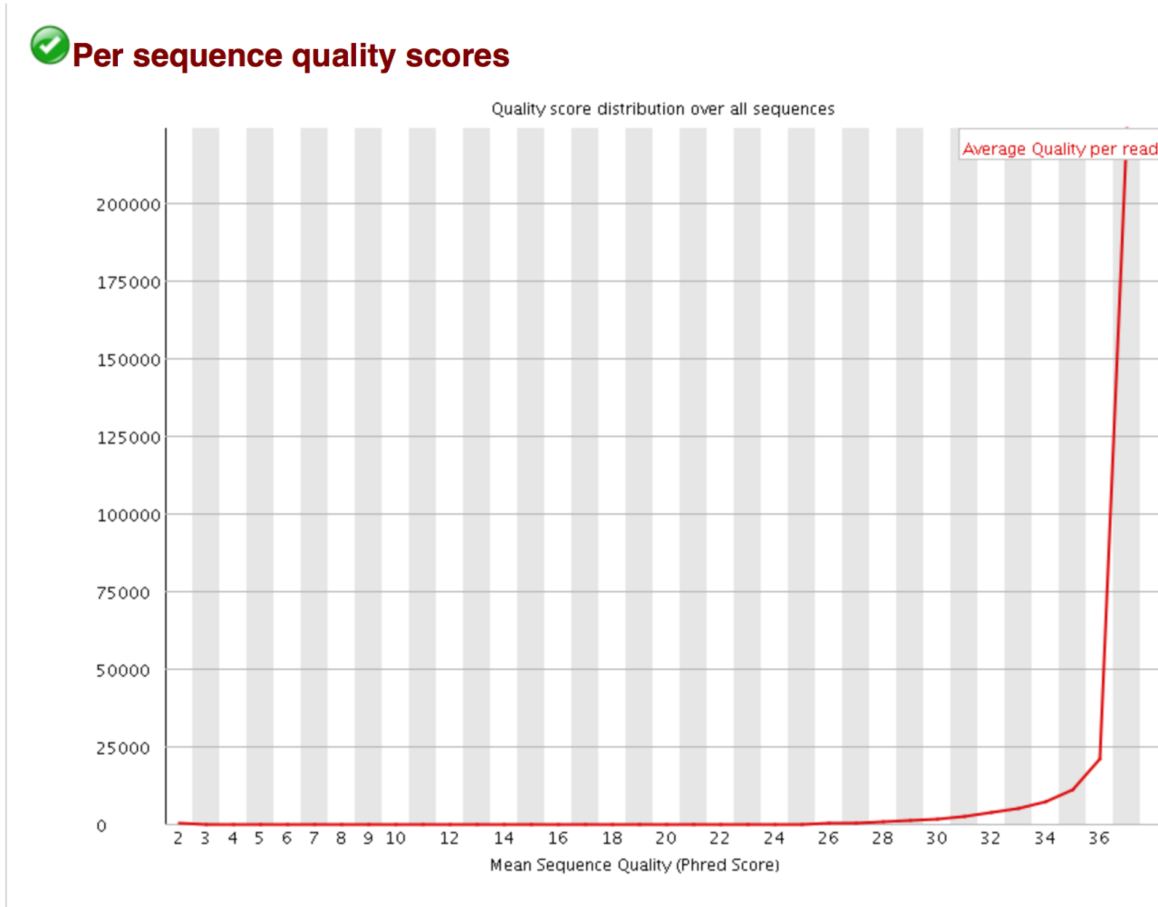


Figure 3.12: Distribution of Quality score across all sequences



3.18.8 Sequencing analysis

Short Illumina reads were quality controlled using FastQC (v0.11.5;). Based on the FastQC report, we used the following pipeline employed in Nullarbor (v1.3dev; (Seeman 2016)*)* for our downstream analysis. Low quality reads as well as adaptors were trimmed from both end of reads for each genome using trimmomatic (v0.36; (Bolger, Lohse et al. 2014)). Reads were then de-novo assembled to generate contigs for each genome using SPAdes (v3.11.1; (Bankevich, Nurk et al. 2012)) and coding sequences (CDS) predicted and annotated by Prokka (v1.12 (Seemann 2014)). Multi locus sequence types for each genome was determined by MLST (v2.8). Resistance genes as well as virulence genes for each genome was inferred using Abricate (v0.7). Also core and accessory genome analysis was determined using roary software (v3.11.2; (Page, Cummins et al. 2015)) and *E. coli* str. K-12 substr. MG1655 reference genome. The detail of tools used to perform analysis is shown in the table 3.4.

3.18.9 Phylogenetic analysis

Sequencing reads were mapped to *E. coli* str. K-12 substr. MG1655 reference genome using BWA MEM (v0.7.17-r1188; (Li and Durbin 2009)) (downloaded from NCBI 8th May 2018). Single nucleotide polymorphisms were called using Snippy (v.4.0 deb2) and approximate maximum likelihood phylogenetic tree reconstructed using FastTree (v.2.1.10; (Price, Dehal et al. 2010)). We used iTOL web tools to annotate the output Phylogenetic tree (https://itol.embl.de/login.cgi?logout=1).

Figure 3.13: Whole Genome Sequencing work flow

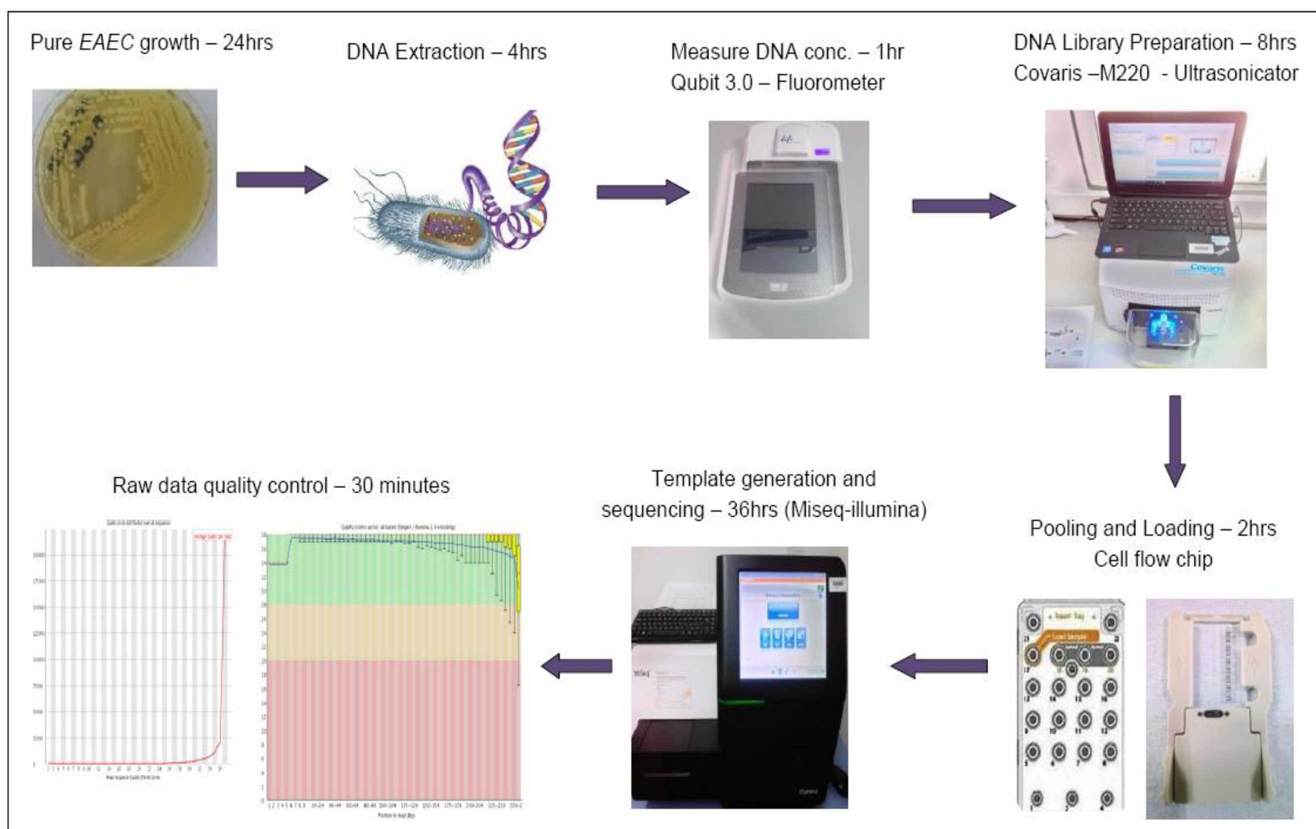
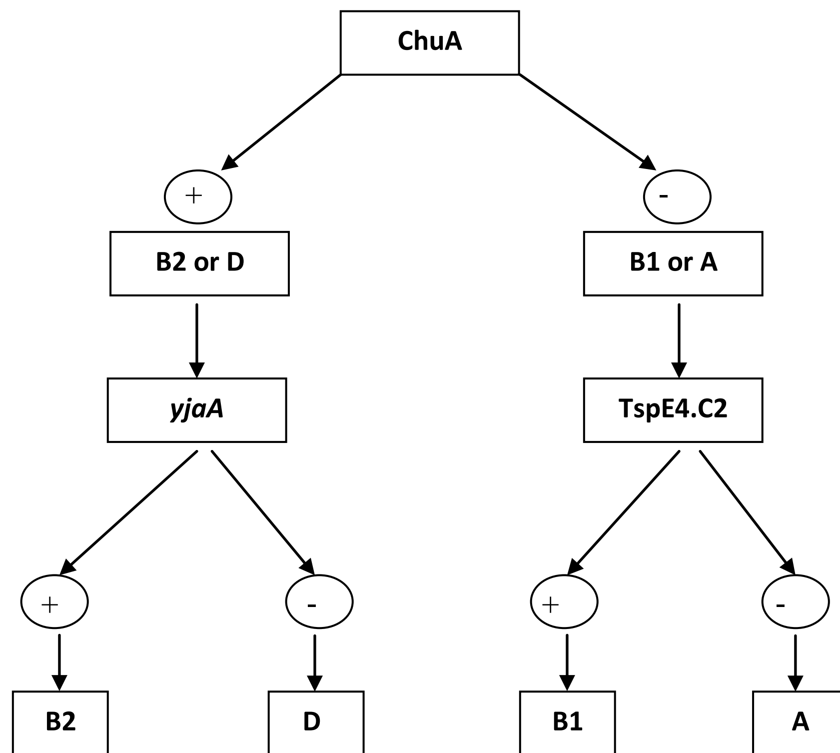


Table 3.4: Bioinformatic analysis performed and detail of tools used

Analysis	Tools	Version
Pipeline software Generate complete reports from sequenced isolates	Nullarbor	1.30-dev
Virulome (Virulence genes) From assembly	Abricate	0.7
Resistome (Resistance genes) From assembly	Abricate + *Resfinder*	0.7
Alignment method Aligning short reads, map more reads with high sequencing error, greater sequence variation. Accept reads only in fasta or fastq	BWA MEM	0.7.17-r1188
Infer core SNP phylogeny Maximum likelihood	FastTree	2.1.10 Double precision (No SSE3), OpenMP (32 threads)
Bayesian variant detector – design to find small polymorphism (SNPs)	FreeBayes	1.1.0-dirty
Species identification (k-mer analysis against known genome database)	Kraken	1.0
MLST – From assembly w/ automatic scheme detection	MLST + *PubMLST*	2.8
De novo assembly	MegaHit	1.1.2
Process Phylogenetic trees Functions include re-rooting, extracting subtrees, trimming, pruning, condensing, drawing	Newick-Utils – Unix shell tools	(unable to determine version)
Annotation Adding features to assembly	Prokka	1.12
Pan genome From annotated consigns	Rosary	86_64-linux-gnu/perl/5.22/Encode.pm line 59.
Generic format for storing large nucleotide sequence alignment	SAM tools	1.7
De novo assembly	Spades	3.11.1
Core genome SNPs	Snippy-core	4.0-dev2
Variants – From reads aligned to reference	Snippy + VFDB	4.0-dev2
Infer core SNP phylogeny SNP distance matrix	snp-dists	0.2
Reads and Write Sequences Extract sequences from database & display sequences	seqret	6.6.0.0
Clean reads – Remove adaptors, low quality bases and reads	Trimmomatic	0.36
Annotate the output phylogenetic tree	iTOL	Web base tool (https://itol.embl.de)

Figure 3.14: Dichotomous decision tree to determine the phylogenic group of an EAEC isolate by the presence or absence of *chuA*, *yjaA* and TspE.C2 from WGS virulence gene result



Chapter 4: Prevalence of Virulence genes among EAEC strains from MSD and non-MSD children

4.1 Introduction

Enteroaggregative *E. coli* (EAEC) is an important causative agent of both acute and persistent diarrhoea among adults and children worldwide (Nataro, Steiner et al. 1998) and it has been among the most common *E. coli* pathotypes causing diarrhoea among children less than five years of age in some developing countries. (Moyo, Maselle et al. 2007). Several outbreaks of EAEC diarrhoea have been reported in both developed and developing nations and infants are the most affected (Cobeljic, Miljkovic-Selimovic et al. 1996; Itoh, Nagano et al. 1997; Pai, Kang et al. 1997; Smith, Cheasty et al. 1997). EAEC has been implicated in travellers' diarrhoea (Adachi, Jiang et al. 2001; Januszkiewicz, Szych et al. 2012) and persistent diarrhoea among human immunodeficiency virus (HIV) infected individuals (Samie, Obi et al. 2007). This pathotype was implicated in a massive outbreak of haemolytic uremic syndrome (HUS) in Germany in 2011 (Bielaszewska, Mellmann et al. 2011). The clinical presentation of EAEC infection is characterized by watery and mucoid diarrhoea with low-grade fever and insignificant vomiting (Bhan, Khoshoo et al. 1989; Paul, Tsukamoto et al. 1994).

The pathogenesis of EAEC diarrhoea is thought to comprise colonization of the intestinal mucosa, followed by elaboration of enterotoxins and cytotoxins and the release of proinflamatory cytokines from infected epithelial cells (Harrington, Strauman et al. 2005; Harrington, Dudley et al. 2006), induced by the EAEC adherence factors called Aggregative Adherence Fimbriae (AAFs). Additionally,

EAEC strains characteristically enhance mucus secretion from the mucosa, potentially trapping the bacterium in a bacterium-mucus biofilm (Nataro, Steiner et al. 1998). A distinctive feature of EAEC is its ability to elicit characteristic stacked brick-like aggregative adherence to HEp-2 or HeLa cells, a test that remains the gold standard to identify this pathotype (Nataro and Martinez 1998). EAEC strains express several genes that may confer virulence and are highly heterogeneous regarding the combination of these virulence genes, which are encoded on the bacterial chromosome or on an EAEC-specific plasmid designated pAA. The majority of EAEC strains harbour a transcriptional activator of the AraC/XyIS fairly called *Agg*R, which control genes on both the plasmid and the chromosome. Among the genes under *Agg*R control includes those that encode the Aggregative Adherence Fimbriae (AAFs) where at least five variants exist. These genes encoding the major structural pilin subunits are designated as *aggA* (AAF/I), *aafA* (AAF/II), *agg3A* (AAF/III), *agg4A* (AAF/IV) and *agg5A* (AAF/V) (Nataro, Yikang et al. 1994; Czeczulin, Balepur et al. 1997; Jonsson, Struve et al. 2015). Other plasmid-borne potential virulence factors include the EAEC heat-stable enterotoxin 1 EAST1 (encoded by the *astA* gene) (Savarino, Fasano et al. 1991), an anti-aggregation protein called dispersin (encoded by the *aap* gene), and a transporter apparatus for dispersin called Aat (encoded by the *aat* genes). EAEC frequently harbour members of the serine protease autotransporters of Enterobacteriaceae (SPATEs), which have been described as enterotoxins and cytotoxins. The heat-labile enterotoxin/cytotoxin called Pet (Plasmid-encoded toxin) (Nataro, Steiner et al. 1998) has been implicated in causing cytotoxic effects on the human intestinal mucosa. Other SPATEs carried by EAEC strains include the cryptic protease called *SepA*, and the mucinase

called *Pic* (protein involved in intestinal colonization) (Henderson, Czeczulin et al. 1999; Kahali, Sarkar et al. 2004) which is encoded on the chromosome. Other important chromosomal gene that encodes virulence markers include (i) *Irp*2 (iron repressible high-molecular-weight protein 2) a protein responsible for yersiniabactin biosynthesis and (ii) flagellin, which interacts with the epithelial cells, leading to the secretion of an intestinal interleukin-8 (Steiner, Nataro et al. 2000). The EAEC genome has been found to be markedly mosaic, thus the various putative virulence factors are found inconsistently among individual strains, suggesting that some strains considered EAEC may be truly virulent, and others not (Nataro, Steiner et al. 1998).

Several studies have shown that EAEC is the most frequently detected *E. coli* pathotype in humans, particularly among children from both developed and developing countries (Presterl, Nadrchal et al. 1999; Knutton, Shaw et al. 2001; Cohen, Nataro et al. 2005). The GEMS comprised of a case-control study of moderate-to-severe diarrhoea among children less than five years of age at four sites in sub-Saharan Africa and three in south Asia showing high frequency of EAEC (Kotloff, Nataro et al. 2013). Although, EAEC was not associated with moderate-to-severe diarrheal disease in GEMS, a subsequent analysis of the association of individual EAEC genes alone and in combination among EAEC isolates from moderate-to-severe diarrhoea cases and controls of GEMS site in Bamako Mali found that *Sep*A protease was associated with moderate-to-severe diarrhoea (Boisen, Scheutz et al. 2012). In this chapter we replicated the analysis by Boisen *et al.* (2012), scoring the presence of twenty-one putative EAEC virulence factors from 428 EAEC isolates randomly selected among 741 EAEC isolates obtained from diarrheal and non-diarrheal children enrolled in the GEMS

study to characterize the virulence genes in children from these isolates from The Gambia. We analyzed these EAEC virulence genes by age strata (0-11, 12-23 and 24-59 months).

4.2 Result

Among all EAEC strains in cases and controls (n = 428), the age and sex distribution were similar among cases and controls except for a lower EAEC prevalence in children above 23 months among cases (table 4.1). Overall, orf61 (*aar*) was the most commonly detected gene, (69.6%). This was followed by the cryptic ORF3 (64%), *capU* (62%), *aggR* (60.1%), *astA* (51.4%), *eilA* (48.3%) and *aap* (46.3%); the rest of the genes were present in less than 40% of isolates (Table 4.2). Analysis of the EAEC virulence genes in all age groups together, showed that only four of the twenty-one genes assayed (*sepA*, *pet*, *astA* and *capU*) were more prevalent among cases. Prevalence of AAF/I encoded by *aggA* gene was slightly higher in cases than controls (29.9% versus 22.9%) (OR 1.4, 95% CI 0.89-2.29, P = 0.106). The frequency of other AAF pilin genes, AAF/II (aafA) and AAF/III (agg3A) were low in both cases and controls but slightly high for AAF/IV (agg4A) in cases compared to controls. However, the AAF usher-encoding gene agg3/4C was similar in cases and controls (36.9% vs 35.4% respectively). Of the five SPATE genes (*sat*, *pet*, *sigA*, *pic* and *sepA*), prevalence of *sepA* (OR 1.6, 95% CI 0.99-2.49, P = 0.041) and *pet* (OR 1.9, 95% CI 0.97-3.56, P = 0.042) genes were higher among diarrhoea cases (Table 4.2).

Table 4.1: Baseline information of study population

Demographic factors	Case (N = 157) No. (%)	Control (N = 271) No. (%)	Total (N = 428) No. (%)	OR (95% CI)	P - Value
<u>Age (month)</u>					
0-11	85 (54.1)	132 (48.7)	217 (50.7)	1.2 (0.82-1.88)	0.278
12-23	61 (38.9)	105 (38.8)	166 (38.8)	1.0 (0.65-1.53)	0.982
24-59	11 (7.0)	34 (12.6)	45 (10.5)	0.5 (0.23-1.10)	0.071

Abbreviation: EAEC, enteroaggregative *Escherichia coli*, OR, odds ratio, CI, confidence intervals

Table 4.2: Distribution of EAEC virulence genes from cases and controls children (age 0-59 month)

Gene Class		Virulence Gene	Case (n=157) No. (%)	Control (n=271) No. (%)	Total (n=428) No. (%)	Odd Ratio (95% CI)	X^2	P-value
pAA PLASMID		***aatA***	51 (32.5)	82 (30.3)	133 (31.1)	1.1 (0.71-1.73)	0.2	0.631
		aggR	97 (61.8)	161 (59.4)	258 (60.3)	1.1 (0.72-1.69)	0.2	0.628
		aaP	77 (49.0)	121 (44.7)	198 (46.3)	1.2 (0.78-1.80)	0.8	0.379
		ORF3	106 (67.5)	168 (62.9)	274 (64.0)	1.3 (0.82-1.98)	1.3	0.251
		capU	108 (68.8)	158 (58.3)	266 (62.2)	1.6 (1.02-2.45)	4.7	0.031
		aar	110 (70.1)	188 (69.4)	298 (69.6)	1.0 (0.66-1.63)	0.1	0.880
	ADHESIN	***aafC***	7 (4.5)	16 (6.0)	23 (5.4)	0.7 (0.25-1.97)	0.4	0.522
		agg3/4C	58 (36.9)	96 (35.4)	154 (36.8)	1.1 (0.69-1.64)	0.1	0.752
		agg3A	10 (6.4)	28 (10.3)	38 (9.8)	0.6 (0.25-1.29)	1.9	0.164
		aafA	3 (1.9)	15 (5.5)	18 (4.2)	0.3 (0.06-1.20)	3.2	0.071
		aggA	47 (29.9)	62 (22.9)	109 (25.5)	1.4 (0.89-2.29)	2.6	0.106
		agg4A	15 (9.6)	16 (6.0)	31 (7.2)	1.7 (0.74-3.75)	1.9	0.160
	TOXINS	***astA***	91 (58.6)	129 (47.6)	220 (51.4)	1.5 (1.00-2.30)	4.3	0.038
		sat	29 (18.5)	56 (20.7)	85 (19.9)	0.9 (0.51-1.47)	0.3	0.583
		sepA	50 (31.9)	62 (22.9)	112 (26.2)	1.6 (0.99-2.49)	4.1	0.041
		pet	24 (15.3)	24 (8.9)	48 (11.2)	1.9 (0.97-3.56)	4.1	0.042
CHROMOSOME		***pic***	55 (35.0)	88 (32.5)	143 (33.4)	1.1 (0.72-1.73)	0.3	0.588
		sigA	18 (11.5)	31 (11.4)	49 (11.5)	1.0 (0.50-1.93)	0.0	0.993
		aaiC	44 (28.0)	97 (35.8)	141 (32.9)	0.7 (0.44-1.09)	2.7	0.099
		air	41 (26.1)	57 (21.0)	98 (22.9)	1.3 (0.81-2.15)	1.5	0.227
		eilA	79 (50.3)	128 (47.2)	207 (48.4)	1.1 (0.75-1.71)	0.4	0.538

The distribution of the characterized virulence genes varied across the age strata. In 0-11 month stratum, prevalence of *pet* (OR 6.9, 95% CI 2.06-29.20, P < 0.001), *aggA* (OR 2.2, 95% CI 1.16-4.29, P = 0.008), and *capU* (OR 1.9, 95% CI 1.02-3.51, P = 0.028) genes were more common in cases compared to controls (Table 4.3). Similar higher prevalence pattern was observed for *pet* (OR 15.0, 95% CI 1.35-750.0, P = 0.003) and *capU* (OR 4.3, 95% CI 1.27-18.54, P = 0.009) when the virulence factors were characterized among the sole EAEC pathogen from MSD children 0-11 month age in cases and controls (Table 4.4).

Prevalence of virulence genes that were proportionately higher in cases compared to controls in children 0-11 months were *sepA* (36.5% vs 26.5%), *astA* (54.1% vs 41.7%), *aggR* (71.8% vs 62.1%), *aap* (56.5% vs 44.7%) and ORF3 (75.3% vs 63.6%). The *astA* gene was found more often in cases (67.2%) than in controls (49.5%) in the age stratum 12-23 months (OR 2.1, 95% CI1.03-4.27, P = 0.026); none of the putative virulence factors were found to be significantly more common in MSD children ≥ 2 years of age (Table 4.3). Furthermore, results obtained from the characterisation of EAEC pathotypes that are sole pathogen among younger children showed significant association of *pet* gene with diarrhoea with OR 15.0, 95% CI 1.35-750, P = 0.003 (Table 4.4).

In addition to considering each virulence factor individually, we also considered the importance of combinations of potential EAEC virulence factors by employing classification and regression tree (CART) analysis. The CART analysis builds a model in stepwise fashion to yield the combination of factors most strongly associated with the queried outcome, in this case the combinations of factors most strongly associated with MSD. Each branch of a CART output tree ends in a terminal "node"; each

observation falls into exactly 1 terminal node; and each terminal node is uniquely defined by a set of rules, such as having or not having a certain factor.

We examined all 21 virulence genes including *aatA, aggR, aaiC, aap*, ORF3, *sat, sepA, pic, sigA, pet, astA, aafC, agg3/4C, aafA, agg3A, aggA, agg4A*, *air, capU, eilA, aar* as well as considering the collective number of virulence loci present (generating a virulence factor score, VFS) (figure 4.1 and 4.2).

As noted, prevalence of the virulence genes were significantly higher in cases compared to controls in children 0-11 months and applying the CART analysis (Figure 2) showed that the presence of *pet (Node 1)*, regardless of the presence or absence of any other scored genotype among the *pet*-positive strains, provided a strong association with diarrhoea. Among the *pet*-negative strains, CART analysis suggested two additional trait clusters that were associated with moderate to severe diarrheal: Node 2 includes those strains with a VFS <=8 in combination with *sepA,* whereas Node 3 includes a VFS > 8, suggesting a combination of typical EAEC factors in addition to the toxin EAST-1 toxin.

Table 4.3: Distribution of EAEC virulence genes in case and control children in three age strata

Virulence genes	0-11 months (N = 217)				12-23 months (N= 166)				24-59 months (n = 45)			
	Case (n = 85) No. (%)	Control (n = 132) No. (%)	OR (95% CI)	P-Value	Case (n = 61) No. (%)	Control(n =105) No. (%)	OR (95% CI)	P-Value	Case (n = 11) No. (%)	Control(n =34) No. (%)	OR (95% CI)	P-Value
aatA	33 (38.8)	48 (36.4)	1.1 (0.60-2.02)	0.714	16 (26.2)	26 (24.8)	1.1 (0.48-2.34)	0.833	2 (18.2)	8 (23.5)	0.7 (0.06-4.71)	0.710
aggR	61 (71.8)	82 (62.1)	1.5 (0.82-2.93)	0.143	33 (54.1)	62 (59.1)	0.8 (0.41-1.62)	0.534	3 (27.3)	17 (50.0)	0.4 (0.05-1.95)	0.187
aaP	48 (56.5)	59 (44.7)	1.6 (0.89-2.88)	0.090	28 (45.9)	46 (43.8)	1.1 (0.54-2.15)	0.793	1 (9.1)	16 (47.1)	0.1 (0.00-0.98)	0.024
ORF3	64 (75.3)	84 (63.6)	1.7 (0.91-3.37)	0.071	37 (60.7)	63 (60.0)	1.0 (0.51-2.06)	0.933	5 (45.5)	21 (61.8)	0.5 (0.10-2.53)	0.341
capU	59 (69.4)	72 (54.6)	1.9 (1.02-3.51)	0.028	42 (68.9)	62 (59.1)	1.5 (0.75-3.18)	0.208	7 (63.6)	24 (70.6)	0.7 (0.14-4.20	0.665
aar	62 (72.9)	98 (74.2)	0.9 (0.48-1.82)	0.831	42 (68.9)	70. (66.7)	1.1 (0.53-2.32)	0.772	6 (54.6)	20 (58.8)	0.8 (0.17-4.24)	0.802
aafC	3 (3.5)	7 (5.3)	0.7 (0.10-2.96)	0.543	3 (4.9)	6 (5.7)	0.9 (0.13-4.18)	0.827	1 (9.1)	3 (8.8)	1.0 (0.01-14.6)	0.978
agg3/4C	34 (40.0)	51 (38.6)	1.1 (0.58-1.91)	0.840	20 (32.8)	33 (31.4)	1.1 (0.50-2.19)	0.856	4 (36.4)	12 (35.3)	1.0 (0.18-5.18)	0.948
agg3A	3 (3.5)	18 (13.6)	0.2 (0.04-0.83)	0.014	5 (8.2)	8 (7.6)	1.1 (0.26-3.96)	0.893	2 (18.2)	2 (5.9)	3.6 (0.22-53.6)	0.212
aafA	1 (1.2)	3 (2.3)	0.3 (0.01-6.51)	0.557	2 (3.3)	10 (9.5)	0.3 (0.03-1.59)	0.134	0 (0)	2 (5.9)	0.0 (0.00-16.8)	0.410
aggA	32 (37.7)	28 (21.2)	2.2 (1.16-4.29)	0.008	13 (21.3)	25 (23.8)	0.9 (0.37-1.95)	0.711	2 (18.2)	9 (26.5)	0.6 (0.05-3.94)	0.578
agg4A	10 (11.8	12 (9.1)	1.3 (0.48-3.55)	0.524	4 (6.6)	2 (1.9)	3.6 (0.49-40.7)	0.121	1 (9.1)	2 (5.9)	1.6 (0.02-33.4)	0.710
astA	46 (54.1)	55 (41.7)	1.7 (0.91-2.96)	0.072	41 (67.2)	52 (49.5)	2.1 (1.03-4.27)	0.026	4 (36.4)	22 (64.7)	0.3 (0.05-1.56)	0.098
sat	20 (23.5)	24 (18.2)	1.4 (0.66-2.84)	0.338	9 (14.8)	24 (22.9)	0.6 (0.22-1.43)	0.207	0 (0)	8 (23.5)	0.0 (0.00-1.69)	0.076
sepA	31 (36.5)	35 (26.5)	1.6 (0.84-2.86)	0.119	16 (26.2)	23 (21.9)	1.3 (0.56-2.79)	0.526	3 (27.3)	4 (11.8)	2.8 (0.33-20.1)	0.217
pet	15 (17.7)	4 (3.0)	6.9 (2.06-29.20)	<0.001	9 (14.8)	16 (15.2)	1.0 (0.34-2.51)	0.933	0 (0)	4 (11.8)	0.0 (0.00-4.78)	0.233
pic	24 (28.2)	34 (25.8)	1.1 (0.58-2.18)	0.687	28 (45.9)	41 (39.1)	1.3 (0.66-2.63)	0.387	3 (27.3)	13 (38.2)	0.6 (0.08-3.18)	0.509
sigA	4 (4.7)	5 (3.8)	1.3 (0.24-6.01)	0.740	10 (16.4)	20 (19.1)	0.8 (0.32-2.04)	0.668	4 (36.4)	6 (17.7)	2.7 (0.42-15.0)	0.194
aaiC	22 (25.9)	34 (25.8)	1.0 (0.51-1.95)	0.983	19 (31.2)	48 (45.7)	0.5 (0.25-1.09)	0.065	3 (27.3)	15 (44.1)	0.5 (0.07-2.47)	0.321
air	26 (30.6)	34 (25.8)	1.3 (0.66-2.32)	0.437	13 (21.3)	21 (20.0)	1.1 (0.45-2.50)	0.840	2 (18.2)	2. (5.9)	3.6 (0.22-53.6)	0.212
eilA	37 (43.5)	54 (40.9)	1.1 (0.61-2.00)	0.702	35 (57.4)	52 (49.5)	1.4 (0.69-2.72)	0.328	7 (63.6)	22 (64.7)	1.0 (0.19-5.38)	0.948

Table 4.4: Characterization of virulence factor in the EAEC pathotype obtained as sole pathogen from cases and control children age 0-11 month

Virulence Gene	0-11 months (n=105)					
	Case (n=25) No. (%)	Control (n=80) No. (%)	Total (n=105) No. (%)	Odd Ratio (95% CI)	X^2	P-value
sat	7 (28.00)	15 (18.75)	22 (20.95)	1.69 (0.49-5.23)	1.0	0.32
sepA	7 (28.00)	22 (27.50)	29 (27.62)	1.03 (0.32-3.03)	0.0024	0.96
Pic	4 (16.00)	17 (21.25)	21 (20.00)	0.71 (0.16-2.52)	0.3	0.57
sigA	2 (8.00)	1 (1.25)	3 (2.86)	6.87 (0.33-410.2)	3.1	0.08
Pet	4 (16.00)	1 (1.25)	5 (4.76)	15.0 (1.35-750.0)	9.14	0.003
astA	12 (48.00)	34 (42.50)	46 (43.81)	1.25 (0.46-3.38)	0.2	0.63
aatA	11 (44.00)	31 (38.75)	42 (40.00)	1.24 (0.45-3.38)	0.2	0.64
aggR	19 (76.00)	52 (65.00)	71 (67.62)	1.71 (0.57-5.80)	1.1	0.30
aaiC	7 (28.00)	17 (21.25)	24 (22.86)	1.44 (0.43-4.39)	0.5	0.48
aaP	15 (60.00)	36 (45.00)	51 (48.57)	1.83 (0.67-5.13)	1.7	0.19
orf3	21 (84.00)	49 (61.25)	70 (66.67)	3.32 (0.98-14.42)	4.4	0.04
aafC	0 (0.00)	4 (5.00)	4 (3.81)	0.00 (0.00-3.07)	1.3	0.25
agg3/4C	11 (44.00)	26 (32.50)	37 (35.24)	1.63 (0.58-4.48)	1.1	0.29
agg3A	0 (0.00)	10 (12.50)	10 (9.52)	0.00 (0.00-1.12)	3.5	0.06
aafA	0 (0.00)	1 (1.25)	1 (0.95)	-	0.3	0.57
aggA	10 (40.00)	18 (22.50)	28 (26.67)	2.30 (0.77-6.57)	3.0	0.08
agg4A	3 (12.00)	7 (8.75)	10 (9.52)	1.42 (0.22-6.89)	0.2	0.63
air	10 (40.00)	18 (22.50)	28 (26.67)	2.30 (0.77-6.57)	3.0	0.08
capU	21 (84.00)	44 (55.00)	65 (61.90)	4.29 (1.27-18.54)	6.8	0.009
eillA	7 (28.00)	34 (42.50)	41 (39.05)	0.52 (0.17-1.52)	1.7	0.20
orf61	19 (76.00)	59 (73.75)	78 (74.29)	1.12 (0.37-3.92)	0.05	0.82

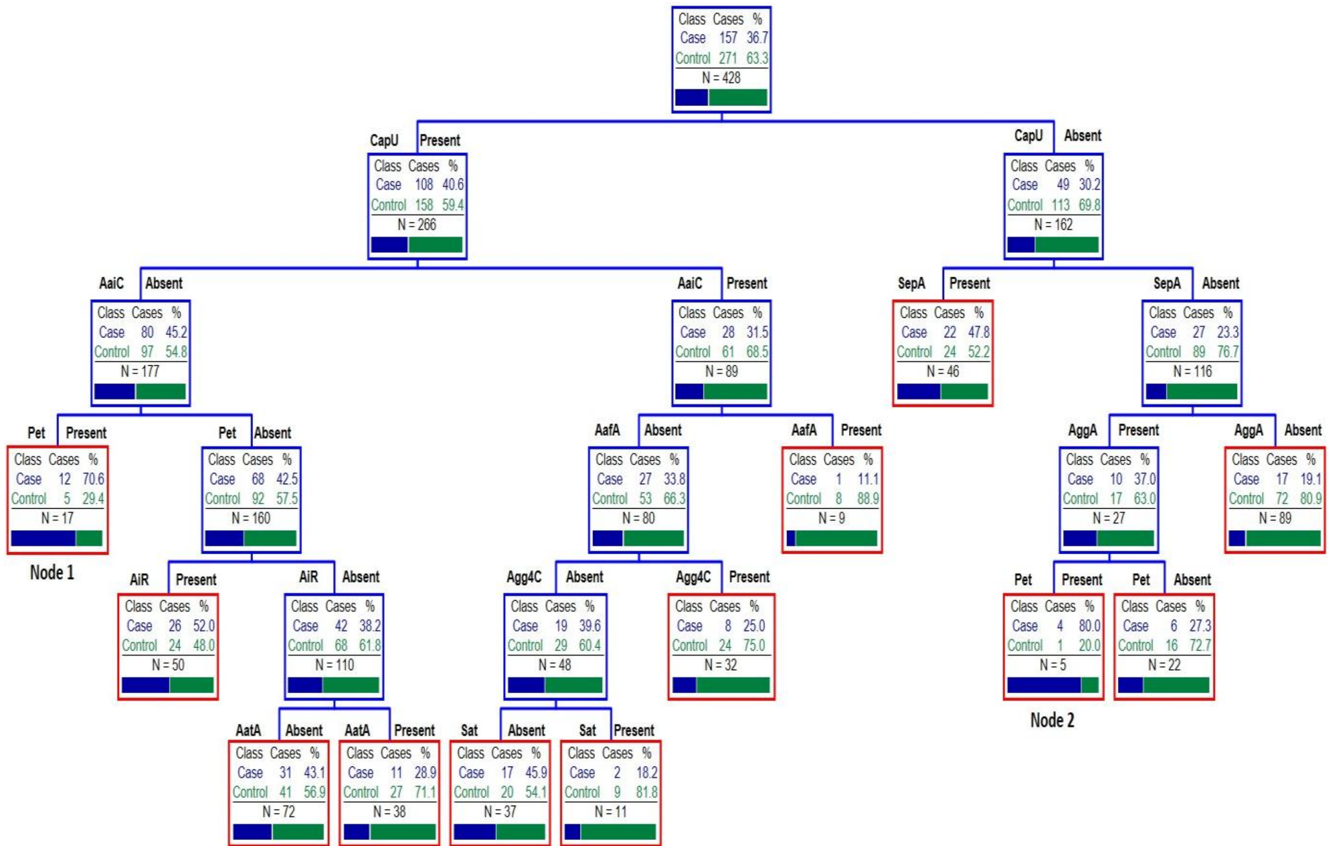


Figure 4.1: Association of virulence factors with diarrhoea in children aged 0-59 months. Each branch of the classification and Regression Tree ends in a terminal "node" (red boxes), and each terminal node is uniquely defined by the presence or absence of a predictive factor such as a gene or virulence factor score

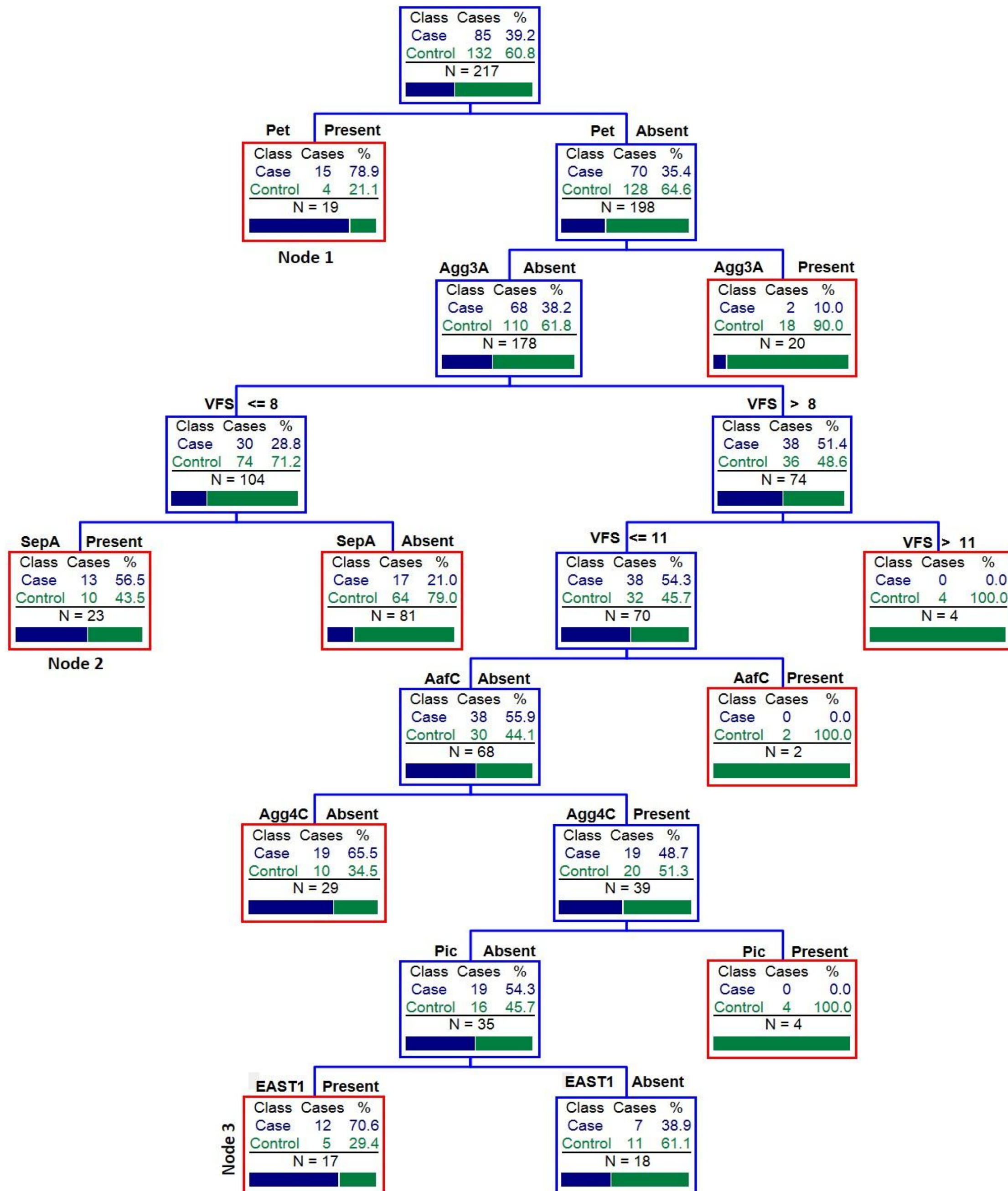


Figure 4.2: Association of virulence factors with diarrhoea in children aged 0-11 months. Each branch of the classification and Regression Tree ends in a terminal “node” (red boxes), and each terminal node is uniquely defined by the presence or absence of a predictive factor such as a gene or virulence factor (VFS)

4.3 Discussion

EAEC is a common cause of diarrhea worldwide (Cennimo, Abbas et al. 2009). The assessment of the 21 genes in the 428 EAEC strains in this study showed that the frequency of most genes correlated well with similar studies, particularly the study from the GEMS Gambia neighboring site, in Bamako Mali (Boisen, Scheutz et al. 2012). In this study, more than half of the study participants were younger than 1 year of age, although there were no statistical differences between cases and controls in 86% of the virulence genes using the p-value. However, when odd ratio (OR), a measure of association is used to assess the association of virulence factors with diarrhoea among the children 0-11 month, we found that over 80% of the virulence genes were associated with diarroeal disease by having OR >1. And a sharp contrast of association was observed among the older age children 24-59 month with only 19% of the virulence genes been associated with diarrhoeal disease having OR >1. Similarly, among 12-23 month old children 52% of the virulence genes with OR >1 were found to be associated with diarrhoea The explanation to this difference is that younger age children (infancy) are more susceptible to EAEC infection majorly due to their immature innate and adaptive immune system, which matures and acquires memory as they grow. This speculation has been corroborated in a study (Philipson Bassaganya et al. 2013) modelling immunity to EAEC which showed the importance of Th17 cells in host responce to EAEC facilitating bacteria clearance. The discovery followed the initial EAEC T cell differentiation model that depicts EAEC infection, antigen presentation, and host adaptive immune response to pathogen. This probably explains stronger immunity with acquisition of memory immune cells of previous EAEC and other enteric infection in the older age children. Therefore, there was a decline in the proportion of virulence genes association with diarrhoea using OR >1 from infant to

older age children (table 4.3). Association between *sepA* gene and MSD in Mali study is stronger compared to this study, and the reason for the difference is unclear but it can be genetic and or environmental factor. SepA is a SPATE protease that was initially found in *S. flexneri* strains (Benjelloun-Touimi, Sansonetti et al. 1995), but has subsequently been found commonly among EAEC (Boisen, Ruiz-Perez et al. 2009). The protease has been implicated in causing increased inflammation in Shigella strains but it may also have enterotoxic activity.

In this study, the virulence genes *aggA* encoding for *AAF/1*, *capU* and *pet*, encoding a member of Class 1 serine protease autotransporters of Enterobacteriaceae (SPATEs) family, were statistically implicated as genes responsible for EAEC diarrhoea in younger children < 12 months.

Our study highlights significant heterogeneity in gene profiles among the EAEC isolates. Of the twenty-one genes targeted, none of the EAEC isolates characterized genetically harbours more than 15 virulence genes. The heterogeneous nature of EAEC enables it to display variation in causing clinical illness, (Cennimo, Abbas et al. 2009) although factors responsible for its virulence are not well understood.

Several studies have shown possible genes that confer virulence on EAEC (Cennimo, Abbas et al. 2009; Opintan, Newman et al. 2010). Our data show three virulence genes associated with diarrhoea in infants. Interestingly, the three incriminated virulence genes are plasmid genes that include plasmid-encoded toxin (*pet*), AAF/1 fimbrial subunit (*aggA*) and hexosyltransferase homolog (*capU*). The Pet toxin is a 108-kDa protease, which secretes enterotoxin that generates high toxicity in human epithelial cells resulting in structural damage to the cell. Following internalization via receptore-mediated endocytosis, pet is delivered to the cytoplasm by means of retrograde

trafficking which is accompanied by cleavage of spectrin known as actin-binding protein fodrin, within microvilli cytoskeleton leading to cell elongation, rounding and finally, the release of cells from the substractum (Navarro-Garcia, Sears et al. 1999; Villaseca, Navarro-Garcia et al. 2000; Navarro-Garcia, Canizalez-Roman et al. 2001; Dutta, Cappello et al. 2002). In Mexico, the *Pet* gene was initially detected from EAEC strain 049766 implicated in a highly virulent outbreak of diarrhoea in which some infants died (Eslava 1993). Also, the reported enterotoxic activity of EAEC induced by Pet is consistent with the secretory diarrhea seen in most patients with EAEC enteritis (Eslava, Navarro-Garcia et al. 1998). A recent report from Iran alluded that pet gene is more prevalent among EAEC strains isolated from adult diarrhoeal patients (Bafandeh, Haghi et al. 2015). Therefore, our findings support the role of *Pet* gene in EAEC causing diarrhea in infants (figure 4.1; Table 4.3). However, earlier EAEC virulence factor study conducted in Southwest Nigeria over a decade ago showed that the Pet gene was equally distributed among EAEC strains isolated from children <5 years with or without diarrheal (Okeke, Lamikanra et al. 2000). Seemingly, our study also showed, no association of Pet with diarrheal disease in the children <5 years but the effect is only seen in EAEC strains isolated from children < 1 year and so the differences between our findings could be due to age stratification, which again explains the poor status of infants adaptive immune system, permitting *pet* enterotoxin to proliferate mucosal epithelial cells leading to diarrhoea in infant.

Generally, innate immune system provides an early first line defence against invading pathogens by involving cells that include neutrphils, monocytes, macrophages and dendritic cells, which all interact with the adaptive immune system. So at birth, immune system is muted in order for the foetus to tolerate only non-shared maternal antigens and to avoid high level of stress and remodelling that takes place during development.

This makes the newborn baby relatively susceptible to bacterial and viral infections. In the adaptive immune system, T cells develop in the thymus which is largest at birth and during first years of life. The function of early-life T cells is different from adult T cells. B cells are present in secondary lymphoid organs and in the bone marrow, they contribute to humoral response of the adaptive immune system. So, most antibody responses, including those to bacterial proteins, bacterial polysaccharides and to polysaccharide-protein conjugate vaccines are dependent on T-cell help.

A study showed that EAEC *pet* gene participates as an immunostimulant molecules for machrophages, which activates both their mobility and cytokine expression (Rocha-Ramirez, Hernandez-Chinas et al. 2016). Prior this study, our understanding was that only *AAF* variants and other virulent genes participate in the activation of early inflammatory response without the participation of *pet* gene which explains why host T-helper cell does not have memory of *pet* antibody.

AAF/I was associated with diarrhoea in the first year of life, also, this study showed significant association of *aggA* with diarrhoea in younger children. Therefore, there is likelihood of synergistic interaction between the enterotoxin producing gene (*pet*) and the adherence factor *aggA* that codes for *aaf/I* to cause damage to mucosal epithelial cells leading to diarrhoeal disease and potentially malnutrition among younger age children. Again, the Shiga toxin producing EAEC strain implicated in the German outbreak expressed *AAF/I* (Scheutz, Nielsen et al. 2011).

Hexosyltransferase homolog (*capU*), a plasmid-encoded protein was significantly high among the younger children. Its role in EAEC diarrhoea is not clearly defined. Notably, the *capU* gene was the third most common gene found (62%) among genes

investigated in this study. This probably highlights the importance of genes acting in concert.

astA encodes a heat-stable enterotoxin (EAST1) that is related to the heat-labile enterotoxin of enterotoxigenic *E. coli*. The relevance of *astA* gene in EAEC diarrhoea has been reported in several studies (Vila, Gene et al. 1998; Paiva de Sousa and Dubreuil 2001; Toshima, Uenaka et al. 2004; Veilleux, Holt et al. 2008), and EAST1 was found to be associated with diarrhoea in combination with other genes in the Mali study (Boisen, Scheutz et al. 2012). *astA* is not restricted to EAEC but is widely distributed among other enteric pathogens (Menard and Dubreuil 2002; Zhou, Ogasawara et al. 2002), as well as commensal *E. coli*.

Classification Regression Tree (CART) Analysis: Following the proportional and statistical analysis of individual virulence factor, CART analysis was employed to investigate combinations of the potential EAEC virulence factors. In children 0-59 months, CART analysis showed the significance association of *pet* with diarrhoea in the presence of *capU* and in the absence of *aaiC* (Node 1) while in the prensence of *capU* and in the absence of *aaiC* and *pet* the *air* and *aatA* genes were not associated with diarrhoea (figure 4.1). Similarly in the same cluster, the *aafA*, *agg4C* and *sat* gene were not associated with diarrhoea in the presence of *capU* and *aaiC*. In the cluster that had *capU* present *sepA* is not associated with disease. Also, when *capU* and *sepA* were absent *aggA* was not associated with disease but *pet* was significantly associated with diasese (Node 2) (figure 1). In children 0-11 months, CART analysis utilised both the combination of virulence factors and virulence factor score (VFS) (figure 4.2). CART analysis among the younger age children (0-11 months) distinctly showed *pet* significant association with diarrhoeal disease (Node 1). While the absence of *pet* and absence of *agg3A* with equal or less than 8 VFS (<=8 vfs) *sepA* was found to be

associated with disease (Node 2). In the absece of pet gene *Agg3A* did not show any association with diarrhoea. Surprisingly, *EAST*1 showed significant association with diarrhoea with VSF<=11 in the absence of many virulence related genes. The common observation from the two trees ploted is the importance of *pet* gene among the diarrhoeal infants in the rural Gambia.

Microbiome: Intestinal microbiome plays a pivotal role in preventing infectious diseases as early as birth (Harris, Haak et al. 2017). But it is unclear whether microbiome play a role in the expression of virulence genes causing disease. An infant may first be exposed to bacteria as early as in utero and upon delivery undergoes rapid intestinal colonisation. The patterns of colonisation are in part non-random and can be shaped by mode of delivery, breastfeeding, geography, genetics, antibiotics and age (Yatsunenko, Rey et al. 2012), and it is evidence that specific bacterial colonisation is required for normal neonatal immune development (Chung, Pamp et al. 2012). Therfore, microbes are important symbiotic modulators of physiological, metabolic and immunological function in the mammalian host (Durack and Lynch 2019). Recent data indicate that the developing gut microbiota of human infants affects the progression of intestinal mucosal IgA responses, and perturbations to these nascent microbial communities cause long-lasting metabolic and immunological dysregulation (Ruiz et al., 2017; Lynn et. al. 2018). For example, the study showed macrolide treatment of conventional, neonatal mice perturbs their gut microbiota with depletion of Bifidobacterium and segmented filamentous bacteria, resulting in decreased numbers of intestinal CD4+IL-17A+ lymphocytes and faecal IgA concentrations. Similar events may have happened in this characterisation study of EAEC where host natural antibiotics or administered antibiotics might have possibly cause down-regulation of expression of some virulence genes. Factually, a study cataloguing functional genes in

the human gut microbiome identified as many as 9.9 million unique microbial genes across 1,200 healthy subjects from three different continents (Li et al. 2014).

Globally, EAEC strains have shown a low to high level of resistance to antimicrobial agents (Mendez Arancibia, Pitart et al. 2009). Our data from the antimicrobial susceptibility investigation highlights high resistant pattern of the EAEC strains against Cotrimoxazole, and Ampicillin. The first line of antibiotics prescribed for patient management in our region are Cotrimoxazole and Ampicillin, which may explain the high resistance against these antibiotics. An increase in resistance of EAEC strains to Chloramphenicol, Nalidixic acid and Quinolones was observed in this study compared to a similar study on a member of enterobacteriaceae family from the same region (Ikumapayi, Antonio et al. 2007) and in eastern Asia (Aslani, Alikhani et al. 2011). Twenty percent of the EAEC strains tested showed multidrug resistance to 3 antimicrobial agents whereas six percent showed resistant to more than 3 antimicrobial agents. This finding is in contrast to a similar study conducted in India, showing 75% of strains with multidrug resistance, i.e. > 3 antimicrobial agents (Raju and Ballal 2009).

The limitations of this study included exclusion of multiple comparisons such as malnutrition and other enteric co-infections. Hence future studies can consider these essential confounders.

4.4 Conclusion

Our study has strengthened the role of *pet* and EAST1 genes of EAEC in the cause of MSD in African infants. The EAEC virulence gene profiles found in this study have also proven the heterogeneity of the genetic component of the EAEC isolates studied. However, further

investigations are needed to establish the specific or combination of gene(s) that are associated with EAEC diarrheal in different age strata, particularly children from developing countries

Chapter 5: A quantitative assessment of clinical significance of EAEC in diarrhoea using TaqMan-QPCR

5.1 Introduction

Diarrhoea is a predominant cause of childhood illness and mortality particularly in developing countries (Okeke 2009). Enteroaggregative *E. coli* (EAEC) is best known for causing acute and persistent diarrhoeal illness in developing countries as well as in travellers and immunocompromised individuals (Harrington, Dudley et al. 2006; Boll, Struve et al. 2013). Major diarrhoeal studies have implicated Enteroaggregative *E. coli* (EAEC) strain among the most important etiologic agents of diarrhoea both in industrialised and low income countries (Pabst, Altwegg et al. 2003; Cohen, Nataro et al. 2005; Rappelli, Folgosa et al. 2005; Nataro, Mai et al. 2006; Boisen, Scheutz et al. 2012). EAEC in diarrhoeal outbreaks is a more common occurrence than ever before and in large scale studies EAEC has been the commonest bacterial pathogen identified in diarrhoeal stool samples (Croxen, Law et al. 2013; Ikumapayi 2014). EAEC infection has been associated with severe intestinal inflammation leading to childhood malnourishment and growth impairment (Steiner, Lima et al. 1998; Roche, Cabel et al. 2010), although, little is known about this claim in West-Africa countries. However, the recent report of an outstanding observational diarrhoea study showed no association of EAEC with diarrhoeal among moderate-to-severe diarrhoea (MSD) children from developing countries particularly those from the Gambia (Kotloff, Nataro et al. 2013).

EAEC is a heterogeneous category of an emerging enteric pathogen (Kaur, Chakraborti et al. 2010). The difficulty in the diagnosis of EAEC causing diarrhoea is inherent in its heterogeneity since strains are found equally in both symptomatic and asymptomatic individuals and/or even more in asymptomatic children (Kotloff, Nataro et al. 2013). To

elucidate our understanding of the cause of diarrhoeal episodes among children from a developing country we seek to evaluate causal relationship between the bacterial quantity and the diarrhoeal symptoms.

This study is nested to a large case-control study (GEMS) where samples for this study were obtained. In order to determine the clinical relevance of the presence of EAEC in a faecal sample from diarrheic individuals, we hypothesize that high bacterial load of EAEC is associated with diarrhoeal disease among children under five. The marker/target genes for the detection of EAEC infection and quantitative-qPCR assays are *aaiC* (a chromosomal) and *aatA* (a plasmid) gene based on GEMS recommendation (Panchalingam, Antonio et al. 2012). The *aaiC* gene (*aggR*-activated island C) encodes a type VI secretion system that is located in a conserved chromosomal sequence in EAEC DNA (Dudley, Abe et al. 2006). The *aatA* (anti-aggregation protein transporter A) gene composes of EAEC-ABC transporter-A which consists of a DNA fragment from the EAEC plasmid which encodes an outer membrane protein of the ABC transporter complex (Lima, Boisen et al. 2013). We target both *aaiC* and *aatA* genes for PCR and qPCR assays in our investigation.

5.2 Result

The Taqman qPCR assay was performed in duplicate on two different days and similar results were obtained from the 160 (80 cases and 80 controls) samples. The average of the obtained results was used in the final analysis.

There was a huge difference in the qPCR detection of *aaiC* and *aatA* among the 160 samples investigated. The two genes combined were qPCR detected and measured in 24 (15%) of the 160 DNA samples, whilst *aaiC* and *aatA* alone were qPCR detected and measured in 16 (10%) and 120 (75%) of the 160 DNA samples respectively. So, we are obligated to choose *aatA* that has 90% qPCR detection and quantification for the analysis of the bacterial load data.

The cut-offs for the high bacterial load (HBL) and low bacterial load (LBL) has been explained in detail in chapter three 'materials and method section' of this thesis. In summary, cut-offs were achieved by constructing a slope from the point where Cq 16 and Cq 38 intercept, to the point it intercept the slope of the standard curve. At the later intercept point Cq 32 is obtained (figure 3.7a). Therefore, detection of a Cq ≥33 is considered low bacterial load (LBL) while detection of a Cq ≥16≤32 is considered high bacterial load (HBL) (figure 3.7a).

Of the 160 samples, 106 (48 [30%] cases and 58 [36.2%] controls) had high bacterial load (HBL) EAEC, whilst, 54 (32 [20%] cases and 22 [13.8%] controls) account for low bacterial load (LBL) EAEC. Proportionately, higher bacterial load EAEC is less in MSD children compared with community matched control.

The crude association of high bacterial load and diarrhoea for the MSD and non-MSD children showed (OR 0.61, 95% CI [0.3132281-1.182877] and p-value 0.143) (Table 5.1).

The univariable analysis showed association of some confounders in the presence of HBL, these include cows and rodents in the household, usage of well-water, presence of a co-infection, presence of other animals that include either a horse, donkey, dog or cat in the household, underweight and lower score Mid Upper Arm Circumference (MUAC) with having EAEC high bacterial load that results in diarrhoea to be statistically significant among MSD children (table 5.1). In the final model, presence of cows in a household, presence of rodents, presence of having a co-infection, underweight and lower score MUAC were found to be the most significant factors associated with having HBL that results in diarrhoea (Table 5.2).

Table 5.1: Univariable analysis showing effect of each confounder on the Association between being a case/control and High bacterial load (HBL)

Variable	Odd Ratio	95% CI	p-value
High Bacterial load (HBL)	0.61	0.3132281 - 1.182877	0.143
Cow + HBL	5.50	1.87856 - 16.11824	0.002
Fowl + HBL	1.93	0.4743 - 7.891345	0.358
Rodent + HBL	3.13	1.396007 - 6.996583	0.006
*Domestic animals + HBL	1.04	0.3310954 - 3.282651	0.943
**Other animals + HBL	3.98	1.478566 - 10.70114	0.006
Well water + HBL	4.80	1.252331 - 18.40323	0.022
Breastfeed + HBL	1.65	0.5441415 - 5.016002	0.376
Co-infection + HBL	2.23	1.125549 - 4.418939	0.022
Underweight + HBL	4.27	1.870679 - 9.742075	0.001
Lower-score-MUAC + HBL	3.58	1.610876 - 7.965095	0.002

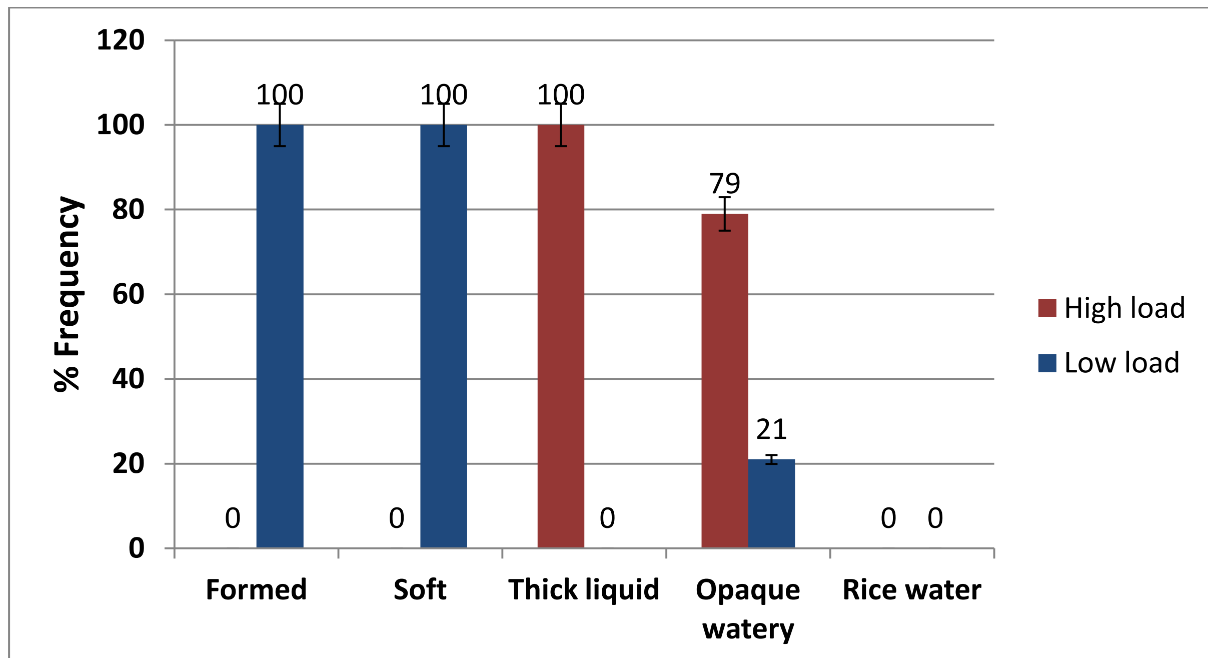
Table 5.2: Multivariable analysis showing effect of all the confounders on the Association between being a case/control and High bacterial load

Variable	Odd Ratio	95% CI	p-value
High Bacterial load	0.45	0.1240591 - 1.668576	0.235
Cow	11.83	1.664311 - 84.03794	0.014
Fowl	2.23	0.1151228 - 43.08074	0.596
Rodent	6.96	1.725073 - 28.10618	0.006
*Domestic animals	0.99	0.1172411 - 6.905467	0.919
**Other animals	2.16	0.5063242 - 8.513867	0.310
Well water	4.33	0.4171896 - 44.95957	0.220
Breastfeed	1.50	0.2571136 - 8.786691	0.651
Co-infection	3.20	1.014459 - 10.10321	0.047
Underweight	4.09	1.036919 - 16.10511	0.044
Lower-score-MUAC	4.94	1.185023 - 20.59792	0.028

*Domestic animals (Goat and Sheep) and ***Other animals (Donkey, Horse, Dog and Cat)

Proportional distribution of stool consistency among MSD children with sole infection of EAEC showed that formed, soft, thick-liquid, opaque-watery and rice-water with high bacterial load account for 0%, 0%, 100%, 79% and 0% respectively, and low bacterial load account for 100%, 100%, 0%, 21% and 0% respectively (figure 5.1). One death was recorded among the high bacterial load MSD children with stool consistency opaque-watery (figure 5.1).

Figure 5.1: proportional distribution of stool consistency in EAEC sole infection among MSD with high and low bacterial load



We investigated a set of earlier detected 21 virulence factors (Ikumapayi, Boisen et al. 2017) for possible association on high or low bacterial load from among MSD and non-MSD children using pathogenicity index (in this case virulence index) calculation. Interestingly, results showed that only *pet* gene was associated with high bacterial load yielding pathogenecity index (PI) 9.27, odd ratio 12.3 and p-value 0.005. The only gene associated with low bacterial load was *astA* gene yielded PI 2.67, odd ratio 6.0 and p-value 0.040 (table 5.3).

Table 5.3: Analysis showing virulence/pathogenicity index of each virulence factor on high and low bacterial load among MSD and non-MSD children age less than 5 years old (n=93)

Virulence genes	High bacterial load (n=69)			Low bacterial load (n=24)		
	Case (n=36) No. (%)	Control (n=33) No. (%)	Pathogenicity / Virulence index	Case (n=12) No. (%)	Control (n=12) No. (%)	Pathogenicity / Virulence index
aatA	13 (36.1)	13 (39.4)	0.92	4 (33.3)	6 (50)	0.67
aggR	25 (69.4)	23 (69.7)	0.99	6 (50)	10 (83.3)	0.60
aap	18 (50)	11 (33.3)	1.50	5 (41.7)	7 (58.3)	0.72
orf3	25 (69.4)	23 (69.7)	0.99	7 (58.3)	9 (75)	0.78
capU	25 (69.4)	19 (57.6)	1.20	8 (66.7)	6 (50)	1.33
orf61	26 (72.2)	25 (75.8)	0.95	9 (75)	8 (66.7)	1.12
aafC	1 (2.8)	0 (0)	NA	1 (8.3)	2(16.7)	0.49
agg3/4C	12 (33.3)	16 (48.5)	0.69	6 (50)	4 (33.3)	1.50
agg3A	2 (5.6)	6 (18.2)	0.31	2 (16.7)	1 (8.3)	2.01
aafA	1 (2.8)	1 (3.0)	0.93	0 (0.0)	0 (0.0)	NA
aggA	12 (33.3)	8 (24.2)	1.38	4 (33.3)	3(25)	1.33
agg4A	2 (5.6)	4 (12.1)	0.46	0	1 (8.3)	NA
astA	18 (50)	15 (45.5)	1.10	8 (66.7)	3 (25)	2.67
sat	8 (22.2)	6 (18.2)	1.22	2 (16.7)	3 (25)	0.67
sepA	14 (38.9)	14 (42.4)	0.92	1 (8.3)	4 (33.3)	0.25
**pet*	10 (27.8)	1 (3.0)	9.27	0 (0)	1 (8.3)	NA)
pic	15 (41.7)	9 (27.3)	1.53	4 (33.3)	5 (41.7)	0.79
sigA	6 (16.7)	2 (6.1)	2.74	2 (16.7)	3 (25)	0.67
aaiC	11 (30.6)	11 (33.3)	0.92	1(8.3)	5 (41.7)	0.20
air	10 (27.8)	12 (36.4)	0.76	5 (41.7)	4 (33.3)	1.25
eilA	13 (36.1)	8 (24.2)	1.50	6 (50)	6 (50)	1.0

*Distinct high pathogenicity index 9.27, odd ratio 12.3 and p-value 0.005

Also, we examined virulence factor score against bacterial load and other important variables that include age, sex, coinfection, stool consistency, type of water and type of animals in the study participant's house that are common to both symptomatic and asymptomatic children. The result showed no statistically significant association of virulence factor score with high bacterial load in both symptomatic and asymptomatic groups (table 5.4 & 5.5).

We report a **fatal case (SID100574) in table 5.4**. The child was infected solely with an EAEC high-bacterial-load infection and was malnourished the demographic characteristics and other details of the child were 10 month old, female, opaque watery stool consistency, had five types of animal (goat, sheep, cow, rodent and fowl) in the house backyard, the EAEC strain involved harbours 9 virulent factors that include *aatA*, *aap*, *orf*3, *orf*61, *agg*A, *ast*A, *sat*, *pet* and *air* among the 21 virulent factors investigated (table 5.4) and was not coinfected with other intestinal pathogens targeted.

Also, another fatal case (SID102095) in table 5.4 had EAEC high-bacterial-load coinfected with astrovirus only but not malnourished. The demographic characteristics and other details were 7 month old, male and opaque watery stool consistency, had 8 types of animal (goat, sheep, dog, cat, cow, rodent, fowl and donkey) in the house backyard and the EAEC strain involved harbours 3 virulent factors that include Orf61, agg3/4C and air.

Table 5.4: Virulent factor score against Bacterial load and other variables among MSD children (n=48)

Sample Identity (SID)	Demography		Bacteria Load	Co-infection	Water Source	Stool Consistency	Animal	Type of Animal Present	pAA plasmid																Chromosome					Virulence Factor Score
															Adhesin						Toxins									
	Age in Month	sex							aatA	aggR	aap	Orf3	capu	Orf61	aafC	Agg3/4C	Agg3A	aafA	aggA	Agg4A	astA	sat	sepA	pet	pic	sigA	aaiC	air	eilA	
100109	10	F	Lo	Y	Tp	Sf	g, s, c, f, d_o, h	**6**				■	■	■		■			■		■							■	■	8
100574	10	F	Hi	N	Tp	Op	g, s, c, r, f	**5**	■		■	■		■					■		■	■		■				■		9
102044	7	F	Hi	Y	we	Sf	r, f	**2**	■															■						2
102055	8	M	Hi	Y	we	TL	g, s, c_a, r, f, d_o, h	**7**		■	■	■	■			■					■			■	■		■		■	10
102105	4	M	Lo	Y	we	Op	g, s, d, c, f	**5**	■	■	■	■	■	■		■					■							■		9
102106	6	M	Hi	N	we	Op	g, s, r, f, d_o, h	**6**		■		■	■	■					■		■		■				■	■		9
102155	6	M	Hi	Y	Tp	Op	g, c_a, r, f, d_o	**5**	■				■	■					■				■					■		6
102444	8	M	Hi	Y	Tp	Op	g, c, r, f, d_o, h	**6**		■	■	■	■	■			■				■		■		■					9
102500	14	M	Hi	N	we	TL	g, d, r, f, d_o	**5**		■		■	■										■		■				■	6
103180	22	M	Lo	Y	we	TL	g, s, d, c, r, f, d_o	**7**		■		■	■				■					■			■	■	■			8
103193	11	M	Hi	N	we	Op	g, s, d, c_a, c, r, f d_o	**8**		■		■	■	■		■						■					■	■		8
103201	8	M	Hi	Y	we	Op	r, f, d_o	**3**	■	■	■	■	■	■	■			■			■			■	■	■				12
103240	8	F	Hi	Y	we	TL		**0**	■	■	■	■		■					■			■							■	8
103275	5	M	Hi	N	we	TL	g, s, d, r, f	**5**	■	■	■	■	■	■					■			■						■		9
103288	12	F	Hi	Y	we	Op	g, r, f, d_o	**4**		■	■	■	■										■						■	6
103448	10	F	Lo	Y	we	TL	g, s, c, r, f	**5**					■								■				■				■	4
100111	7	F	Lo	Y	Tp	Op	g, s, c_a, r, f, d_o, h	**7**	■	■	■	■	■	■		■					■	■							■	10
100513	20	F	Lo	Y	Tp	TL	g, s, c_a, r, f, d_o, h	**7**	■	■	■	■	■	■					■		■							■	■	10
102095	7	M	Hi	Y	we	Op	g, s, d, c_a, c, r, f d_o	**8**						■		■												■		3
102465	8	F	Hi	Y	we	Op	g, c_a, f	**3**	■				■	■							■			■	■					6
103039	13	M	Hi	Y	Tp	Op	g, s, r, f, d_o, h	**6**		■		■		■							■				■	■	■	■		8
103445	20	M	Hi	Y	we	Op	g,s,d, c_a, c,r,f,d_o, h	**9**		■	■		■								■			■	■	■				7
103467	9	F	Hi	Y	we	Op	g,s,d, c_a, c,r,f,d_o, h	**9**		■	■	■		■		■					■		■			■	■			9
103693	7	M	Hi	N	Tp	Op	g, s, c, r, f,	**5**	■	■	■	■	■	■		■			■	■	■									10
100119	7	M	Lo	Y	Tp	Op	s, d, c, r, f,	**5**			■				■	■					■		■						■	6
100232	14	F	Lo	N	Tp	Fo	g, s, c, r, f, d_o, h	**7**					■	■					■		■							■		5
100484	8	M	Lo	Y	Tp	Op	g, s, d, c, r, f, d_o, h	**8**						■															■	2
100679	6	M	Hi	N	we	Op	g, s, r, f, d_o	**5**					■	■		■					■	■		■	■			■		8
102258	4	F	Hi	Y	Tp	Op	g, s, f	**3**		■	■	■	■				■				■	■	■						■	9
102334	18	M	Lo	Y	we	Sf	g, s, c, r, f, h	**6**	■	■	■	■		■		■	■				■				■	■				10
102602	21	M	Hi	Y	we	Sf	g, s, c_a, r, f, d_o	**6**	■	■	■		■	■							■		■		■		■		■	10
100191	6	M	Hi	Y	Tp	Op	g, s, c, r, f, d_o, h	**7**		■		■	■	■								■	■				■		■	8
100890	6	M	Hi	Y	we	Op	R, f, h	**3**		■	■	■	■	■					■			■			■				■	9
102191	6	M	Hi	Y	we	TL	g, s, d, c_a, r, f, d_o,h	**8**		■	■	■		■					■		■			■	■		■			9
102192	10	M	Hi	Y	we	TL	g, s, c_a, r, f, d_o, h	**7**		■		■	■	■					■				■					■		7
102821	11	F	Hi	Y	Tp	Op	g, s, r, f, d_o	**5**		■	■	■	■										■							5
103276	43	M	Hi	N	Tp	Op	g, r, f, d_o	**4**	■				■	■		■										■			■	6
103663	16	M	Hi	N	we	Op	g, s, c_a, c, r, f	**6**					■																■	2
100313	5	M	Hi	N	Tp	TL	g, s, d, c, r, f, d_o, h	**8**	■	■		■	■	■		■					■								■	8
100794	10	M	Hi	Y	we	Op	g, s, d, r, f, d_o	**6**	■	■	■	■	■	■		■					■		■	■	■		■		■	13
100796	10	F	Hi	Y	we	Op	g, s, d, r, f, d_o	**6**				■				■				■			■		■					5
102090	24	F	Hi	Y	we	TL	g, s, c, r, f	**5**		■		■		■					■											4
103016	19	F	Hi	Y	Tp	Sf	g, s, r, f, d_o, h	**6**		■	■	■	■	■					■		■				■	■				9
100713	10	M	Hi	N	Tp	Op	g, s, r, f, d_o, h	**6**	■	■	■	■				■								■					■	7
102428	22	M	Lo	N	we	TL	g, c_a, c, r, f	**5**		■		■		■		■			■						■			■		7
102788	21	M	Hi	N	Tp	Op	g, d, r, f, d_o, h	**6**						■		■					■		■		■		■			6
102845	5	F	Hi	N	Tp	TL	g, s, f	**3**					■	■					■								■	■		5
103446	28	F	Lo	Y	we	TL	s, r, f, d_o	**4**					■	■																2

Key: **c** – cow, **c_a** – cat, **d** – dog, **d_o** – donkey, **f** – fowl, **g** – goat, **h** – horse, **r** – rodent and **s** – sheep
Hi – high bacterial load, **Lo** – low bacterial load, **Tp** – tap-water, **we** – well-water, **Y** – yes, **N** – no,
Op – opaque-watery, **Sf** – soft and **TL** – thick-liquid, ■ Present, □ Absent

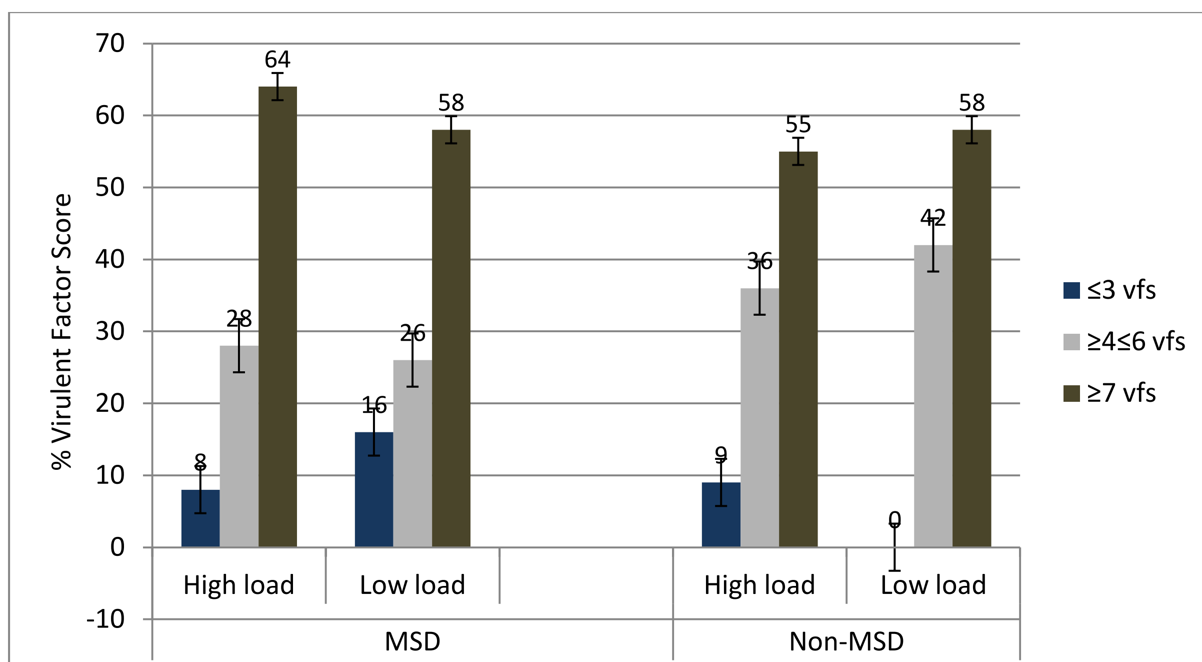
Table 5.5: Virulent factor score against Bacterial load and other variables among Asymptomatic children (n=45)

Sample Identity (SID)	Demography		Bacteria Load	Co-infection	Water Source	Stool Consistency	Animal	Type of Animal Present	pAA plasmid																Chromosome					Virulence Factor Score
															Adhesin						Toxins									
	Age in Month	sex							aatA	aggR	aap	Orf3	capu	Orf61	aafC	Agg3/4C	Agg3A	aafA	aggA	Agg4A	astA	sat	sepA	pet	pic	sigA	aaiC	air	eilA	
100099	3	M	Lo	Y	Tp	TL	g, s, r, f	**4**					■		■	■										■			■	5
100123	10	F	Hi	Y	Tp	Sf	g, s, c, f	**4**	■	■	■	■	■	■		■					■		■						■	10
102061	8	M	Hi	Y	we	Op	g, s, d, r, f, d_o, h	**7**		■		■	■	■			■					■			■					7
102157	8	M	Hi	Y	Tp	Op	r, f	**2**	■												■									2
103018	10	M	Hi	N	Tp	Sf	g, r, f, d_o, h	**5**		■		■	■	■		■					■		■				■	■		9
103177	18	M	Hi	Y	Tp	Sf	g, s, f, d_o	**4**		■		■				■	■								■	■	■			7
103192	18	M	Hi	N	we	Fo	g, r, f, d_o	**4**					■	■								■						■		4
103203	10	M	Hi	N	we	TL	g	**1**					■	■		■				■			■				■			6
103208	20	F	Lo	N	Tp	TL	g, s, c_a, r, f	**5**	■	■	■	■										■							■	6
103237	6	F	Hi	N	we	Fo	r, f	**2**		■		■	■	■		■							■					■		7
103285	6	M	Hi	Y	we	TL	g, s, r, f, d_o	**5**	■	■	■	■	■	■		■				■			■							9
103394	8	M	Hi	Y	we	Sf	g, d, c_a, c, r, f, d_o, h	**8**	■	■	■	■	■			■				■			■							8
100127	5	F	Lo	N	Tp	TL	g, s, r, f	**4**	■	■	■	■	■	■		■												■	■	9
100525	23	F	Hi	N	Tp	Op	g, r	**2**		■			■								■		■				■		■	6
102484	6	F	Hi	N	we	TL	g, d, r, f, d_o	**5**	■														■							2
103042	13	M	Lo	N	Tp	TL	g, s, d, c_a, r, d_o, h	**7**		■		■	■	■		■					■				■	■		■		9
103273	18	M	Lo	N	Tp	Sf	g, s, r, f, d_o, h	**6**		■		■		■					■					■	■		■			7
103460	19	M	Hi	Y	we	TL	g, s, d, c_a, c, r, f, d_o	**8**		■	■	■	■	■		■					■		■			■	■	■		11
103700	8	M	Hi	N	Tp	TL	g, c, r, f, d_o	**5**	■	■	■	■	■	■		■			■		■				■				■	11
100138	5	M	Hi	N	we	Op	g, s, d, f	**4**		■		■		■		■				■	■							■		7
100235	17	F	Lo	N	Tp	Sf	g, s, ca, f	**4**	■	■	■	■	■	■					■				■		■	■			■	11
100255	8	M	Hi	N	we	TL	g, s, d, c, r, f, d_o	**7**		■		■		■							■				■			■		6
100495	9	M	Hi	N	Tp	Sf	g, s, d f	**4**	■	■			■	■							■								■	6
102287	6	F	Hi	N	Tp	Sf	g, s, r, f	**4**		■		■	■	■			■				■		■		■		■			9
102341	18	M	Hi	N	we	Fo	r, f	**2**			■	■		■							■								■	5
102614	22	M	Hi	N	we	Fo	g, s, r, f	**4**		■		■		■					■				■		■		■			7
100896	7	M	Hi	N	we	Fo	g, s, c, f	**4**	■	■		■		■					■		■		■					■		8
102210	7	M	Hi	N	Tp	Sf	g, s, d, c_a, f	**5**		■	■	■	■	■			■		■		■	■	■		■					11
102211	9	M	Hi	Y	Tp	Sf	g, s, d, c_a, f	**5**		■	■	■	■	■		■	■		■		■				■		■	■		12
102670	9	F	Hi	N	Tp	Sf	g, f	**2**	■	■																				2
102990	7	F	Hi	N	Tp	Op	g, s, d, r, f, d_o	**6**			■	■		■		■	■										■			6
103396	8	M	Lo	N	Tp	Fo	g, s, r, f	**4**	■	■	■	■		■	■							■							■	8
100018	10	M	Hi	N	Tp	TL	g, s, r, f	**4**	■	■	■	■	■	■		■			■			■						■		10
100318	5	M	Hi	N	Tp	TL	s, c, r, f	**4**	■	■		■		■							■								■	6
100801	10	M	Lo	Y	Tp	TL	d, c_a, f, d_o, h	**5**	■	■	■	■	■	■		■				■			■		■		■	■		12
100800	9	F	Hi	Y	Tp	Sf	d, c_a, r, f, d_o, h	**6**		■		■		■		■							■		■		■	■		8
102103	24	F	Hi	Y	we	Fo	g, r, f, d_o, h	**5**	■				■									■							■	4
102154	7	M	Hi	Y	Tp	Sf	g, r, f, d_o, h	**5**					■	■					■									■		4
103027	23	F	Lo	N	Tp	Fo	g, s, r, f	**4**		■		■		■					■		■						■			6
103598	15	F	Hi	N	Tp	Fo	g, s, r, f, d_o, h	**6**					■			■			■										■	4
100715	9	M	Hi	Y	Tp	Sf	g, f, d_o, h	**4**	■	■	■	■		■		■						■								7
102433	21	M	Lo	N	we	TL	g, s, f, d_o, h	**5**		■	■			■			■						■		■		■	■		8
102792	21	M	Lo	Y	Tp	Op	g, s, r, f, d_o, h	**6**	■	■	■	■																		4
102849	4	F	Hi	Y	Tp	TL	g, s, f, d_o, h	**5**				■		■				■									■	■		5
103451	27	F	Lo	Y	we	Sf	g, s, r, f, d_o	**5**					■								■	■	■				■		■	6

Key: c – cow, $\mathbf{c_a}$ – cat, **d** – dog, $\mathbf{d_o}$ – donkey, **f** – fowl, **g** – goat, **h** – horse, **r** – rodent and **s** – sheep
Hi – high bacterial load, **Lo** – low bacterial load, **Tp** – tap-water, **we** – well-water, **Y** – yes, **N** – no,
Fo – formed, **Op** – opaque-watery, **Sf** – soft and **TL** – thick-liquid, ■ Present, □ Absent

Our examination of the proportional association of Virulent Factor Score (VFS) and bacterial load among the MSD and non-MSD showed no association. In the MSD group virulent factor score was proportionately more in HBL compared with LBL for vfs ≥4≤6 (28% vs. 25%) and vfs ≥7 (64% vs. 58%), but in the non-MSD group virulent factor score ≥4≤6 and ≥7 were proportionately more in LBL compared with HBL which are (42% vs. 36%) and (58% vs. 55%) respectively (figure 5.2).

Figure 5.2: proportional association of virulent factor score and bacterial load among MSD and non-MSD children



5.3 Discussion

EAEC is a worldwide recognised diarrhogenic strain (Kaur, Chakraborti et al. 2010). Several studies have reported the implication of EAEC in diarrhoea among children and adults in both industrialised (Smith, Cheasty et al. 1997; Presterl, Nadrchal et al. 1999; Knutton, Shaw et al. 2001; Cohen, Nataro et al. 2005) and low income countries (Okeke, Lamikanra et al. 2000; Okeke, Ojo et al. 2003; Kahali, Sarkar et al. 2004; Araujo, Tabarelli et al. 2007). However, in the GEMS that was conducted in four African and three Asian countries from 2007-2010, it became clear that the relationship between presence of EAEC and disease among MSD children from low income countries is not absolute (Kotloff, Nataro et al. 2013) so a more sensitive, specific and practical diagnostic methods were needed to investigate the association of EAEC with disease. The gold standard for the identification EAEC is the Hep-2 adherence test (Nataro, Kaper et al. 1987), but it is limited to reference laboratories, as it requires special facilities and skill. Despite many diagnostic methods developed, limited studies have had the opportunity to use a well-defined case-control data to quantify bacterial load to show that the relationship between the EAEC high bacterial load and diarrhoea is either causal or not. To date, there are no available data that validate a robust, easy and affordable method for the detection of infectious EAEC strain.

In this study, we investigated the relevance of bacteria load measurement using Taqman-QPCR as a reliable diagnostic tool to identify true pathogenic EAEC strains that cause diarrhoea. We choose *aatA* that has 90% qPCR detection and quantification for the analysis of the bacterial load data. Although it has been shown that plasmids may vary in gene content and have the potential to transfer to unrelated bacteria (Dobrindt 2005) nonetheless, previous studies have utilised *aatA* maker genes to detect and quantify EAEC presence (Nataro, Mai et al. 2006; Chattaway, Harris et al. 2013; Liu, Kabir et al. 2014). We used Cycle threshold (C_t) value as an indicator of bacterial load and defined a cut off that yielded 60% sensitivity and

27.5% specificity. These values clearly imply that estimation of bacterial load by C_t values of a qPCR for marker (*aaiC* and *aatA*) genes is a poor diagnostic test for EAEC infection that cause diarrhoea. Interestingly, a similar investigation conducted on routine faecal samples in the United Kingdom (Chattaway, Harris et al. 2013) corroborate our finding, the specificity in that study was also 60%. In our study, the results obtained for the crude model logistic regression analysis was in agreement with the diagnostic values as there was no association of EAEC high bacterial load with disease (OR 0.61 and p-value 0.143).

The multivariable analysis showed that presence of cows and rodents in a household, having a co-infection, underweight, and lower score mid-upper-arm-circumference (muac) were associated with higher odds of diarrhoea. One possible explanation to this outcome is that if an EAEC infected child presented these variable symptoms there is liklyhood that the child harbourghs high bacterial load EAEC that can possibly result into diarrhoeal disease.

No doubt, the complexity of interpreting pathogen isolated from faecal samples of children with diarrhoea and compared with asymptomatic colonisation is huge. That is why three (Environmental, bacterial and host) factors were often considered in many studies for suitable interpretation. There are numerous reports highlighting the role of environmental factors in diarrhoeal disease in which contaminated drinking water was implicated (Baker, O'Reilly et al. 2016), including a study conducted in The Gambia which showed filtering water through a cloth and of storing drinking water significantly associated with diarrhoeal episode in children (Baker 2011). Although well-water was not associated with EAEC high bacterial load in our multivariable analysis but a study has specifically incriminated EAEC in well water. For example, in 1996, a village outbreak of diarrhoea in India was epidemiologically associated with the drinking water from open well contaminated with EAEC (Pai, Kang et al. 1997). An important environmental factor observed was the association of cow and rodent with high bacterial load EAEC among the diseased individuals which signals that these

animals may serve as risk factors for EAEC diarrhoea. A possible explanation for the assumption is that over 90% of the households in the area/region where this study was conducted consumed untreated raw cow milk that may have been exposed to EAEC contamination and rodents such as rats and mice common in the households often suck from spill milk that occurs during processing and aliquotting for consumption. Augmenting this speculation, a study implicated cheese made from unpasteurised sheep milk as likely source of EAEC infection in Italy (Scavia, Staffolani et al. 2008). Although EAEC strains are considered to be adapted to the human host (EFSA 2015) hence, no strains of EAEC pathotype was detected in studies that examined the faecal samples obtained from sick and healthy calves and lambs in Kashmir India (Wani, Hussain et al. 2013). Similarly, EAEC was not detected in the investigation of healthy cattle, sheep and pigs in UK slaughter house (Cassar, Ottaway et al. 2004). However, a study from a low income country have implicated animals such as cattle, chicken and pigs as possible risk factors for EAEC infection (Kagambega, Martikainen et al. 2012). So in a poor sanitation setting it can be possible that animals become exposed to EAEC originating from human waste. Another observation from our study revealed that goat and sheep which are commonly found in many households in rural Gambia are not associated with EAEC high bacterial load infection and it is not clear whether one or both these animals provide herd-effect against EAEC infection a speculation that requires further investigation.

Studies from low income countries have shown cases of mixed infection particularly EAEC co-infecting with other diarrhoeal pathogens (Adachi, Jiang et al. 2001; Kotloff, Nataro et al. 2013). However, other than few cases that involved malnourishment, our data showed that EAEC did not, independently but in the presence of another pathogen, particularly rotavirus and shigella cause disease. Our finding is strengthened by a study that uses three different advanced molecular methods that include PCR-Luminex, multiplex real-time PCR and

TaqMan array card to test for 15 enteric pathogens including EAEC (Liu, Kabir et al. 2014). Despite EAEC co-infecting with other pathogens the three methods showed no association of EAEC with symptomatic diarrhoeal children.

Studies have given an assertion that a significant proportion of global malnutrition result from disorder of intestinal absorptive function resulting from repeated enteric infection (Black, Allen et al. 2008), and other studies have linked EAEC persistent diarrhoea to malnutrition and decreased physical and congnitive development in children (Guerrant, Oria et al. 2008). Likewise, studies have reported that malnutrition predisposes to EAEC and vice versa (Roche, Cabel et al. 2010), consistent with the trend seen in the results from our study which showed a high odds ratio of 4.09 and 4.94 for underweight and lower-score-MUAC respectively. Also, the results were further augmented with malnutrition and under-nutrition accounting for 21.25% (17/80) and 26.25% (21/80) respectively, although the two conditions are equally distributed between high and low bacterial load EAEC strains as indicated in the table 5.3. However, studies have showed that malnutrition predisposes host to diarrhoea caused by the pathotypes of *E. coli* (Nataro 2006) that include EAEC and vis versa.

In the proportional estimate of abnormal and normal stool consistency among EAEC sole infected MSD children we found out that abnormal thick-liquid account for 100% in the high load MSD children and zero percent in the low load MSD children. In the case of abnormal opaque-watery, high load MSD children account for 79%, low bacterial load MSD children account for 21% (figure 5.1). This finding throws more light on our understanding of a major symptom of EAEC characterised with opaque watery and thick-liquid or mucoid supporting previous studies (Nataro, Mai et al. 2006; Estrada-Garcia, Perez-Martinez et al. 2014).

We are compelled to give account of the demography and other characteristics of a fatal case with opaque-watery stool consistency. The patient was a 10 month old female child

harbouring EAEC of high bacterial load with no other pathogen other than EAEC that consist of 9 virulent factors score (vfs) which are *aatA, aap, orf3, orf61, aggA, astA, sat, pet* and *air*. The patient had neither malaria infection nor bacteria growth from blood culture result, her source of drinking is tap-water and her compound harbours five type of animals that include goat, sheep, cow, rodent and fowl (table 5.4). A critical observation among these characteristics is the unique combination of virulent factors involved and despite the strain fatality it has no aggR factor which makes it atypical EAEC. So, a close descriptive study of the gene combination among symptomatic and asymptomatic children (table 5.4 & 5.5) suggest that combination of *aatA*, *pet, aggA, orf3* and *orf61* are required in an EAEC strain to cause a disease that can possibly result to fatality in this geographic area. This assertion requires further investigation since the scope of this study does not cover gene expression and gene functional and interaction investigation. More importantly host factors have to be considered as malnutrition was the major recognised risk factor in this particular case. There was another fatal case with similar characteristics of stool consistency opaque-watery and high bacterial load of EAEC but co-infected with astrovirus and virulence factor score of three (*orf61*, *Agg3/4C* and *air*) genes, and the patient is 7 month old male without malnourishment and malaria, and no bacterial growth from blood culture (table 5.4). A total of three deaths have been observed in Gambia specific GEMS data that were attributed to EAEC co-infecting with astrovirus only. These incidences may not have happened by chance, so a thorough investigation is required to unveil potential mechanism of synergistic virulence been exhibited from the combination of the two organisms that often result to fatality.

The association of virulent factor score and high or low bacterial load among MSD and non-MSD children was proportionately investigated. There was no major/significant difference in VFS between the two groups but the irony was that there was high proportion of vfs ≥7 and vfs ≥4≤6 in the high bacterial load compared to low bacterial load respectively in the MSD

group, while in the non-MSD group there was high proportion of vfs ≥7 and vfs ≥4≤6 in the low bacterial load compared to high bacterial load, but in the vfs ≤3 it was a reverse (figure 4). This result showed that there is high likelihood of having vfs ≥4 in EAEC strain with high bacterial load and vfs ≤3 in EAEC strain with low bacterial load from diarrhoeal children compared to non-diarrhoeal children. Several studies have showed virulent factor score and disease in EAEC strains (Samie, Obi et al. 2007; Boisen, Scheutz et al. 2012; Lima, Boisen et al. 2013; Jensen 2017) but none has shown association with bacterial load. However, this observation is not absolute for diagnosis.

Our analysis of pathogenicity or virulence index of virulence factors in cases and controls of both high and low bacterial load showed *pet* as the only virulence gene that has high pathogenicity index of relevance and strong significant odd ratio and p-value. This outcome is not surprising because the *pet* gene was implicated in previous studies (Eslava, Navarro-Garcia et al. 1998; Lima, Boisen et al. 2013; Bafandeh, Haghi et al. 2015; Ikumapayi, Boisen et al. 2017). However it may be important to considere the *pet* gene as a diagnostic marker when investigating infectious EAEC in this region. It is not clear whether host genetic factor play a role in the EAEC infection in the presence of *pet* gene even though its presence in EAEC strains is small but its significance in the EAEC diseased persons particularly children from developing countries is huge. There can be a potential role of pet-toxin in relation to nutrients and other competing organisms. Studies have shown that some bacteria play a role in competing for the consumption of a limited resource by one strain restricting supply of nutrient to other competitors. The mechanisms used can be either through increased nutrient uptake or through the extracellular secretion of molecules that harvest nutrient (MacLean RC. and Gudeji I. 2006). An example of the former is *Escherichia coli* that can metabolically shift from fermentation to respiration when oxygen is present, generating high growth rates but low yield, allowing them to absorb nutrient faster than their competitors (Ghoul M and Mitri

S. 2016). The example of the latter competitive strategy is the production of digestive enzymes to degrade complex nutrient molecules, or sideropgores, which are iron-scavenging molecules that access insoluble iron. Pet is speculated to play a role in malnutrition using the latter strategy. For example it can be possible that enterotoxin and cytotoxin secreted by pet are use as a weapon to eliminate their competitors, and concurrently causing under nutrition among children. Pet toxins promote host inflammation that impedes commensal survival. The inflammation caused by toxin-mediated diarrhoea significantly decreases the number of commensal microbiota in the intestine, and in turn, increases the chance of colonization and proliferation of incoming pathogens because of less competition (Lupp C et al. 2007). This can be an area of future investigation of pet toxins. So, we suggest that *pet* gene be added to other diagnostic marker such as *aatA*, *aaiC* and *agg*R for detecting EAEC infection among children from developing countries.

We have examined the utility of qPCR to assess the health implication of EAEC from among the diarrhoeal children from The Gambia using observational case-control data. The results obtained showed possible inappropriateness of TaqMan qPCR technique to link EAEC to diarrhoeal disease. A limitation of this study was our inability to consider using *aggR* marker to run similar assay. Although, this was done during pilot study and the results obtained was similar to the result obtained in this study.

5.4 Conclusion

Our study showed importance of EAEC pathotype that causes diarrhoea but has not been adequately diagnosed resulting in the under estimating the medical importance of EAEC in the region. It is obvious that there is critical relationship between EAEC strains that has

unique combination of virulence factors and malnutrition that requires thorough investigation. Additionally, further investigation on the host genetic factor interacting with EAEC infection will probably explain role of EAEC in malnourished children. Our analysis examined the use of qPCR to diagnose EAEC causing MSD in children but the method revealed poor sensitivity and specificity. So, it is either that TaqMan-based qPCR is not a useful diagnostic tool for the EAEC that cause diarrhoea in children or that the appropriate marker gene for the bacterial load approach is yet to be identified. Although we speculate that a qPCR assay that targets *pet*, aap, *aatA* and *aggR* to obtain bacterial load result may be adequate to confirm infectious EAEC that cause diarrhoea in children who are under 5 years old from West-Africa region. Many diagnostic tools have been developed and some are in progress yet our current understanding about the diagnosis of EAEC is that non-labour intensive diagnostic tool that identifies true infectious EAEC strain remains a challenge mainly due to heterogeneity of EAEC strains. Therefore, we recommend further investigation for the development of diagnostic tools that distinguish pathogenic EAEC strain from non-pathogenic EAEC strain.

Chapter 6: Biofilm production in EAEC strains from diarrhoeal and non-diarrhoeal children

6.1 Introduction

In *E. coli*, the formation of a Biofilm requires three major components that include synthesis of curli, which allow cells to bind to various kinds of surfaces and to each other (Prigent-Combaret, Prensier et al. 2000), colanic acid, a viscous capsular exopolysaccharide that allows the formation of voluminous biofilms (Danese, Pratt et al. 2000), and type 1 pili which are needed for the initial attachment of bacteria to inert surfaces and to other cells (Prigent-Combaret, Prensier et al. 2000). Pili are proteinaceous appendages on the surfaces of bacteria, these structures accomplish adhension by overcoming electrostatic repulsive forces between substratum surfaces and the bacterial envelopes (Pratt and Kolter 1998).

Bacterial biofilms are distinct structures that have mushroom-shaped micro-colonies encased in a highly hydrated matrix of exopolymeric substances, polisacharrides and protein that are produced by the resident microorganisms, and with flat layers firmly adhered to the surface or specific intracellular microcolonies (Busscher, Bos et al. 1995; Garnett and Matthews 2012) (figure 6.1). The structure is common to many pathogenic bacteria and of huge importance for medicine and infectious disease (Goldberg 2002). The transcriptional *lacZ* reporter-gene fussions were used to described first set of specific genes that are involved in up or down-regulated in biofilm bacteria (Davies, Chakrabarty et al. 1993), the phenomenon that led to the understanding that bacteria attachment initiates the expression of a set of genes that culminates in a biofilm phenotype (figure 6.1) (Sauer 2003; Garnett and Matthews 2012). During biofilm formation many species of bacteria are able to communicate with one another through specific mechanism called quorum sensing, a system of stimulus to co-ordinate different gene expression (Garnett and Matthews 2012). Compared with their planktonic

(non-adherent) counterparts, biofilm associated cells are characterised by diverse functions that include enhanced resistance to conventional biocides or detergents, antimicrobial treatment, and host immune defense responses. In addition, a biofilm provides bacterial cells with high-osmolarity conditions, oxygen limitations and high cell density (Adamus-Bialek, Kubiak et al. 2015). Biofilms have been implicated in the colonisation of different medical devices and to be associated to human diseases, such as cystic fibrosis, burn wound infection and chronic otitis media with effusion, valve endocarditis (Sauer 2003) and persistent diarrhoea caused by Enteroaggregative *E. coli* (Tokuda, Nishi et al. 2010). It is extremely difficult to eradicate biofilms from living host due to its ability to tolerate antimicrobial agents within concentration range of 10-1000 times require to kill genetically planktonic bacteria and highly resistant to phagocytosis (Lewis 2001). Advances in the genetic and molecular basis of bacterial community behaviour can be employed to develop therapeutic targets-biofilm as a means to control infection due to formation of biofilm.

Enteroaggregative *E. coli* (EAEC) has been shown to cause acute and persistent diarrhoeal particularly among children from developing countries (Okeke 2009). The strain has been implicated in diarrhoeal outbreaks that have resulted in fatal cases in Europe and in Japan (Wakimoto, Nishi et al. 2004; Bielaszewska, Mellmann et al. 2011). Diarrhoea caused by EAEC is usually watery, and it can be accompanied by mucus and or blood. Colonisation of EAEC can occur in the mucosa of both the small and large intestines, which can resulted to mild inflammation in the colon (Nataro, Steiner et al. 1998). Biofilms formed by EAEC are distinct from biofilms formed by non-pathogenic *E. coli* in that they can form biofilms in the absence of common factors that include flagella, curli and antigen 43 (Ag43) (Sheikh, Hicks et al. 2001). EAEC biofilms are encased in a thick mucus layer on the surface of enterocytes (Croxen and Finlay 2010). Likewise, EAEC has the capacity to free up itself by penetrating

the mucus layer through mucolytic activity of the protein involve in intestinal colonisation (pic) a class 1 SPATE family (Gutierrez-Jimenez, Arciniega et al. 2008). A few genes that are either plasmid-borne or chromosomal that encode proteins involving in the formation of biofilms have been identified and this includes genes that encode a type VI secretion system of which detail is unclear (Aschtgen, Bernard et al. 2008).

Many diagnostic tools have been considered to detect EAEC that cause diarrhoea, but the gold standard remains characteristic-phenotypic aggregative adhension that involves the formation of a stacked-brick pattern of HEp-2 cells mediated by the genes found on a family of virulence plasmids called pAA plasmids. We aim to evaluate qualitative and quantitative screening of biofilm to identify EAEC causing diarrhoea among children less than five years old in rural Gambia. Few studies have employed phenotypic quantitative screening of biofilm to identify EAEC during outbreak of diarrhoea and for epidemiological studies (Wakimoto, Nishi et al. 2004; Boisen, Struve et al. 2008). Result from such study showed that quantitative of biofilm to identify EAEC causing diarrhoea can be reliable. The technique is very useful for direct detection of polysaccharide production as spectrophotometric measurements provide quantitative information on the ability of bacterial strains to rapidly grow while adhering to the substratum.

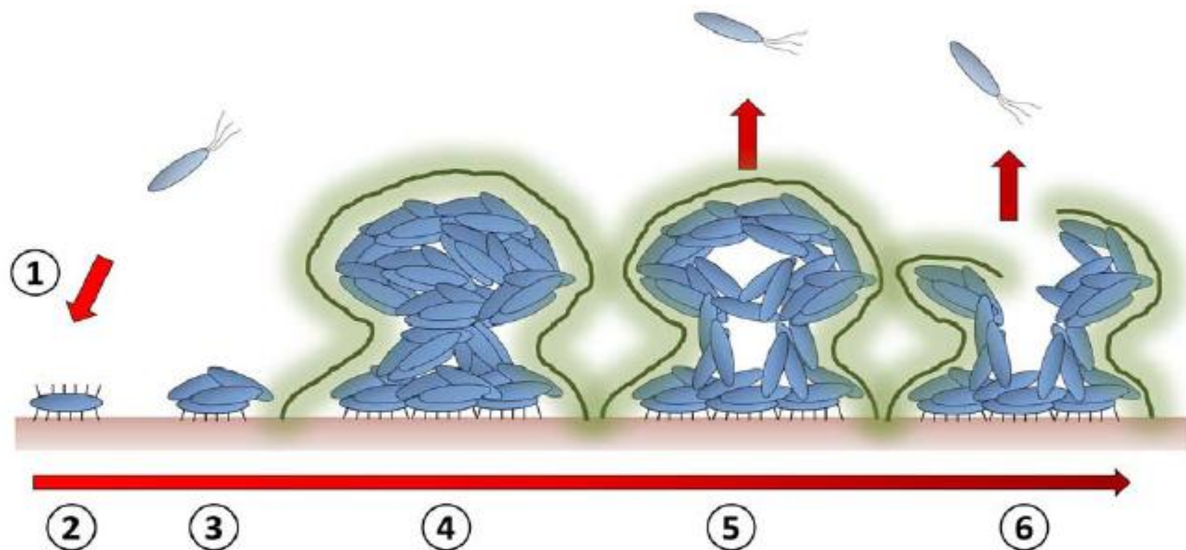


Figure 6.1: Schematic representation of Biofilm life-cycle: (1) free swimming bacteria (2) the bacterial cells adhered reversibly to the surface, at this stage (3) bacteria attachment became irreversible, this step is mediated mainly by exopolymeric substance, and the cells lose their flagella-driven motility, cells being to divide and the expression of further macromolecules allows them to stick together in small micro colonies. (4) These colonies grow and secrete a complex mixture of carbohydrates, proteins and lipids that encapsulate the bacteria. The biofilm matrix (fuzzy outline) provides protection and stability for the maturing biofilm. (5) When the biofilm reaches maturity, a number of factors will have developed a heterogeneous arrangement of cells and molecules within the biofilm, and given rise to solvent filled cavities and channels. This can result to dispersal of cells from the cellular mass. (6) Upon signal from the environment (waste build up or demand for nutrients, for example), molecules are released that cause lysis and matrix dissemination. Many planktonic cells are now released to find a new habitat**. (Adapted from James A. Garnett and Steve Mathews 2012 – Interaction in Bacterial Biofilm Development: A structural Perspective. Current Protein and Peptide Science, 2012, 13, 739-755)**

6.2 Results

Biofilm formation was established using three phenotypic methods of which two were qualitative and one quantitative. The two qualitative methods are test-tube and Congo red agar (CRA) test (Freeman, Falkiner et al. 1989) whilst the quantitative is tissue culture plate (TCP) method (Christensen, Simpson et al. 1985) and all of these methods were described in detail in the materials and method section of this thesis (3.16.2, 3.16.3 and 3.16.4). The positive (*E. coli 042*) and control (*E. coli HB101*) reference bacterial strains were correctly identified by the three methods for the presence and absence of biofilm.

Of the 400 EAEC isolates tested for the identification of biofilm, test-tube method detection rate was 6.5% (26/400), CRA account for 47% (188/400) and TCP was 54.75% (219/400) (table 6.1 & 6.2). In this study we adopted the data generated by the TCP method for all analysis relating to biofilm due to its reliability and widely approved screening technique although study showed that it has low specificity (Stepanovic, Vukovic et al. 2000). The biofilm mean (optical density) OD_{570nm} value for the three positive controls (3.485, 3.697 and 3.296) EAEC 042 was 3.4926 ±0.2006, whereas the mean OD_{570nm} value for the three negative controls (0.5149, 0.5238 and 0.5324) EAEC HB101 was 0.5237±0.0.0087. A cut-off OD_{570} was obtained by taking the average of all the ODs of the negative control EAEC HB101 and thrice the value of standard deviation (SD) of the negative control was added to it. Therefore the cut-off used was the mean OD_{570nm} (0.5237) of the negative control + (3 x 0.0087) which, is 0.54. Therefore, EAEC strains are classified biofilm producer if the OD_{570} reading was ≥0.54 and EAEC strains are classified non-biofilm producers if the OD_{570} readings was ≤0.53. For all the EAEC strains investigated, the OD_{570nm} readings ranged from 0.2691 to 3.4956.

Therefore, result from data generated by the TCP method showed that , biofilm producing EAEC are found more in controls 61% (134/219) compared to cases 39% (85/219) and similar proportionate of result 64% (116/181) and 36% (65/181) showed in non-biofilm producing EAEC for controls and cases respectively (figure 6.2 A&B) yielded p-value 0.550. Similarly, result obtained from sole-EAEC isolates (without co-infection) showed more biofilm producing EAEC in controls 69% (66/96) conmpared to cases (31%) (30/96) with similar proportional distribution among cases 24% (19/80) and controls 76% (61/80) in sole-EAEC infection with non-biofilm producing EAEC (figure 6.2 C&D) yielded p-value 0.269.

We studied the distribution of 21 virulence genes in both the biofilm producing and non-biofilm producing EAEC isolates among cases and controls and found no evidence of an association with severity of diarrhoeal disease (table 6.3). However, further investigation revealed weak evidence (p=0.047) of an association between a virulence gene (*aatA*) and biofilm production (Table 6.4). Additionally, we investigated possible association of EAEC biofilm producer with virulence genes that have aggR gene in the background, interestingly, the result showed *aat*A, *aa*P, *ORF*63 and *ORF*61 genes to be highly significant among diarrhoeal children (table 6.5).

Also, we performed similar analysis on association of antimicrobial resistance with biofilm formed (BF +) and non-biofilm formed (BF-) EAEC strains from diarrhoeal and non-diarrhoeal children (table 6.6) but there was no evidence of an association. Again, comparison analysis of antimicrobial resistance and non-resistance among cases and controls between biofilm producing EAEC and non-biofilm producing EAEC were performed. The result showed similar distribution pattern of resistance among the resistant and non-resistant EAEC in the two groups (figure 6.3A, 6.3B, 6.4A & 6.4B).

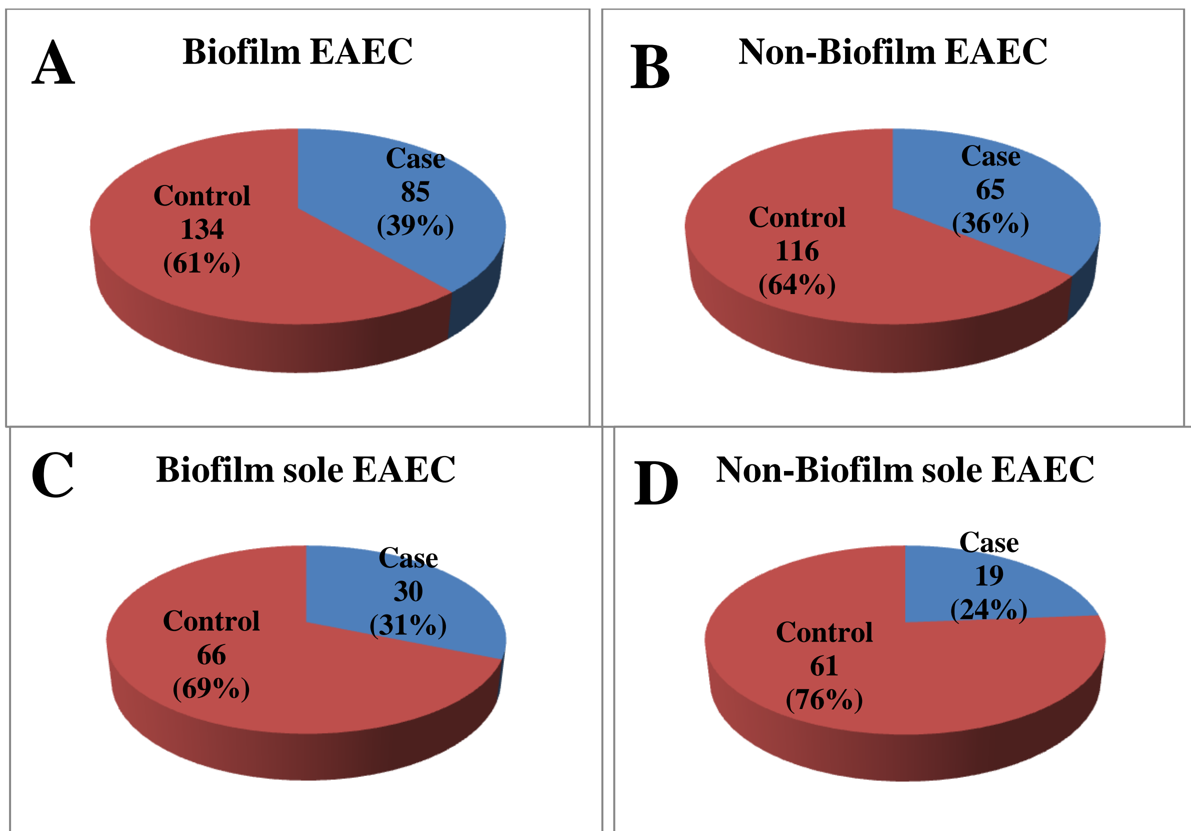
Table 6.1: Evaluating the use of CRA for biofilm screening tests using TCP as gold standard

Congo-Red-Agar	**Tissue culture plate**			
	Positive	Negative	Total	%
Positive	106	113	219	54.75
Negative	82	99	181	45.25
Total	188	212	400	100
Sensitivity = 106/188 x 100 = 556.4% Specificity = 99/212 x 100 = 46.7% Positive Predictive Value (PPV) is 106/219 x 100 =48%, Negative Predictive Value (NPV) is 99/181 x 100 = 55%				

Table 6.2: Evaluating the use of TT for biofilm screening tests using TCP as gold standard

Test-Tube	**Tissue culture plate**			
	Positive	Negative	Total	%
Positive	13	206	219	54.75
Negative	13	168	181	45.25
Total	26	374	400	100
Sensitivity = 13/26 x 100 = 50% Specificity = 168/374 x 100 = 45% Positive Predictive Value (PPV) is 13/219 x 100 = 6% Negative Predictive Value (NPV) is 168/181 x 100 = 92%				

Figure 6.2 (A, B, C and D): Proportional distribution of Biofilm in EAEC strains from cases and controls



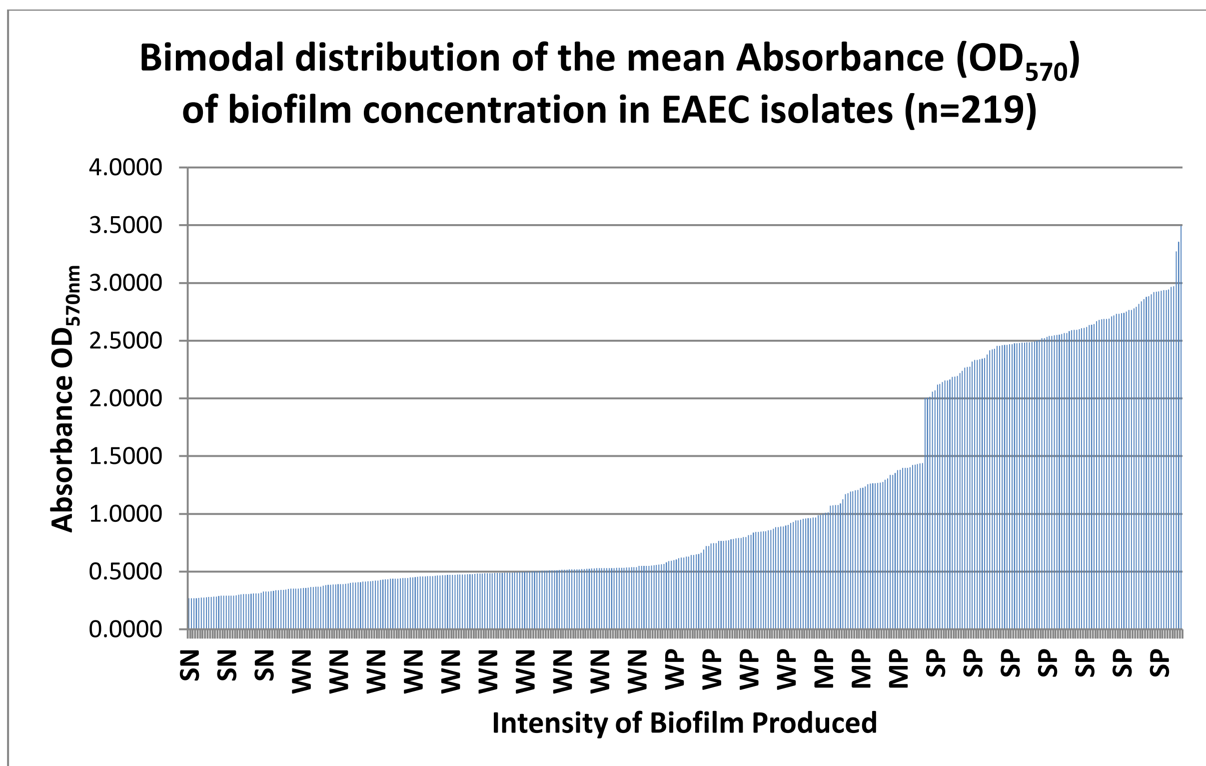
The statistical strength for both biofilm producing EAEC and non-biofilm producing EAEC in both cases and control (A and B) yielded p-value =0.550 which indicates that biofilm production were more among controls compared to cases.

Likewise in both biofilm producing sole EAEC infection and non-biofilm producing sole EAEC infection, biofilm iwas produced far more in controls compared to cases (C and D) yielded p-value 0.269.

Table 6.3: Categigorising biofilm producing EAEC base on the intensity of biofilm produced (n=219)

Intensity of biofilm produced	Case (n=85) n (%)	Control (n=134) n (%)	P – value	
Strong biofilm produced	39 (46)	65 (49)	0.704	
Moderate biofilm produced	10 (12)	31 (23)	0.035	
Weak biofilm produced	36 (42)	38 (28)	0.032	

Figure 6.2a: Bimodal distribution of the mean Absorbance



Key: SN – Strong Negative; WN – Weak Negative; WP – Weak Positive; MP – Moderate Positive and SP – Strong Positive

Table 6.4: Distribution of virulence genes in Biofilm producing and Non-Biofilm Producing EAEC strains among Cases and Controls

Virulence Gene	Biofilm Producing EAEC (n=219)			Non-Biofilm Producing EAEC (n=181)		
	Case (n=85) No. (%)	Control (n=134) No. (%)	P-value	Case (n=65) No. (%)	Control (n=116) No. (%)	P-value
aatA	34 (40)	47 (35)	0.461	16 (25)	33 (28)	0.577
aggR	54 (63.5)	89 (66)	0.661	41 (63	69 (59)	0.634
aaP	44 (52)	69 (51)	0.968	32 (49)	48 (41)	0.307
orf3	57 (67)	91 (68)	0.895	47 (72)	74 (64)	0.243
capU	56 (66)	83 (62)	0.554	50 (77)	69 (59)	0.017
aar	59 (69)	98 (73)	0.551	45 (69)	85 (73)	0.561
aafC	6 (7)	9 (7)	0.922	1 (1.5)	6 (5)	0.223
agg3/4C	29 (34)	44 (33)	0.844	26 (40)	48 (41)	0.856
agg3A	3 (4)	16 (12)	0.031	6 (9)	11 (9)	0.955
aafA	2 (2)	11 (8)	0.073	1 (1.5)	4 (3.5)	0.451
aggA	26 (31)	37 (28)	0.635	20 (31)	21 (18)	0.050
agg4A	6 (7)	6 (4)	0.413	9 (14)	10 (9)	0.271
astA	50 (59)	65 (49)	0.136	36 (55)	54 (47)	0.254
sat	17 (20)	28 (21)	0.873	10 (15)	27 (23)	0.206
sepA	27 (32)	37 (28)	0.510	21 (32)	24 (21)	0.082
pet	16 (19)	14 (10)	0.078	8 (12)	9 (8)	0.314
pic	27 (32)	40 (30)	0.764	26 (40)	47 (41)	0.945
sigA	8 (9)	16 (12)	0.559	9 (14)	13 (11)	0.602
aaiC	23 (27)	51 (38)	0.093	21 (32)	44 (38)	0.449
air	27 (32)	31 (23)	0.158	13 (20)	26 (22)	0.704
eilA	42 (49)	65 (49)	0.896	31 (48)	52 (45)	0.710

Table 6.5: Association of virulence factor and the formation of biofilm (BF +) and the non-formation of biofilm (BF -) among EAEC isolates from diarrhoeal and non-diarrhoeal children (n=400)

Virulence Gene	**Diarrhoeal Children (n = 150)**			**Non-Diarrhoeal Children (n= 250)**		
	BF + (n=85) No. (%)	**BF - (n=65) No. (%)**	**P-Value**	**BF + (n=134) No. (%)**	**BF - (n=116) No. (%)**	**P-Value**
aatA	34 (40)	16 (24.6)	0.047	47 (35)	33 (28)	0.262
aggR	54 (63.5)	41 (63)	0.954	89 (66)	69 (59)	0.256
aaP	44 (51.8)	32 (49)	0.758	69 (51)	48 (41)	0.110
orf3	57 (67)	47 (72)	0.489	91 (68)	74 (64)	0.493
capU	56 (65.9)	50 (76.9)	0.141	83 (62)	69 (59)	0.691
aar	59 (69)	45 (69)	0.980	98 (73)	85 (73)	0.979
aafC	6 (7)	1 (2)	0.112	9 (6.7)	6 (5)	0.608
agg3/4C	29 (34)	26 (40)	0.458	44 (32.8)	48 (41)	0.162
agg3A	3 (3.5)	6 (9)	0.145	16 (12)	11 (9.5)	0.532
aafA	2 (2)	1 (1.5)	0.724	11 (8)	4 (3.5)	0.113
aggA	26 (31)	20 (31)	0.980	37 (27.6)	21 (18)	0.075
agg4A	6 (7)	9 (13.9)	0.169	6 (4.5)	10 (8.6)	0.181
astA	50 (59)	36 (55)	0.673	65 (48.5)	54 (46.6)	0.757
sat	17 (20)	10 (15)	0.465	28 (21)	27 (23)	0.650
sepA	27 (32)	21 (32)	0.943	37 (27.6)	24 (20.7)	0.203
pet	16 (19)	8 (12)	0.280	14 (10.5)	9 (7.8)	0.463
pic	27 (31.8)	26 (40)	0.295	40 (30)	47 (40.5)	0,077
sigA	8 (9)	9 (13.9)	0.395	16 (12)	13 (11)	0.856
aaiC	23 (27)	21 (32)	0.484	51 (38)	44 (38)	0.983
air	27 (32)	13 (20)	0.106	31 (23)	26 (22)	0.892
eilA	42 (49)	31 (48)	0.834	65 (48.5)	52 (44.8)	0.560

After adjusting for multiple testing using Bonferroni method (0.05/21) that resulted in cut of P-value 0.002, only aafC gene was found to be associated with diarrhoea in the presence of biofilm production. BF+ (Biofilm present) and BF- (Biofilm absent).

Table 6.6: Association of EAEC biofilm producer and diarrhoea with virulence genes having AggR in the background

Virulence genes	Diarrhoeal Children (n = 85)			Non-Diarrhoeal Children (n= 134)		
	BF + / aggR+ (n=54) No. (%)	**BF + / aggR- (n=31) No. (%)**	**P-Value**	**BF + / aggR+ (n=89) No. (%)**	**BF + / aggR- (n=45) No. (%)**	**P-Value**
aatA	31 (57)	3 (10)	<0.001	43 (48)	17 (38)	0.246
aap	39 (72)	8 (26)	<0.001	60 (67)	21 (47)	0.020
orf3	48 (89)	14 (45)	<0.001	74 (83)	31 (69)	0.058
capU	38 (70)	18 (58)	0.249	58 (65)	25 (56)	0.279
orf61	41 (76)	13 (42)	0.001	68 (76)	28 (62)	0.085
aafC	2 (4)	4 (13)	0.110	6 (7)	3 (7)	0.986
agg3/4C	20 (37)	9 (29)	0.453	31 (35)	13 (29)	0.489
agg3A	3 (6)	0 (0)	0.181	13 (15)	3 (7)	0.180
aafA	1 (2)	1 (3)	0.687	10 (11)	1 (2)	0.072
aggA	21 (39)	5 (16)	0.028	29 (33)	8 (18)	0.070
agg4A	4 (7)	2 (6)	0.868	5 (6)	1 (2)	0.369
astA	30 (56)	10 (32)	0.038	47 (53)	18 (40)	0.161
sat	13 (24)	4 (13)	0.215	15 (17)	13 (30)	0.105
sepA	24 (44)	6 (19)	0.019	32 (36)	5 (11)	0.002
pet	11 (20)	5 (16)	0.630	11 (12)	3 (7)	0.308
pic	18 (33)	9 (29)	0.681	30 (34)	10 (22)	0.170
sigA	6 (11)	2 (6)	0.478	12 (13)	4 (9)	0.438
aaiC	17 (31)	6 (19)	0.225	38 (43)	13 (29)	0.120
air	17 (31)	10 (32)	0.940	22 (25)	9 (20)	0.540
eilA	22 (41)	20 (65)	0.034	37 (42)	28 (62)	0.023

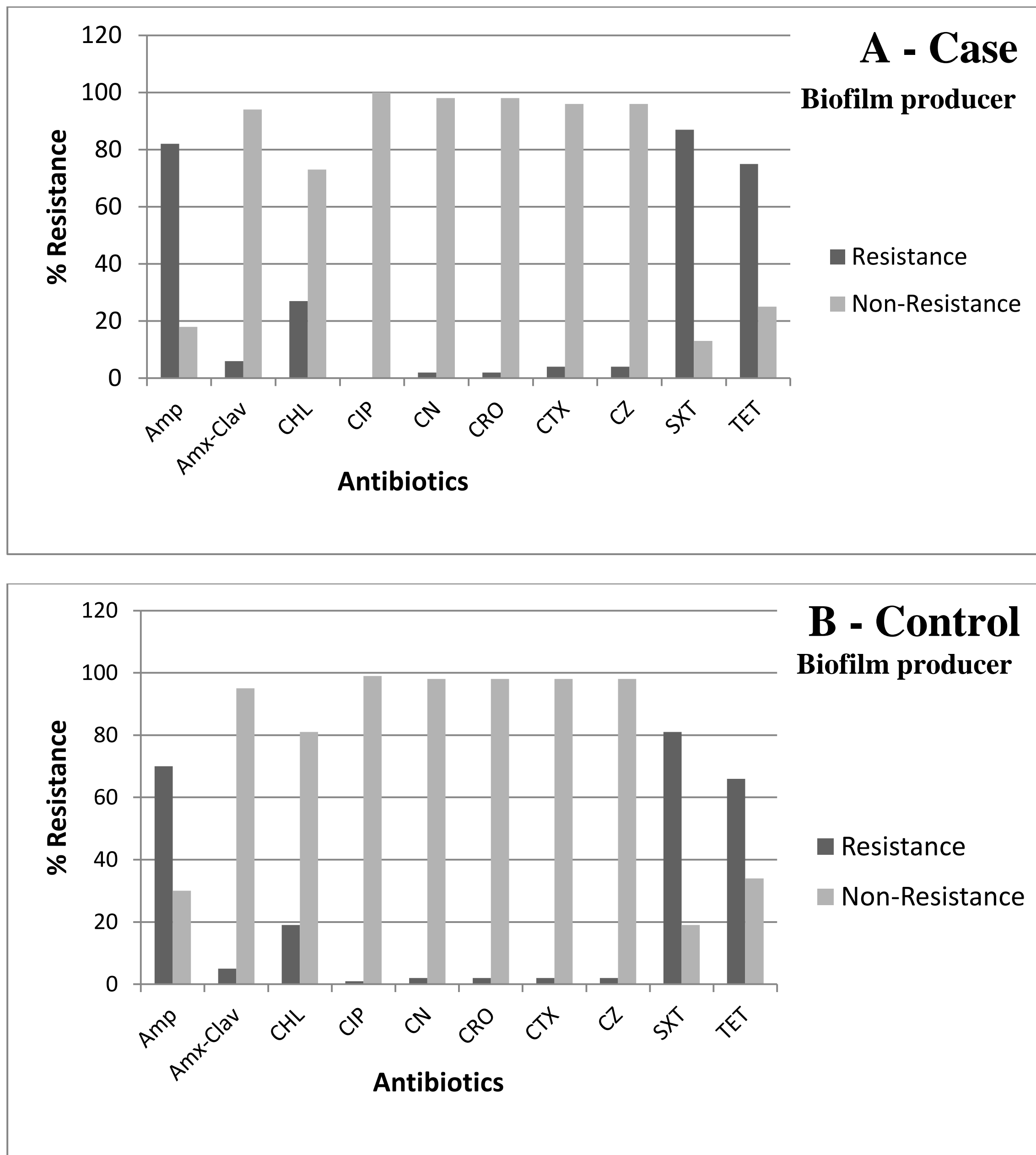
Note: In order to adjust for multiple testing Bonferronni method was applied to obtain a cut off p-value (0.05/20) 0.0025. Thus, four virulent genes that include *aatA*, *aap*, *orf3* and *orf61* were found to be statistically associated with diarrhoea in the presence of biofilm production and aggR in the background. aggR+ (aggR present) and aggR- (aggR absent); BF+ (Biofilm present) and BF- (Biofilm absent).

Table 6.7: Association of Antimicrobial resistance and the biofilm formation (BF +) and the non-biofilm formation (BF-) from diarrhoeal and non-diarrhoeal children (n=400)

Antibiotics	Diarrhoeal Children (n=150)		Non-diarrhoeal Children (n=250)	
	Resistance BF+ (n=85) No. (%)	Resistance BF- (n=65) No. (%)	Resistance BF+ (n=134) No. (%)	Resistance BF- (n=116) No. (%)
Ampicillin	70 (82)	54 (83)	94 (70)	87 (75)
Amoxicillin/Clav	5 (6)	4 (6)	6 (4.5)	6 (5)
Chloramphenicol	23 (27)	15 (23)	25 (19)	30 (26)
Ciprofloxacin	0 (0)	0 (0)	1 (1)	3 (2.6)
Gentamicin	2 (2)	1 (1.5)	2 (1.5)	1 (0.8)
Ceftriaxone	2 (2)	0 (0)	2 (1.5)	2 (1.7)
Cefotaxime	3 (4)	2 (3)	3 (2)	1 (0.8)
Ceftazidime	3 (4)	0 (0)	3 (2)	3 (2.6)
Co-Trimoxazole	74 (87)	55 (85)	109 (81)	91 (78.4)
Tetracyline	64 (75)	49 (75)	89 (66)	89 (76.7)

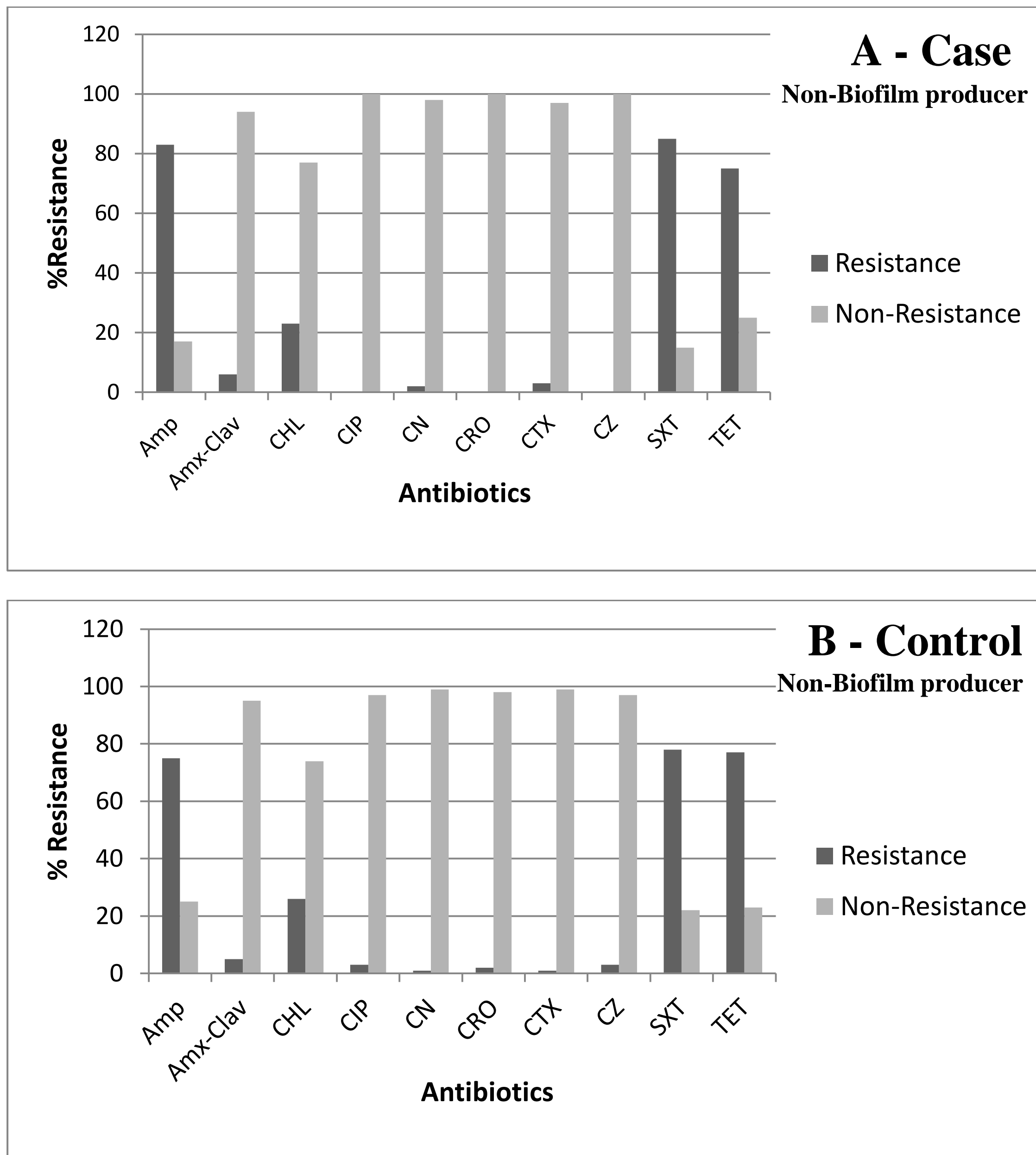
The above table 6.6 showed equal distribution of antibiotic resistance among biofilm producing and non-biofilm producing diarrhoeal children. There was no distinct evidence of biofilm formation limiting effect of at least six antibiotics such as amoxicillin-clavulanic acid, ciprofloxacin, gentamycin, ceftriaxone, cefotaxime and ceftazidime against the EAEC strains.

Figure 6.3 (A & B): Comparison of Antibiotic resistance pattern of biofilm producing- EAEC between cases and controls (n=219)



Key: Amp –Ampicillin, Amx-Clav – Amoxacillin-Clavulanic acid, CHL – Chloramphenicol, CIP – Ciprofloxacin, CN – Gentamicin, CRO – Ceftriaxone, CTX – Cefotaxime, CZ – Ceftazidime, SXT – Co-Trimoxazole and TET – Tetracycline

Figure 6.4 (A & B): Comparison of Antibiotic resistance pattern of **non-biofilm producing EAEC** between cases and controls (n=181)



Key: Amp –Ampicillin, Amx-Clav – Amoxacillin-Clavulanic acid, CHL – Chloramphenicol, CIP – Ciprofloxacin, CN – Gentamicin, CRO – Ceftriaxone, CTX – Cefotaxime, CZ – Ceftazidime, SXT – Co-Trimoxazole and TET – Tetracycline

6.3 Discussion

An important characteristic of EAEC strains is adherence to small and large intestinal mucosal surfaces in a thick aggregating biofilm (Nataro, Hicks et al. 1996). The ability of EAEC to form biofilms helps the strain to survive within the mucus layer, consequently colonise the host by evading host immune apparatus, presenting a barrier to the host antibacterial factors and hinders therapeutic agents to reach target (Torres, Zhou et al. 2005). EAEC producing biofilms often contribute to mucoid stool, malnutrition and persistent colonisation with prolong diarrhoea mostly among children from low income countries, (Kaur, Chakraborti et al. 2010). The goal of this investigation was to determine the prevalence of EAEC producing biofilm and association of various EAEC virulence factors with biofilm production among EAEC strains obtained from diarrhoeal and non diarrhoeal children from rural Gambia. Many studies have shown the importance of biofilm production by EAEC as a diagnostic tool to screen EAEC pathotypes causing diarrhoea illness (Wakimoto, Nishi et al. 2004; Mohamed, Huang et al. 2007). In this study, three detection methods that include test-tube (TT), Congo-red agar (CRA) and tissue culture plate (TCP) were used. The CRA and TT methods were evaluated using tissue culture plate method as gold standard considering its high sensitivity and specificity in previous studies (Wakimoto, Nishi et al. 2004; Bangar and Mamatha 2007; Hassan, Usman et al. 2011). Using the total number of EAEC that formed biofilm as denominator we obtained TT positive rate 12% (26/219) and CRA 86% (188/219). Because of the high rate (86%) for CRA, the result demonstrates usefulness of CRA in screening EAEC for the production of biofilm. However, when the evaluation was carried out using measures of diagnostic test of accuracy the results showed the CRA sensitivity 56.4% and specificity 46.7%, result for TT showed sensitivity 50% and specificity 45%. Thus, results for the CRA and TT suggest that the two methods may not be appropriate to screen for biofilm in EAEC isolates. One of the studies that

evaluates different detection methods of biofilm among clinical isolates recommended TCP over TT and CRA due to its reliability and sensitivity (Hassan, Usman et al. 2011). However, some previous studies have shown high sensitivity and specificity of CRA and TT to detect biofilm among enterobacteriacae that include *E. coli* isolates implicated in the catheter associated urinary tract infectios (Dadawala 2010; Niveditha, Pramodhini et al. 2012). Additionally, studies have demonstrated high detection rate of biofilm in Staphylococcus species using CRA and TT methods with high sensitivity and specificity rate (Oliveira and Cunha Mde 2010). However, studies looking at formation of biofilm in EAEC from cases and controls using phenotypic screening assay were scarce. Overall, results from the present study showed higher rate of EAEC forming biofilm accounting for 54.75% (219/400) compared to non-Biofilm EAEC 45.25% (181/400). Interestingly, also the result revealed that biofilm-EAEC in healthy control account for 61% (134/400) compared to cases 39% (85/400). These proportional differences were consistent in all the three methods used. In this study, we plot a graph to demonstrate biamudal distribution of absorbance readings of concentration of biofilm produced by EAEC isolates. The graph showed the order of abscence to presence of biofilm with steady increase in concentration (figure 6.2a).

Several studies have employed the use of multiplex PCR to detect many virulence genes in *E. coli* isolates to identify EAEC strains (Cerna, Nataro et al. 2003; Boisen, Scheutz et al. 2012; Ikumapayi, Boisen et al. 2017). The role of some of these virulence genes in the formation of biofilm in EAEC strains have been elucidated (Sheikh, Czeczulin et al. 2002; Boisen, Struve et al. 2008), these studies found that the fimbrial adhesion AAF/II that encodes *aafA* gene is required for formation of stable biofilms and in that same study *fis* and *yafK* were found to have the capacity to activate biofilm formation by regulating the transcription of the AAF/II biogenesis and activator aggR (Sheikh, Czeczulin et al. 2002). AggR is considered a global regulon for EAEC virulence genes that include *aatA* and *aap* that play important role in the

formation of biofilms (Boisen, Struve et al. 2008) and Dispersin, encoded by *aap* speculated to counteract aggregation mediated by AAFs to form biofilms (Sheikh, Hicks et al. 2001). Surprisingly however, the fact was that there was no statistical significant difference in the distribution of 21 virulence genes in Biofilm producing and Non-Biofilm Producing EAEC strains among Cases and Controls (table 6.3). This outcome suggest that biofilm formation among EAEC strains is a common phenomenon irrespective of whether the EAEC infected children from this region became symptomatic or remain healthy. Our this observation was contrary to previous study (Mohamed, Huang et al. 2007) and the reason for this outcome is unclear but could be due to singly studying the distribution of genes rather than study the combination of genes. One particular study has demonstrated the incidence of *aatA* in biofilm-EAEC strains and was found in all the EAEC strains that strongly produced biofilm of OD_{570nm} >0.2 in that study (Wakimoto et al. 2004) but this did not corroborate with the findings from our study. However, interesting results from this study was that some virulence genes that have *agg*R in the background together with the formation of biofilm showed strong association with diarrhoea (table 6.5). The virulence genes involved are *aatA* p-value <0.001, *aap* p-value <0.001, *orf3* p-value <0.001 and orf61 p-value 0.00172. This significant outcome suggests that biofilm-produced by typical-EAEC (EAEC with *agg*R) combined with one or more of its regulatory genes are likely to be more virulent compared to the biofilm-produced by atypical-EAEC (EAEC without *agg*R). These unusual findings were also reported in a similar study that showed significant association of *in vitro* production of biofilms from EAEC isolates with *agg*R and its regulated genes that include *astA*, *pet*, *aap*, *irp2*, and *set1*A in the background, although the EAEC isolates in that study were obtained from travellers (with or without diarrhoea) to EAEC endemic region (Mohamed, Huang et al. 2007). Thus, it is tempting to speculate that the presence of biofilm by typical EAEC indicates a functional *aggR* gene. The limitation of this study is our inability to utilise PCR method to detect genes

that are soley responsible for the expression of biofilm in EAEC strain to serve as gold standard method.

This study further examined the effect of biofilm production on the antimicrobial resistance pattern among diarrhoeal and non-diarrhoeal children but the result showed no significance difference among the two groups Table 6.6). Similar distributions were observed when we compared antibiotic resistance pattern in biofilm producing-EAEC and non-biofilm producing- EAEC between cases and controls (figure 6.3A, 6.3B, 6.4A and 6.4B). There was a huge proportional resistance of ≥75% observed against Ampicillin, Tetracycline and Sulphamethoxazole-trimethoprim which proven the presence of antimicrobial-resistant-EAEC strains circulating in the community. There are limited investigations that have specifically shown antibiotic resistance distribution among biofilm and non-biofilm producing EAEC strains to compare our finding against. However, studies have demonstrated biofilm and non-biofilm producing urinary *E. coli* been 100% resistant against Ampicillin and 90% against Sulphamethoxazole-trimethoprim (Hassan, Usman et al. 2011). Also, another biofilm producing *E. coli* study showed resistance against Sulphamethoxazole-trimethoprim 83% and Tetracycline 75% (Ponnusamy, Natarajan et al. 2012). These even distributions of antibiotic resistance among biofilm and non-biofilm producing *E. coli* corroborate our findings. However, Hassan *et al* study has shown ciprofloxacin to be 95% resistant in biofilm producing *E. coli* compared to 50% resistant in non-biofilm producing *E. coli*. Likewise Asian biofilm producing *E. coli* study showed 100% resistance against Chloramphenicol and Amoxyclav (Amoxicillin-clavulanic acid), 86% against Gentamicin and Cefotaxime, and 84% against Ceftazidim (Ponnusamy, Natarajan et al. 2012). These studies report demonstrated effect of biofilm production by uropathogenic *E. coli* against the antibiotics therapeutic action. This disproportional rate of resistance among biofilm producing and non-biofilm producing *E. coli* is not seen in our study. Although our current

study focuses on EAEC strains obtained from children less than 5 years old, besides, many studies from both developed and low-income countries have corroborated our findings by reporting that EAEC strains to be multidrug resistance mostly against Ampicillin, Sulphamethoxazole-trimetprime and Tetracycline (Kahali, Sarkar et al. 2004; Nguyen, Le Van et al. 2005; Aslani, Alikhani et al. 2011; Ali, Ahmed et al. 2014; Davoodabadi, Abbaszadeh et al. 2015; Hebbelstrup Jensen, Roser et al. 2016). Therefore, mechanisms of spread of EAEC-antibiotic resistant strain need to be well understood so that measures to deaccelerate the dissemination of the resistant strains are established and implemented.

6.4 Conclusion

The current study has demonstrated that TCP is probably the method of choice when screening for bacterial biofilm in epidemiologic samples. Also, the study showed that biofilm producing and non-biofilm producing EAEC is common among both diarrhoeal and non-diarrhoeal children, and that this study discovered that biofilm producing EAEC were found more among non-diarrhoeal children compared to diarrhoeal children. However, *in vitro* production of biofilm EAEC isolates that have *agg*R with one or more of its genes (*aatA*, *aap*, *orf*3 and *orf*61) that it regulates in the background are associated with moderate to severe diarrhoeal children. Additionally, the study highlight even distribution of antimicrobial resistance among biofilm and non-biofilm producing EAEC isolates from both cases and controls which suggest that biofilm production may not always be responsible for the spread EAEC resistance against the three antibiotics that include Ampicillin, Tetracycline and Sulphamethoxazole-trimethoprim.

Chapter 7: Antibiotic resistance patterns of EAEC strains from diarrhoeal and non-diarrhoeal children from rural Gambia

7.1 Introduction

Antibiotics, which literarily means 'against life,' are recognised as great warriors of modern medicines. These drugs have helped treated and cured infections that have threatened human life throughout history. The global beneficial effect of the advent of antibiotics is immense and well applauded. In developing countries where sanitation is critical, antibiotics have reduced morbidity and mortality caused by food-borne and other poverty related infections (CRSR 2005). Thus, the last century saw a stunning swing in human fortunes against bacteria. Humans have used antimicrobials often inappropriately, such as for promoting weight gain in farm animals leading to immense selective pressures. A bacterium that develops a point mutation which impacts resistance or the incorporation of a mobile genetic element which provides a resistance gene will have a clear edge over its susceptible peers in the presence of antibiotic. Indeed, this is survival of the fittest. Many studies have performed antibiotic susceptibility testing on bacterial strains belonging to enterobacteriacae family (Aslani, Alikhani et al. 2011) particularly on urinary gram-negative rods commonly referred to as coliforms that include the well known *E. coli* pathotype Uropathogenic *E. coli* (UPEC). However, our knowledge of the antibiotic resistant patterns of diarrhoegenic *E. coli* (DEC) is limited and variable.

Therefore, this chapter is set to study the antibiotic resistance patterns of EAEC strains from diarrhoeal and non-diarrhoeal children in order to elucidate our understanding on the magnitude of the cases of multidrug-resistant-EAEC strains associated to diarrhoea among children in the rural Gambia and the likely risk factors that contribute to the spread.

EAEC strains are recognised cause of diarrhoeal disease in developed as well as developing countries (Okeke 2009). Currently, EAEC has been associated with a wide range of diarrhoeal syndrome that include watery to invasive diarrhoea that can be acute or persistent (Nataro 2011). The strains can be recovered from symptomatic and asymptomatic individuals even among younger children in developing countries (Kotloff, Nataro et al. 2013). This likely explains the strain heterogeneity and role variation of host genetic factor, thus, some carrier become reservoir for the EAEC strain that can cause disease in more susceptible persons. At present, human is the only known reservoir for EAEC as other implicated risk factors such as animal have not been substantiated. The emergence and maintenance of diarrhoea caused by multidrug-resistant pathogenic bacteria have become significant public health problems that often results to increased morbidity, high mortality and huge health care costs due to treatment failures and longer hospital stays (Cho, Lim et al. 2011). Furthermore, global dissemination of plasmid-borne extended-spectrum Beta-lactamases (ESBLs) among enterobactraecae particularly *E. coli* is of huge concern (Imuta, Ooka et al. 2016). EAEC resistance to antimicrobial agents has been demonstrated in some part of the world (Hebbelstrup Jensen, Olsen et al. 2014) but little is known about pattern of EAEC resistance to antibiotic among children in sub-Sahara Africa. Heterogeneity, characteristics of the strain make the development of vaccine difficult. Hence, antibiotic likely remains useful therapeutic agent to treat persistent as well as acute diarrhoea caused by enteric bacterial pathogens that include EAEC.

The aim of this study is to investigate the prevalence of antibiotic resistance in EAEC strains from under five year old diarrhoeal and non-diarrhoeal children to investigate possible association of antibiotic-resistant-EAEC strains to diarrhoeal disease. Additionally, we investigated risk factors such as virulence genes and backyard animals that may contribute to the spread of antibiotic-resistant-EAEC among children in rural setting.

7.2 Results

In this study minimum inhibitory concentration (MIC) e-test method was performed on 20 (5%) of the 400 EAEC isolates using available e-test strips of eight different antibiotics which include Ampicillin, Sulfamethaxazole-Trimethoprim, Chloramphenicol, Ciprofloxacin, Ceftaxime, Ceftazidime, Gentimacin and Tetracycline. Also, disk diffusion method was performed on the 400 EAEC isolates that include the 20 EAEC isolates MIC e-test was performed. The antibiotic sensitivity results obtained from the two methods had 100% concordance (Table 7.1a).

The outcome of antimicrobial resistance patterns performed using ten antibiotics against each of 400 EAEC strains that constitute 150 (37.5%) diarrhoeal and 250 (62.5%) non-diarrhoeal children showed almost equal distribution among the two groups. Although high rate of resistance to Ampicillin, Sulphamethoxazole-Trimethoprim and Tetracycline were observed in both diarrhoeal and non-diarrhoeal groups, and the Ampicillin resistance showed moderate association to MSD (p-value 0.019), while Sulphamethoxazol-trimethoprim and Tetracycline had p-value 0.128 and p-value 0.368 respectively (Table 7.1).

The rate of multidrug-resistant EAEC to two or more antibiotics is very high in all study children account for 90% and 83% in cases and controls respectively (figure 7.1). The proportionate grading for categories of drug resistance (DR) EAEC strains ranging from 0DR to 5DR, revealed the category of 3DR EAEC strain that account for 48% MSD and 41% non-MSD children, whilst 0DR, 1DR and 5DR were less than 10% in both MSD and non-MSD (figure 7.2).

Investigation of resistant patterns of the ten antibiotics among cases and controls of two age strata 0-11 month and 12-59 month showed similar distribution patterns among the two

groups except for Ampicillin (p-value 0.037) and Amoxicillin-clavulanic acid (P-value 0.051) in cases and controls of age stratum 0-11 months (Table 7.2).

Two categories of MDR that involved ≥2MDR and ≤1DR were studied against three age strata 0-11, 12-23 and 24-59 month. The result showed strong evidence (p-value 0.0009) of association between ≥2MDR and age stratum 0-11month and there was steady decline of significance towards older age groups 12-23 and 24-59 months (table 7.3).

The analysis of two categories of MDR (≤1DR and ≥2MDR) against six age strata of study children further revealed the age stratum that harboured the highest rate of ≥2MDR which was age 7-11 months (36%) and highest rate of ≤1MDR was in 18-23 months (34%) (figure 7.3). Further analysis revealed comparison margin between ≥2MDR and ≤1DR within age stratum 7-11 month yielded p-value 0.010. The analysis of ≥2MDR that focuses on the six age strata of MSD children alone maintain age stratum 7-11 month harbouring the highest rate ≥2MDR at 39% followed by age stratum 3-6 months at 20% while the rest age strata 0-2, 12-17, 18-23 and 24-59 harbours 1%, 16%, 19% and 5% respectively (figure 7.4).

Further analysis of five categories of MDR EAEC strains that includes 0DR, 1MDR, 2MDR, 3MDR, and ≥4MDR against the six age strata of all study children was performed. The frequency value obtained clearly showed that 3MDR EAEC strains was the most common that account for 70 and more concentrated in the age stratum 7-11 months, followed by 2MDR, ≥4MDR, 1DR and 0MDR which account for 28, 26, 7 and 4 respectively (figure 7.5). Interestingly, similar pattern of distribution were obtained for analysis focusing on MSD children (figure 7.6) and non-MSD children (figure 7.7).

Of the twenty-one virulence genes investigated against MDR only two showed evidence of association with MDR and these are *agg3A* and *sepA* genes (table 7.4). However, further analysis on the association of virulent genes with MDR revealed that combination of 2, 3 and

or 4 virulence genes were significantly associated with multiple drug-resistant. Hence, combination of virulent genes pic+sepA (P+S) showed over 40% association with ≥4MDR, while combination of genes *aatA+Pet* (A+P), *aatA+aggR+Pet* (A+A+P), *aatA+Pet+aggA* (A+P+A) and *aatA+Pet+aggR+aggA* (A+P+A+A) were associated with 3MDR in the proportion 50%, 60%, 60% and 75% respectively (figure 7.8).

Investigation on the likelihood of animals as risk factor for the spread of MDR exonerate most animals that include goat, sheep, dog, cat, cow, rodent and fowl. Whilst donkey and horses are considered possible risk factors (Table 7.5)

There are six villages where the study participants were recruited. These villages were analysed against MDR in order to establish whether any of the villages contribute to the spread of multi-resistant EAEC strains by harbouring high rate of study participants with ≥2MDR. The result showed no evidence of association with spread of MDR-EAEC in all of the six villages in the study (table 7.6).

Table 7.1a: Comparing antibiotic sensitivity results from disk diffusion and MIC E-test methods

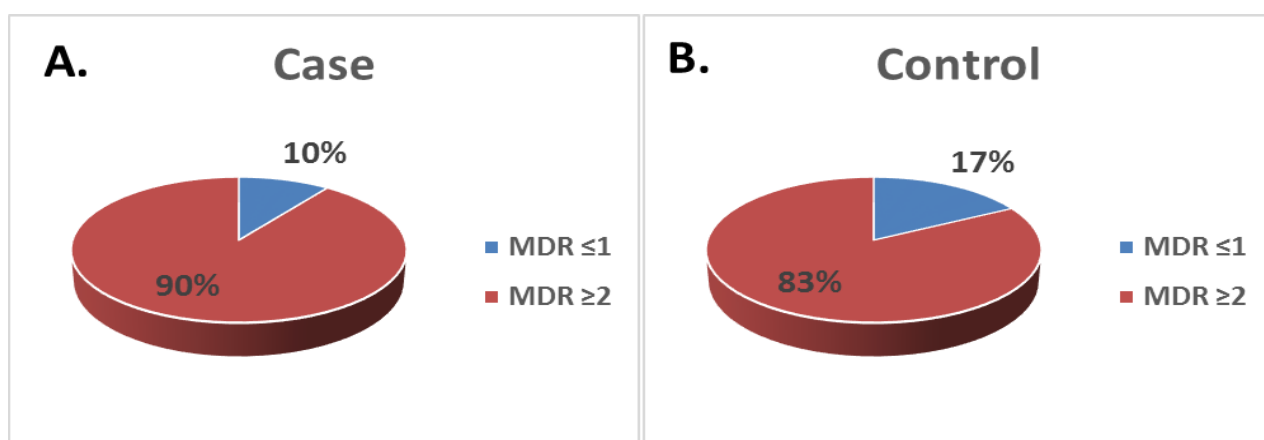
SID	Ampicillin		Ciprofloxac		Cefotaxime		Chloramph		Gentamicin		Trim-Sulfam		Tetracyclin		Ceftazidime	
	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion	MIC E-test	Disk diffusion
100041	R	R	S	S	S	S	S	S	S	S	R	R	R	R	S	S
100055	I	I	S	S	S	S	S	S	S	S	R	R	R	R	S	S
100096	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
100103	I	I	S	S	S	S	S	S	S	S	R	R	R	R	S	S
100107	R	R	S	S	S	S	R	R	S	S	R	R	R	R	S	S
100138	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
100161	R	R	S	S	S	S	S	S	R	R	R	R	R	R	S	S
100245	I	I	S	S	S	S	S	S	S	S	S	S	S	S	S	S
100404	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
100715	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
100716	R	R	S	S	I	I	S	S	S	S	S	S	R	R	S	S
100722	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
102274	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
102602	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
102866	R	R	S	S	S	S	S	S	S	S	R	R	R	R	S	S
102871	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
103069	S	S	S	S	S	S	S	S	S	S	R	R	S	S	S	S
103070	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
103240	R	R	S	S	S	S	S	S	S	S	R	R	R	R	S	S
103446	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
ATCC25922	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S
ATCC25923	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S	S

Table 7.1b: Antimicrobial resistance patterns of EAEC strains from diarrhoeal and non-diarrhoeal children

Antimicrobial agent	Concentration in µg	Diarrhoeal children (n = 150)	Non-Diarrhoeal children (n = 250)	Odd Ratio	P-Value
		Resistance n (%)	Resistance n (%)		
Ampicillin*	10µg	124(83)	181 (72)	1.81	0.019
Amox-Clav	30µg	9(6)	12 (5)	1.26	0.602
Cefotaxime	30µg	5 (3)	4 (2)	2.12	0.257
Ceftazidime	30µg	3 (2)	6(2.4)	0.82	0.793
Ceftriaxone	30µg	2 (1)	4 (2)	0.83	0.831
Chloramphenicol	30µg	38 (25)	55 (22)	1.20	0.444
Ciprofloxacin	5µg	0(0)	4 (2)	0.00	0.119
Gentamicin	10µg	3 (2)	3(1)	1.68	0.523
Sulphamethoxazole-Trimethoprim*	25µg	129 (86)	200 (80)	1.54	0.128
Tetracycline*	30µg	113 (75)	178 (71)	1.3	0.368

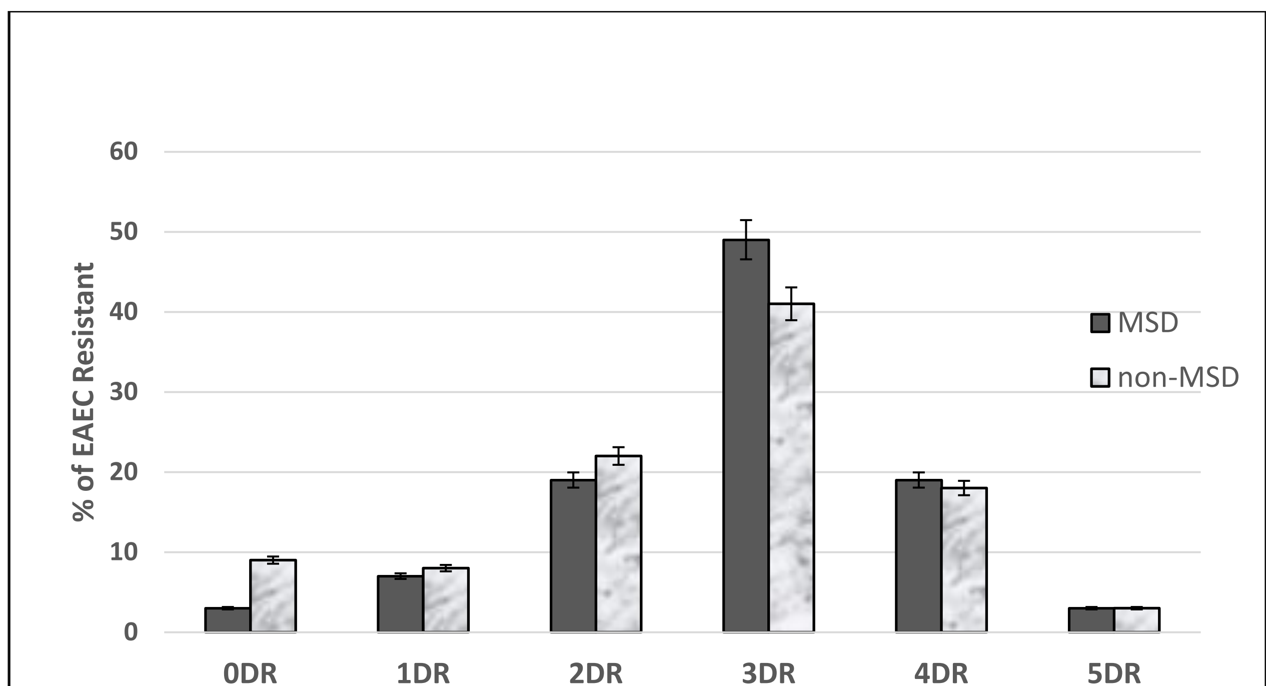
The EAEC resistance patterns against the ten antibiotics are almost equally distributed between diarrhoeal and non-diarrhoeal children with the high rate of resistance attributed to three antibiotics that include Ampicillin, Cotrimoxazole-trimethoprim and Tetracycline in the two groups.

Figure 7.1: Distribution of Multi-Drug-Resistant EAEC strains from case and control



Multiple drug resistance of two or more antibiotics is exceptionally high in all study children account for 90% and 83% in cases and controls respectively

Figure 7.2: Antibiotic multi-resistance patterns of EAEC strains from MSD and non-MSD children



The 0, 1 and 5 drug resistance (DR) were less than 10% in both MSD and non-MSD. However, three drug resistant (3DR) was proportionately higher 48% in MSD compared to 41% in non-MSD. There are no multi-drug resistance of six or more antibiotics observed.

Table 7.2: Antibiotic Resistant Pattern of EAEC strains from two age strata diarrhoeal and non-diarrhoeal children

Antimicrobial agent	**Resistant in age 0-11 month (n = 205)**			**Resistant in age 12-59 month (n =195)**		
	Case (n=83) n (%)	**Control (n=122) n (%)**	**P - value**	**Case (n=67) n (%)**	**Control (n=128) n (%)**	**P - value**
Ampicillin	75 (90.4)	97 (79.5)	0.037	49 (73.1)	84 (65.6)	0.284
Amoxicillin-Clavulanic Acid	7 (8.4)	3 (2.5)	0.051	2 (3.0)	9 (7.0)	0.244
Cefotaxime	4 (4.8)	4 (3.3)	0.576	1 (1.5)	0 (0)	--
Ceftazidime	3 (3.6)	5 (4.1)	0.860	0 (0)	1 (0.8)	--
Ceftriaxone	2 (2.4)	4 (3.8)	0.717	0 (0)	0 (0)	---
Chloramphenicol	22 (26.5)	28 (23.0)	0.560	16 (23.9)	27 (21.1)	0.655
Ciprofloxacin	0 (0)	2 (2.0)	--	0 (0)	2 (1.6)	---
Gentamicin	1 (1.2)	2 (1.6)	0.799	2 (3.0)	1 (0.8)	0.235
Sulphamethoxazole-Trimethoprim	76 (92.0)	101 (83.0)	0.072	53 (79.1)	99 (77.3)	0.778
Tetracycline	66 (79.5)	92 (75.4)	0.492	47 (70.2)	86 (67.2)	0.673

The resistance patterns of the EAEC strains to the ten antibiotics in the two age strata (0-11 months and 12-59 months) of unmatched case and control children were almost equally distributed. After adjusting for multiple testing (Bonferronni method) resistant to all of the ten antibiotics were not statistically significant between the case and control group in the two age strata. In the age strata 0-11 and 12-59 months the cut off p-value was (0.05/9) 0.0056 and (0.05/6) 0.0083 respectively

Table 7.3: Distribution of MDR EAEC strains among three age strata of study children

Age in Month	≥2 MDR (n=342)	≤1DR (n=58)	Odd Ratio	P - value
0-11	189 (55.26)	16 (27.59)	3.24	**0.00009**
12-23	123 (35.96)	30 (51.72)	0.52	0.022
24-59	30 (8.77)	12 (20.69)	0.36	0.006

Figure 7.3: Proportional distribution of MDR EAEC strains among age strata of study children (MSD and non-MSD combined)

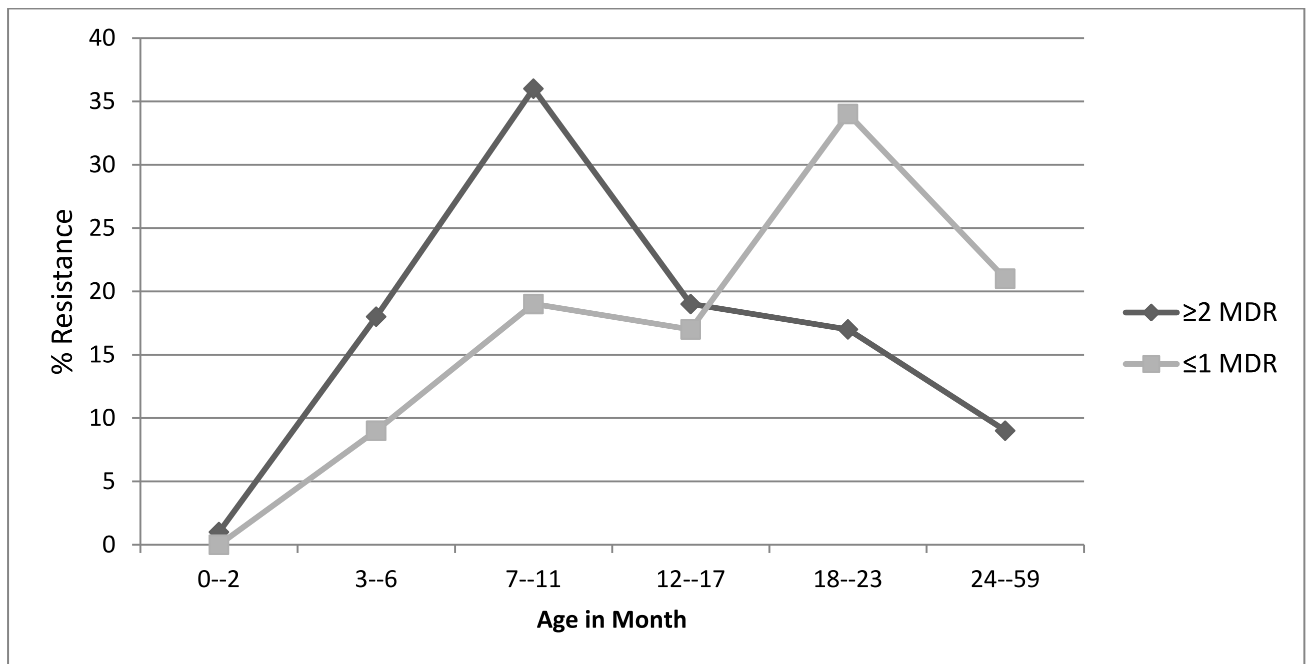
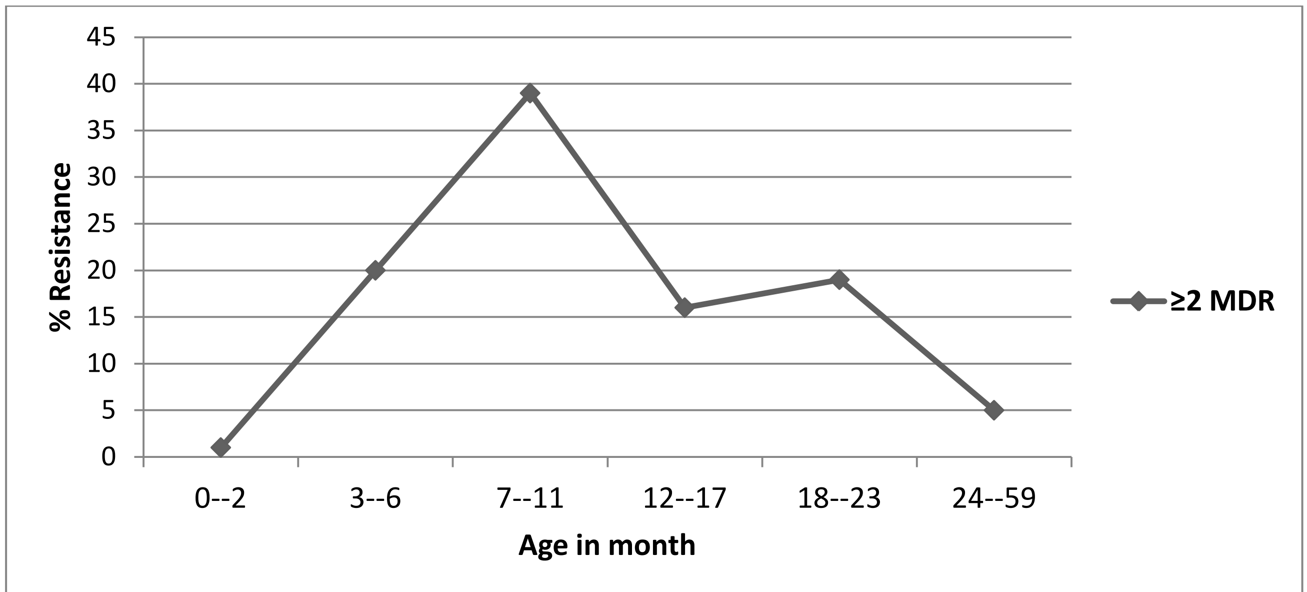


Figure 7.4: Rate of multi-drug-resistance in six age strata of study children revealed the age stratum 7-11 month that MDR is most common and there was some evidence (p-value 0.01) of a difference in resistance between ≥2MDR and ≤1DR. Similar finding was seen in figure 7.5 that is specific for only MSD children where sole ≥2MDR was investigated.

Figure 7.4: Proportional distribution of MDR-EAEC strains among age strata of MSD children only



It is obvious that two or more antibiotics are been resistant against by the EAEC isolates and the children most vulnerable are those within the age 7-11 month old.

Figure 7.5: Distribution of MDR EAEC among MSD and non-MSD study children

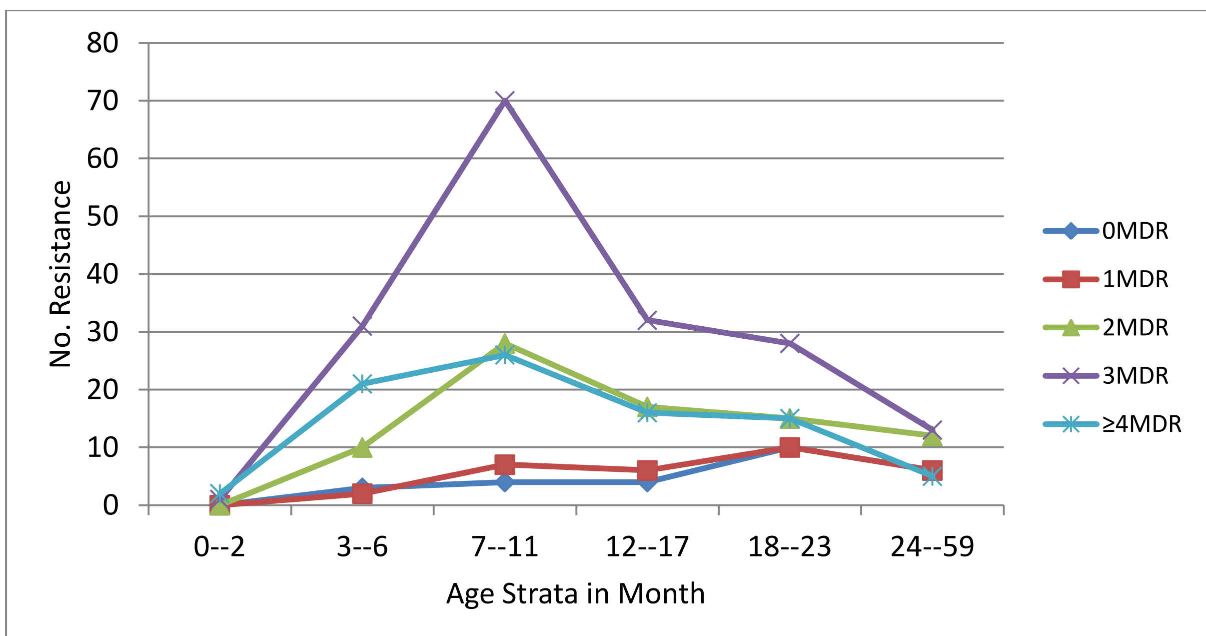


Figure 7.6: Distribution of MDR EAEC among MSD children

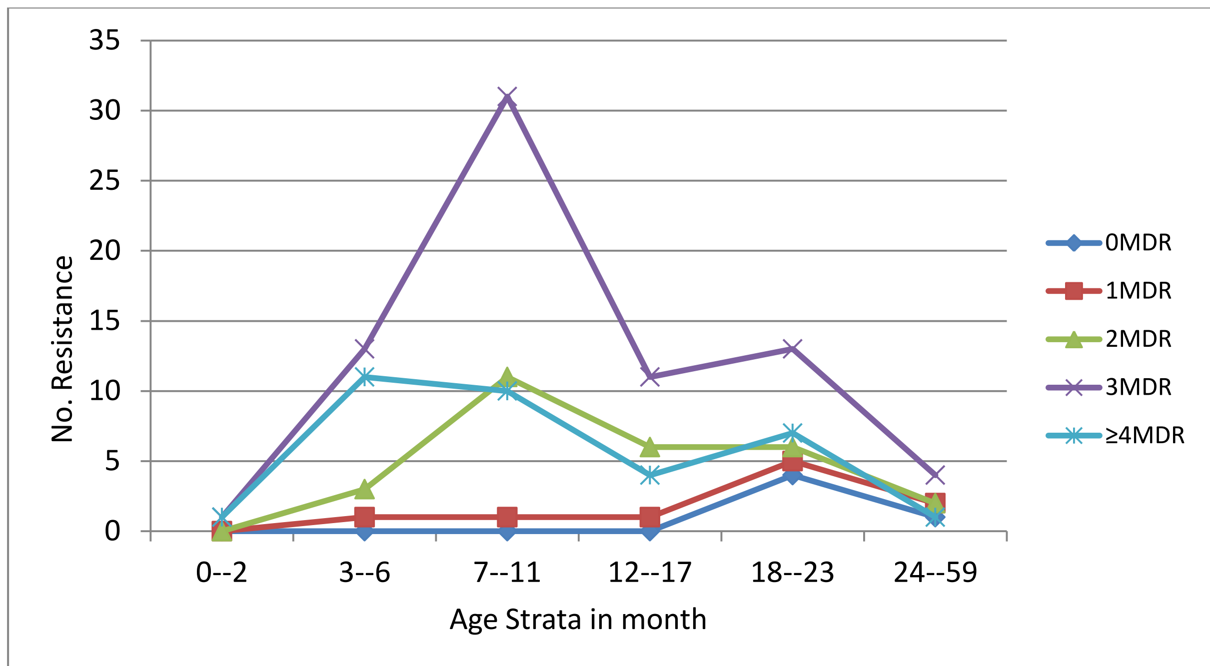


Figure 7.7: Distribution of MDR EAEC among non-MSD children

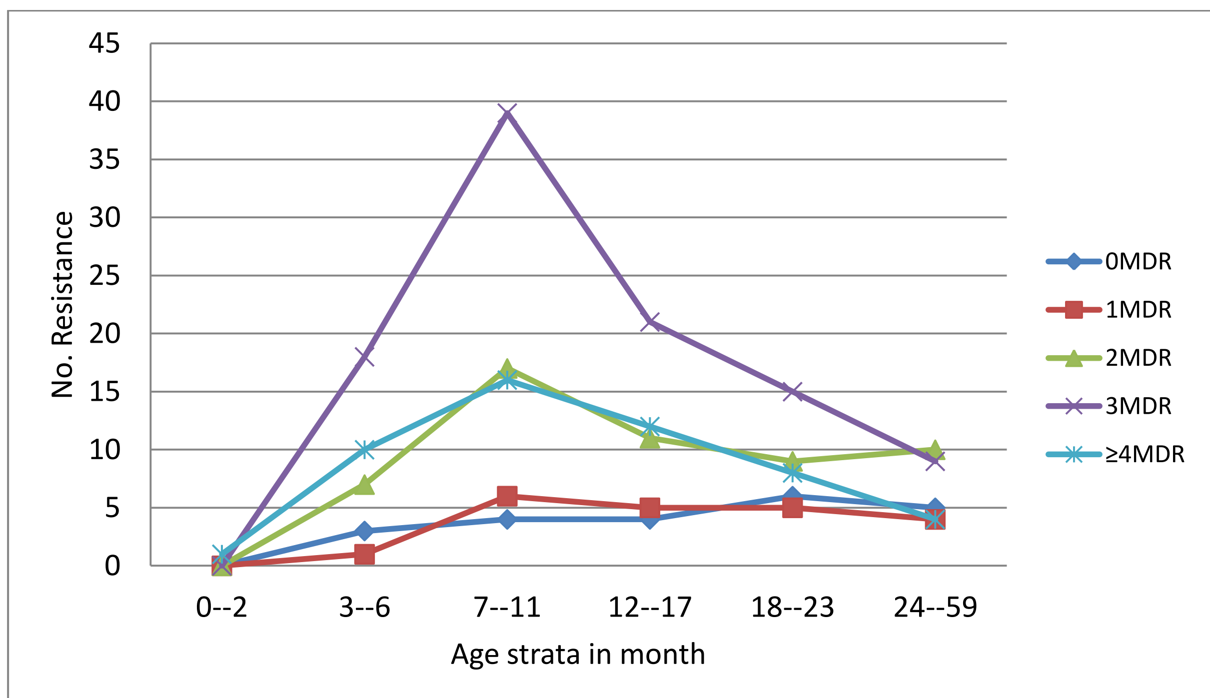


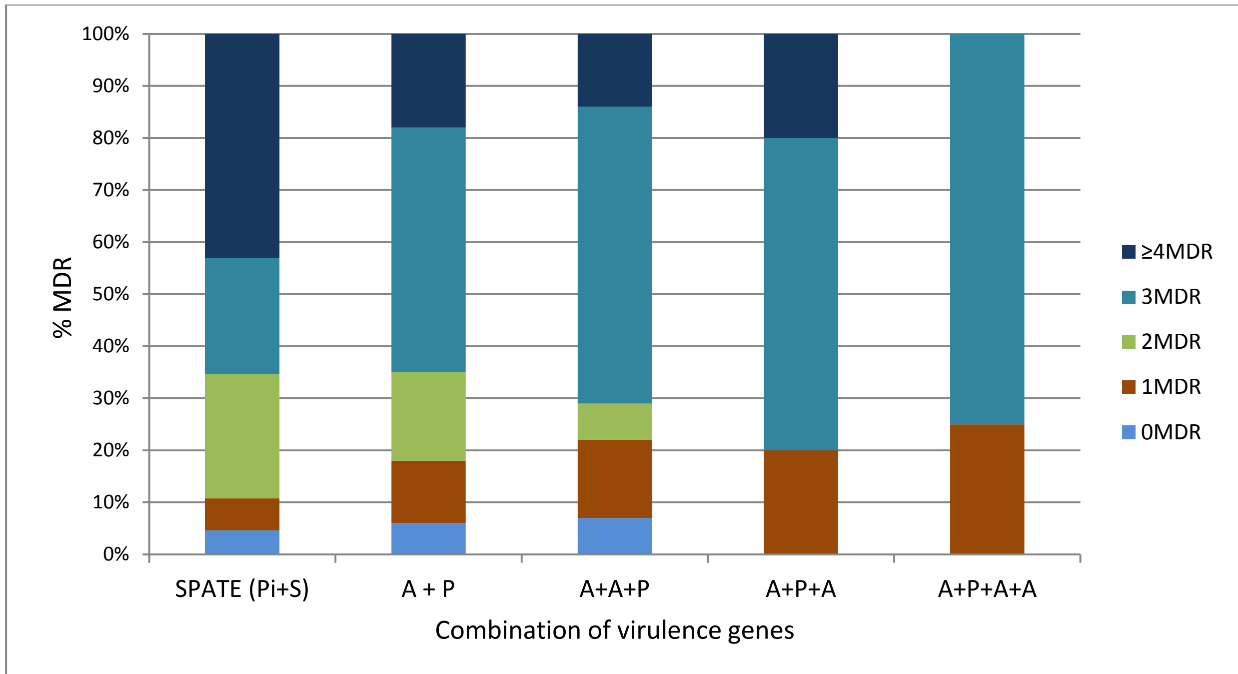
Figure 7.6, 7.7 and 7.8 showed the patterns of different categories of MDR among six age strata. Result obtained for the figure 7.6 that combined data for both MSD and non-MSD, figure 7.7 MSD only and figure 7.8 non-MSD only showed similar patterns of multi-drug-resistance in the six age strata with age stratum 7-11 month showing highest for 3MDR, 2MDR and ≥4MDR.

Table 7.4: Association between Multidrug resistant EAEC strains and the corresponding virulence factors

Virulence factor	≥2 MDR (n=342)	≤1 MDR (n=58)	Odd Ratio	P-value
aatA	115 (33.63)	15 (25.86)	1.45	0.243
aggR	219 (64.04)	34 (58.62)	1.26	0.429
aaP	166 (48.54)	27 (46.55)	1.08	0.779
ORF3	235 (68.71)	34 (58.62)	1.55	0.129
capU	223 (65.20)	35 (60.34)	1.23	0.474
aar	248 (72.51)	39 (67.24)	1.29	0.409
aafC	19 (5.56)	3 (5.17)	1.08	0.905
agg3/4C	132 (38.60)	15 (25.86)	1.80	0.062
agg3A	35 (10.23)	1 (1.72)	6.49	0.036
aafA	16 (4.68)	2 (3.45)	1.37	0.676
aggA	86 (25.15)	18 (31.03)	0.74	0.344
agg4A	28 (8.19)	3 (5.17)	1.63	0.427
astA	174 (50.88)	31 (53.45)	0.90	0.717
sat	73 (21.35)	9 (15.52)	1.47	0.309
sepA	100 (29.24)	9 (15.52)	2.24	0.029
pet	38 (11.11)	9 (15.52)	0.68	0.335
pic	123 (35.96)	17 (29.31)	1.35	0.325
sigA	37 (10.82)	9 (15.52)	0.66	0.299
aaiC	125 (36.55)	14 (24.14)	1.81	0.066
air	86 (25.15)	11 (18.97)	1.43	0.309
eilA	156 (45.61)	34 (58.62)	0.59	0.066

All of the 21 virulence genes studied were found not to be associated with MDR following adjustment for multiple testing (Bonferroni method) that resulted in cut off p-value (0.05/21) 0.002

Figure 7.8: Association of class II (pic & sepA) SPATE proteins and other specific gene combination with multi-drug-resistant EAEC strains



Key: (Pi+S) – Pic + SepA, A+P – aatA + Pet, A+A+P – aatA+aggR+Pet, A+P+A – aatA+Pet+aggA and A+P+A+A – aatA+Pet+aggR+aggA. Observation showed specific gene combination resulting into more MDR EAEC.

Table 7.5: Association of animals as risk factors for multidrug resistant EAEC strains from children

Animals	Multidrug Resistant EAEC strains			
	≥2 MDR (n=342) No. (%)	≤1 MDR (n=58) No. (%)	OR	P-value
Goat	263 (76.90)	45 (77.59)	0.96	0.908
Sheep	255 (74.56)	42 (72.41)	1.11	0.729
Dog	90 (26.32)	20 (34.48)	0.67	0.197
Cat	59 (17.25)	11 (34.48)	0.39	0.002
Cow	83 (24.27)	13 (22.41)	1.10	0.759
Rodent	245 (71.64)	39 (67.24)	1.23	0.495
Fowl	303 (88.60)	54 (93.10)	0.57	0.305
Donkey	178 (52.05)	21 (36.21)	1.91	0.025
Horse	113 (33.04)	10 (17.24)	2.36	0.015

After adjusting for multiple testing all the nine animals in the table 7.5 were not implicated as part of risk factors for the spread of MDR EAEC in human. The cut off p-value was (0.05/9) 0.0056

Table 7.6: Multidrug Resistance Patterns among study participant's recruitment zones (villages)

Zone (Village)	Multidrug Resistant EAEC strains			
	≥2 MDR (n=342) No. (%)	≤1 MDR (n=58) No. (%)	Odd Ratio	P-value
Zone-1	122 (35.67)	17 (29.31)	1.34	0.346
Zone-2	68 (19.88)	9 (15.52)	1.35	0.435
Zone-3	47 (13.74)	11 (18.97)	0.68	0.296
Zone-4	43 (12.57)	7 (12.07)	1.05	0.914
Zone-5	39 (11.40)	7 (12.07)	0.94	0.883
Zone-6	23 (6.73)	7 (12.07)	0.53	0.153

Following scrutiny of study participants base on their respective villages, data showed that MDR is evenly distributed among villages without any of the villages solely responsible for spread of MDR.

7.3 Discussion

EAEC usually associated with watery diarrhoea that is often persistent (Cohen, Nataro et al. 2005). In children antimicrobial therapy is used in cases of severe diarrheal disease to reduce the duration of illness, particularly if associated with persistent diarrhoea. Since the characterisation of the EAEC pathovar in 1987 (Nataro, Kaper et al. 1987), many clinical isolates of EAEC developed multiple antibiotic resistance (Bangar and Mamatha 2008). In our investigation, resistant EAEC strains to two or three antibiotics that include Ampicillin, Sulphametaxazole-trimethoprim and tetracycline were greater than eighty percent. This finding corroborate with studies conducted in low income countries where high levels of resistance to Tetracycline, Spectinomycin, Streptomycin, Sulphamehoxazole-trimethoprim and Ampicillin were reported (Sang, Oundo et al. 1997; Vila, Vargas et al. 1999; Aslani, Alikhani et al. 2011). Studies from high income countries such as Denmark and Japan have shown similar finding (Jensen 2017; Kubomura, Misaki et al. 2017).

This study, employed two antibiotic sensitivity testing methods which are disk diffusion on the 400 EAEC isolates and MIC e-test and on selected 20 EAEC isolates. The 100% concordance results obtained for the two methods showed that the less expensive disk diffusion method is reliable, valid and can still be recognised as a method of choice (table 7.1a) but required experienced laboratory personnel.

The highest number of antibiotics that constitute multi-resistance in this study was five and the most frequent and proportionately common was equal or greater than three antibiotics (≥3MDR) combined. Data from this study showed significant association of Ampicillin-resistance to moderate-to-severe diarrhoea among children (p-value 0.019) (table 7.1). This may be due to direct exposure of sick children to frequent use of Ampicillin drug compared to other drugs. There was high prevalence of resistance to Cotrimoxazole-trimethoprim and

Tetracycline but without indication of association to diarrhoea. In our quest to understand the distribution of resistant EAEC strains among the 10 antibiotics in two age strata (0-11 and 12-59 months) of both case and control groups, we observed that strains from diarrhoeal infant were more resistant to ampicillin yielded p-value 0.037 compared to other antibiotics. This is a moderate association consistent with previous studies (Nguyen, Le et al. 2005; Aslani, Alikhani et al. 2011).

To further identify factors associated with an increased risk of spread of multidrug-resistant EAEC strains we stratified age of studied children in to three (0-11, 12-23 and 24-59 months) and in to six (0-2, 3-6, 7-11, 12-17, 18-23 and 24-59). Results from the three age strata of children showed association of age 0-11 months with high rate of ≥2MDR given p-value 0.00009. In our analysis of six age strata of children, age 7-11 months was fund implacably associated with spread of ≥2MDR yielded p-value 0.010. Further analysis showed high frequency of 2MDR, 3MDR and ≥4MDR been associated with 7-11 months age stratum in all study children combined, MSD children and non-MSD children (figure 7.6, 7.7 & 7.8). This kind of association in this age group is rare in previous investigations although a study in Demark has shown similar association implicating 6 month old children as carrier of multiple drug resistant bacteria (Hebbelstrup Jensen, Roser et al. 2016). It is important to understand the risk factors involved in the association of multiple resistance strains in a particular age group 7-11 months. We speculate factors that include crawling period for children, underdeveloped immunity, non-exclusive breastfeeding, host genetic and genetic aspect of antimicrobial resistant bacteria and environmental factors may be the cause. At the crawling age increased mobility provides greater freedoms; however children may unknowingly infect themselves by coming into contact with unsanitary objects or surfaces. The immunity of children below 12 months is not fully developed, as a result the less developed host defence apparatus that are often overwhelmed by the infected antibiotic

resistant strains. Likewise, non-exclusive breast feeding or complete absence of breast feeding contributes, to lower immunity. In The Gambia rural community, the proportion of exclusive breastfeeding is about 55% (NaNA 2015).

In the genetics of antimicrobial resistance, intrinsic bacterial property and acquired resistance mechanisms play major role in the spreading of antibiotic resistance. The acquired bacterial antibiotic resistance often stems either from a mutation of cellular genes or the acquisition of foreign resistance genes or a combination of the two. In our findings, we speculate that acquisition of foreign resistance genes strongly associated with high cases of drug resistance strains in younger children via direct case-contacts that include mothers, siblings, other household members and animals.

Of the twenty-one virulence genes screened to identify virulence genes that may be associated with multiple antibiotic resistance, only two genes found to be possibly associated with MDR EAEC and they are *agg3A* and *sepA* with p-value 0.036 and 0.029 respectively. The two genes are plasmid genes, *agg3A* is adhesin producer and belongs to aggregative adherence fimbrial (AAF) family whilst *sepA* is a toxin and belongs to class II serine protease autotransporter (SPATE) protein family. Studies have implicated *sepA* as one of the important genes responsible for acute diarrhoea caused by EAEC among less than five years old from developing countries (Boisen, Scheutz et al. 2012; Ikumapayi, Boisen et al. 2017). No literature however, has reported possible association of the two genes with multiple drug resistance. Therefore, the mechanism of association with MDR requires further investigation. Additionally, observation from our analysis uncovered association of gene combination with MDR. For example, combination of *pic* (protein involve in colonisation) and *sepA* (Shigella extracellular protein A) showed association with MDR of four or more antibiotics (≥4MDR) whilst combination of gene *aatA+Pet*, *aatA+aggR+Pet*, *aatA+Pet+aggA* and *attA+Pet+aggA+aggR* are associated with MDR of three antibiotics (3MDR) and strongest in

the later (figure 7.9). Again, this pattern of gene combination has not been implicated to have association with MDR before now. Therefore, more investigations are needed to substantiate our findings.

The transfer of resistant bacteria to humans by farm animals is well recognised. Studies have shown that resistant bacteria in farm animals reach consumers through meat products (Bartlett, Gilbert et al. 2013). The implication of resistant bacteria on environmental microbiome is also well documented, about 90% of the antibiotics given to livestock are excreted in urine and stool which in turn widely dispersed through fertilizer, ground-water and surface runoff (CDC 2013). So, non-judicious use of antibiotic in animal can cause spread of resistant bacteria. Our study implicated two animals that include horse and donkey been associated with MDR EAEC of two or more (≥2MDR) antibiotics. The explanation that supports the likelihood association of the two animals with MDR was that these animals are well look after and often treated with antibiotics against infectious diseases as they are used as a form of transportation and farm work to generate income. Humans are known reservoir for EAEC infecting other humans however, study from BurkinaFaso has implicated animal such as cattle, chickens and pigs as EAEC reservoir (Kagambega, Martikainen et al. 2012).

Six villages were included in this study, study participants recruited from these villages were analysed for possible spread of MDR. However, statistical analysis of the distributions of MDR in the 6 villages showed even distribution of MDR in all the villages (table 7.6). Our finding was that no particular village was implicated solely in the spread of MDR EAEC strains.

Many studies have demonstrated that the Extended-spectrum beta-lactamase (ESBL)-producing Enterobacteriaceae is increasing (Overdevest, Willemsen et al. 2011) and has led

to numerous cases of healthcare associated infections and many death in developed countries such as United State of America (CDC 2013). However, 56% (28/50) of the 50 randomly selected EAEC isolates for whole genome sequencing revealed only blaTEM-1B gene which belongs to beta-lactamase families predominant in 1980s and early 1990s and are presently recognise as non-ESBLs derivatives (Paterson and Bonomo 2005). Interestingly, there was no ESBLs CTX-M complex or CTX-M15 among the 50 EAEC isolates sequenced which suggest infrequent used of quinolone or other third generation antibiotics in The Gambia rural regions both in human and animal likely, due to limited source of income to purchase the expensive third generation antibiotics.

The high incidence of multi-drug resistant isolates of EAEC can be attributed to many factors that include sub-lethal exposure of bacteria to antibiotics. This often occurs due to indiscriminate use of antibiotics in humans and animals leading to immense selective pressure on bacteria, a phenomenon that accelerates the emergence of bacterial resistance (Goldstone and Smith 2017), such as horizontal gene transfer (HGT) by acquisition of resistance genes from other microorganism and the spread of antibiotic resistance from chromosomal mutations.

7.4 Conclusion

Antimicrobial drug resistant bacteria are widespread across the globe. Many national and international public health organisations have described the rapid emergence of resistant bacteria as a crisis that could result into a catastrophic situation. Our study has epidemiologically showed infants less than 1 year old, particularly, those between 7-11 months are at risk of been overwhelmed with multidrug-resistant EAEC and similar trend is speculated for other multidrug-resistant enterobactriaceae. Additionally, the study has

demonstrated that animals such as donkey and horses and even cattle can be responsible for the spread of MDR bacteria in developing countries. Although a bigger and carefully planned study is recommended to validate our observation.

To prevent the continued emergence and spread of MDR many factors need to be employed primarily through antimicrobial stewardship and reducing the indiscriminate use of antibiotics in farming. Additional measure that encourage reference laboratories, referral hospital laboratories and regional health laboratories to establish a continuous and sustainable antimicrobial surveillance platform that has the potential to transform national and regional public health action on the burden of antimicrobial drug resistant bacteria.

Chapter 8: Whole Genome Sequencing of Selected EAEC strains from case and control study children

8.1 Introduction

Identification and characterization of microbial pathogens are pivotal for our understanding of host-pathogen interaction, diagnosis, drug target and drug development and safety of patients. Unfortunately, many available molecular tests are unable to detect emerging genetic features in rapidly evolving infectious agents that spread in humans, animals and the environment. Unrecognised pathogens can cause outbreaks that can put patients and healthcare providers at risk (Deurenberg, Bathoorn et al. 2017).

During the last 20 years, molecular diagnostic methods have evolved tremendously and played important roles in medical microbiology laboratories. Sequence analyses can be used to answer different diagnostic questions, such as the genetic relationship of either bacteria or viruses, the detection of mutations that cause resistance against antivirals and or antibiotics in viral and bacterial genomes and identification of bacteria through sequence analyses of the 16S rDNA (Bush 2013). **Sanger sequencing** (first generation sequencing) use the principle of amplifying single gene or genomic region using specific primers (Deurenberg, Bathoorn et al. 2017). This same sequencing approach is applied for the identification of pathogens in clinical samples. Unfortunately, the approach has poor specificity if use on non-invasive samples such as faeces which has normally multiple microbial species. In such cases, results obtained by Sanger sequencing are unreliable. Also, the cost of Sanger sequencing required to achieve the investigational tasks needed is high and the turnaround time is long. The standard Sanger sequencing (traditional dideoxynucleotide chain termination) identifies linear sequences of nucleotides by electrophoretic separation of randomly terminated extension products (Rizzo and Buck 2012). The reactions can read DNA fragments of 500 bp to 1000

bp in length, and this method is still in used routinely for sequencing small amounts of DNA fragments and it is the gold standard for clinical cytogenetics (Kingsmore and Saunders 2011). However, the challenge is the requirement for electrophoretic separation of DNA fragments for reading DNA sequence content in Sanger-based sequencing becomes a bottleneck for the throughput, increasing time and limiting the number of reactions that can be run in parallel (Rizzo and Buck 2012).

Next-generation sequencing (NGS) refer to as high-throughput DNA sequencing technologies which are capable of sequencing large numbers of different DNA sequences in a single reaction for example in parallel. All NGS technologies monitor the sequential addition of nucleotides to immobilized and spatially arrayed DNA templates but differ substantially in how these templates are generated and how they are interrogated to reveal their sequences (Linnarsson 2010). NGS allows sequencing of the whole genome of numerous pathogens in one sequence run, either from bacterial isolates of different patients, or from multiple species present in patient material from one individual (metagenomics). Advantage of NGS to Sanger sequencing is that a single protocol can be used for all pathogens for both identification and typing applications. Thus, NGS has proven useful in medical microbiology laboratories and for infection prevention measures (Zhou, Lokate et al. 2016). There are three basic methodological steps of NGS which are; Template preparation (genomic DNA or cDNA, library preparation and library amplification), Sequencing and imaging (either Ion torrent PGM [personal Genome Machine] that uses pH change or MiSeq which uses fluorescence) and Data analysis (Grada and Weinbrecht 2013). However, different sequence platform vendors have devised different strategies to prepare the sequence libraries into suitable templates as well as to detect the signal and ultimately read the DNA sequence. So the different strategies to generate the sequence reads also lead to differences in the output capacity for the different platforms (Buermans and den Dunnen 2014).

NGS studies of microorganisms typically follow one of two general strategies: targeted amplicon sequencing (TAS) or whole genome sequencing (WGS) (figure 8.1). Target Amplicon sequencing approach uses target-specific primers for PCR-mediated amplification, so that the genomic regions of interest are enriched and selectively sequenced. This approach is often used to interrogate well-characterized genomic regions to identify known drug resistant mutant as well as disease-causing mutations for diagnosis of pathological conditions. Sequencing for *de novo* assembly of whole genome relies on non-targeted library preparation or fragmentation. This method is usually performed when microorganism are unknown and or when the aim is to determine the genomic content and functional potential of the organism under investigation (Lefterova, Suarez et al. 2015).

Whole genome sequencing (WGS) is a molecular method that determines the exact order of nucleotides present in a given DNA or RNA (Grada and Weinbrecht 2013). The method often applied to primary or complex polymicrobial specimens that include clinical, environmental and food specimens for the identification of culture-independent pathogen as well as characterization of the microbial population. In addition, WGS of bacterial genomes divulge the presence of antimicrobial resistance (AMR) genes, virulence genes or genes associated with virulence and pathogenicity and to discover new genetic mechanisms that dictate bacterial virulence (Rossen, Friedrich et al. 2018). It is on this bedrock we performed WGS on 51 EAEC in order to determine these isolates genetic content. The first complete genome sequence of EAEC, targeting strain 042, the prototypical member of EAEC-pathotype was achieved in 2009 by Chaudhuri and colleagues revealing major features of the *E. coli* 042 genome (Chaudhuri, Sebaihia et al. 2010). Also, the study showed genomic and phylogenetic comparisons of *E. coli* 042 with other *E. coli* strains leading to the detection of previously uncharacterised virulence factors.

In this study, we aim to explore WGS to investigate the antimicrobial resistance genes contain in the EAEC isolates which phenotypic resistance assay has been conducted, and do a correlation between the data generated from both phenotypic (disk diffusion) and genotypic (WGS) characterisation to evaluate the concordance of resistance. Also, we will study the distribution of Multi-locus Sequence Types (MLST) circulating among the EAEC isolates in the rural Gambia. Additional investigation includes genetic relationship of the EAEC strains.

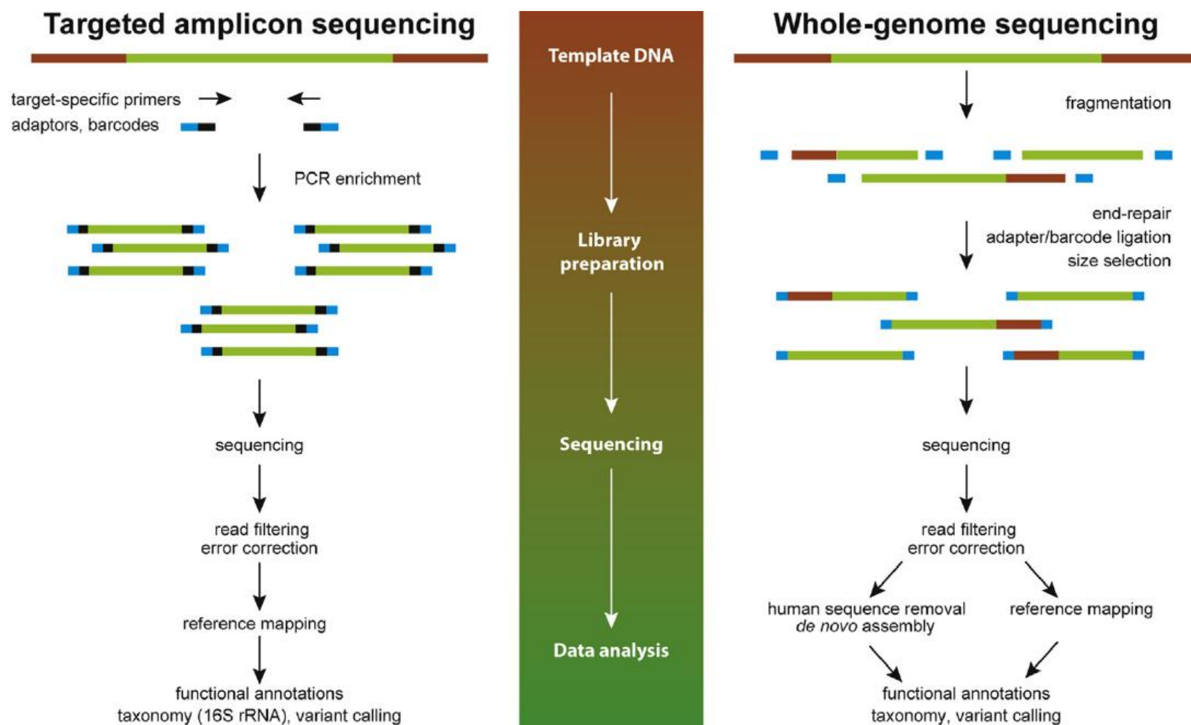


Figure 8.1: Sketch sequencing approaches for diagnosis and monitoring of infectious diseases. Adapted from (Lefterova, Suarez et al. 2015).

Targeted amplicon sequencing uses target-specific primers for template (green bars) enrichment, followed by primers that are partially complementary to the target-specific primers (black bars) and contain sequencing adaptors and bar codes (blue bars).

Whole-genome sequencing uses enzymatic or mechanical fragmentation, followed by end pair to allow ligation of primers that contain sequencing adaptors and bar codes (blue bars) Size selection allows only fragments of a predefined length to be used for sequencing. However, Bioinformatics removal of human sequences is required because the nucleic acids of the organism of interest often constitute <1% of the nucleic acid pool. Fragmentation libraries can also be made from PCR-enriched amplicon.

The application of Whole-genome shotgun sequences (WGS) involves sampling the chromosomes that make up one genome, and the WGS assembly which is the reconstruction of sequence up to chromosome length done by computer software that generate the results shown in the tables and figures in the result section of this chapter. De novo assembly does not require reference genome

8.2 Results

The application of next-generation sequencing (NGS) was used to conduct whole-genome sequencing using MiSeq-Illumina, we sequenced 51 EAEC isolates out of which 50 turned out to be genuine *E. coli* as one isolate resulted in *Stenotrophomonas maltophilia*. *E. coli* reference genome NC_011748 *E. coli* 55989 chromosome with 5,154,862 bp in size was included in the analysis as reference strain. The summary of the WGS result regarding genome content is shown in the table 8.1. The contig (contiguous consensus sequence derived from the assembly of many short, overlapping DNA fragments) readings for the DNA sequenced were of good quality for genome interogation. The percentage Guanine Cytosine (G+C) contents were optimal. Also, the isolates N50 (half of the total length of the collection of all contigs), protein-coding sequence (CDS) or coding region and percentage aligned bases were obtained. Overall, the averages of genome contents or parameters yielded useful information about the EAEC isolates genome. For example the average for contigs, reads, GC, bp, N50, CDS, aligned bases and %-aligned bases were 439.29, 689729, 51.16%, 5120782, 64127.35, 4773.18, 4359349 and 84.57% respectively (Table 8.1). None of the aligned base for each of the DNA isolates sequenced genome was more than 4743331bp whilst the reference aligned base was 5154862 well above each of the 50 DNA isolates. The number of virulence genes per isolates ranged from 55 to 110, and number of resistance genes per isolates ranged from 1 to 15. However, the total number of genes observed in each of the study EAEC isolates ranges from 4,182 to 6,097.

Resistance genes

This study revealed thirty-seven (37) different resistance genes. From WGS data we look at the distribution of the identified resistance genes in diarrhoeal and non-diarrhoeal children to ascertain association with diarrhoea but results obtained showed no association except for a Trimethoprim resistance gene dfrA14_1 which showed p-value 0.010. Also, there are three resistance genes that were only found in diarrhoeal children, and these are $bla_{TEM\text{-}1C_}$, $bla_{SHV\text{-}1_18}$, and Sul1_1 although in small number. Also, three different resistance genes that include blaOXA-1_1, aac(3)-IIa and dfrA17_1 were detected only among the non-diarrhoeal children (Table 8.2) which are in small number too.

Comparison of resistance between presence of resistance genes using whole genome sequencing and resistance using phenotypic (disk diffusion) assay showed significant correlations between phenotypic antimicrobial resistance and each antimicrobial resistance genes that were correspond to antibiotics tested. The correlation investigation showed 100%, 100%, 76%, 97% and 95% concordance with Ampicillin, Gentamicin, Chloramphenicol, Sulphamethoxazole-trimethoprim and Tetracycline respectively (Table 8.3).

Multi-locus Sequence Type (MLST)

The MLST analysis of structural seven housekeeping genes (*adk, fumC, gyrB, icd, mdh*, *purA* and *recA*) revealed twenty-nine distinct sequence types (STs) and thirteen clonal complexes (ST-complex) from the 50 EAEC isolates sequenced. The most frequent ST was ST38 which was 16% (8/50) and distributed evenly among diarrhoeal and non-diarrhoeal children. The next common ST in the study was ST10 which account for 6% (3/50) two of which were

found in non-diarroeal children. Also there are 12 dual STs and 15 single STs, majority of these were unevenly distributed among cases and control (Table 8.4 & Table 8.5).

Sequence type and Resistance genes

Resistance gene *Sul2_2* is found in all the ST38 while the *Tet-D* and *blaTEM* genes account for seven and six respectively in the ST38. *Tet-B* gene was found in the 3 ST10 but *Tet-A* was found in two of the three ST10. The two ST3032 has six resistance genes that include *blaTEM, CatA1, Sul1_2, Sul2_2, dfrA12* and *Tet-A*. The two ST111 has four resistance genes that include *blaTEM, CatA1, Sul2_3* and *Tet-A*. Also the two ST131 has *blaTEM, Sul2_2, dfrA8_1* and *Tet-B* in common, while the two ST156 has *blaTEM, Sul1_2, dfrA1_30* and *Tet-A* in common. At least one resistance gene is found common to two *ST349, ST394, ST678, ST1291* and ST3018 which are *Tet-B, blaTEM, Sul2_2, dfrA1_1* and *dfrA1*_1 respectively (table 8.6).

Pairwise core SNP distance is graphed in histogram format as well as in phylogeny as representation of the SNPs distances (figure 8.2 and figure 8.3). The farthest SNP is at position 487 and contains 2% SNPs while the highest percentage of SNPs (20%) is at distance position 280 (figure 8.3). The core SNP density revealed 471951 SNPs across the 5.2 million base pair (bp). The genome position of the highest SNPs density (3,948 SNPs) is at 2.34M-2.36M, while the lowest SNPs density (166 SNPs) is at genome location 3.42M (figure 8.4).

Pan genome graph computed using Roary software showed 51 taxa and 22,045 clusters that include core and accessory genes (figure 8.5). The dark left side of the graph showed genes

common to all isolates, hence, the darker the blue colour to the left side the more common the genes so, this group of genes are regarded as core-genes. Also, the paler the blue colour to the right side the more differences are the genes and this group of genes are regarded as accessory genes. The highest number of genes registered was 6,097 contain in isolate SID100590 while the lowest gene recorded was 4,182 found in isolate SID100096. The pan genome phylogenetic tree was constructed using a distance matrix based on the presence or absence of genes in the isolates (figure 8.6).

The annotated phylogenetic tree showed relatedness of the strains and their commonality details regarding year the strains isolated, sequence type, sequence type complex, number of antibiotic resistant to and resistance genes contained (figure 8.7a and 8.7b). The annotation was done using interactive tree of life (Letunic and Bork 2016) *an online* (https://itol.embl.de/personal_page.cgi) *tool for the display and annotation of phylogenetic and other trees.*

Plasmid investigation

Of the 50 EAEC isolates investigated to identify genetic components using whole genome sequencing (WGS), 47 harbours plasmid. Thirty-four of the 47 isolates harbours plasmid associated with resistance genes. However, 32 EAEC isolates have antibiotic resistance genes associated with the detected plasmids. Coincidentally, the 32 EAEC consist of 16 cases and 16 controls and 10 distinct plasmid associated resistance genes were distributed among the 32 EAEC isolates (table 8.7). The three EAEC isolates that had no plasmid are females of two controls and a case with age 6, 19 and 28 months respectively.

Table 8.1: Summary of genome content of the 50 Isolates Examined

Sample ID	Year of Isolation	Origin in Gambia	ST	Contigs	Reads	G+C Content	bp	N50	CDS	Aligned Bases	% Aligned Bases
100020	2007	102	10	341	752558	51.7	5092123	41988	4790	4523874	87.76
100096	2008	101	2162	135	809278	50.6	4632118	76820	4133	4226716	81.99
100119	2008	103	14	349	1076290	50	5318912	46267	5026	4177708	81.04
100125	2008	102	2707	1101	779624	51.8	4879617	8432	4426	4657004	90.34
100127	2008	102	131	174	748868	50.8	5025243	114540	4722	4130745	80.13
100138	2008	103	200	175	608330	50.8	5053235	109100	4812	4572848	88.71
100182	2008	101	73	174	630814	50.8	5049330	118871	4655	4130142	80.12
100191	2008	104	58	129	981712	51.3	4775355	131337	4406	4569449	88.64
100342	2008	105	131	216	688330	51	5160817	145157	4877	4107319	79.68
100382	2008	104	394	869	620850	51.9	4900370	10481	4403	4210767	81.69
100404	2008	102	2622	236	1025872	49.1	4985608	107951	4579	4268305	82.8
100415	2008	103	38	254	896092	50.7	5404046	110276	5038	4475169	86.81
100427	2008	103	222	1460	872504	51.1	5729908	9144	5411	4743331	92.02
100503	2008	103	2067	356	911726	51.1	5244148	49441	4976	4636867	89.95
100569	2008	102	10	259	870236	50.8	5127539	72495	4843	4443136	86.19
100590	2008	102	3444	1972	510362	49.5	6344745	6760	6054	4523394	87.75
100715	2008	105	73	265	571534	50.3	5180187	49693	4799	4132871	80.17
100722	2008	101	6907	140	946636	51.5	4655989	88452	4324	4245256	82.35
100796	2008	105	3032	132	613194	50.7	5171336	124870	4724	4408160	85.51
102031	2009	101	38	279	566012	50.6	5188333	43014	4818	4375354	84.88
102098	2009	101	38	307	472684	50.9	5365437	57209	4996	4448305	86.29
102106	2009	101	38	315	612080	50.8	5367356	48111	4990	4457968	86.48
102191	2009	104	2704	502	488414	49.5	5179127	22265	4883	4135098	80.22
102274	2009	101	517	123	959836	50.7	4825593	112857	4498	4518971	87.66
102296	2009	102	38	846	481110	50.9	5291567	12763	4909	4447504	86.28
102375	2009	104	394	2052	366674	50.6	5845401	5032	5522	4516037	87.61
102425	2009	101	3018	233	567074	51.2	5055435	51977	4650	4422738	85.8
102469	2009	102	3018	1123	275602	51.4	4936793	6989	4483	4363593	84.65
102602	2009	103	222	1265	658718	51.4	4876817	7239	4456	4626879	89.76
102705	2009	101	1312	367	351646	50.8	4888741	24553	4565	4406444	85.48
102742	2009	102	111	316	484450	50.9	5008195	35975	4713	4533387	87.94
102806	2009	101	3032	193	507192	51.1	5168696	65709	4724	4394450	85.25
102820	2009	103	111	165	725744	51.5	5017709	117311	4732	4541878	88.11
102871	2009	103	58	220	863704	50.8	5046801	76384	4724	4641815	90.05
102906	2009	103	678	329	719200	51.1	5187229	44460	4880	4952539	96.08
102951	2009	106	31	222	678740	51.1	5200279	75077	4838	4402428	85.4
103016	2010	106	678	205	747746	50.1	5085193	78163	4773	4864334	94.36
103047	2010	106	349	195	779672	51.5	5097266	91200	4798	4376479	84.9
103069	2010	106	1291	313	739394	51.1	5139442	51283	4852	4557229	88.41
103070	2010	106	1291	1521	488756	49.1	4989253	5034	4734	4523081	87.74

Continue

Sample ID	Year of Isolation	Origin in Gambia	ST	Contigs	Reads	G+C Content	bp	N50	CDS	Aligned Bases	% Aligned Bases
103076	2010	101	2141	288	827428	50.9	5400256	75086	5079	4344936	84.29
103275	2010	101	156	305	487564	50.6	4991005	34622	4686	4551111	88.29
103276	2010	104	156	159	731316	51.4	4991009	99209	4692	4560074	88.46
103278	2010	104	10	193	836276	50.9	4910239	66617	4582	4452867	86.38
103400	2010	101	38	197	548974	50.3	5234471	127805	4854	4421501	85.77
103446	2010	101	196	94	1272482	50.5	4583263	126431	4245	4446534	86.26
103530	2010	102	38	423	840518	50.2	5390647	37619	5014	4478512	86.88
103691	2010	101	349	193	473444	51	5094603	85955	4785	4362597	84.63
103693	2010	102	2178	316	410882	51.1	4914690	34201	4590	4552339	88.31
103709	2010	103	38	275	622328	50.9	5335918	65014	4958	4434509	86.03
AVERAGE				439	689729	51.16%,	5120782	64127	4773	4359349	84.57%

Origin in The Gambia

Key:

101 – Basse,

102 – Gambisara,

103 – Koina

104 – Fatoto

105 – Yorobawol

106 – DamphaKunda

Table 8.2: Distribution of Antimicrobial resistance genes among diarrhoeal and non-diarrhoeal children

Group of Antimicrobial agent	Antimicrobial agent	Resistance gene	Case (n = 23) n (%)	Control (n = 27) n (%)	Total (n = 50) n (%)	OR (95% CI)	P-value
Beta-lactams	Ampicillin	*bla*OXA-1_1	0 (00)	2 (7)	2 (4)	--	0.182
		bla_{TEM-1B_1}	14 (61)	14 (52)	28 (56)	1.44(0.407-5.194)	0.522
		bla_{TEM-1C_5}	1 (4.4)	0 (00)	1 (2)	--	0.273
		bla_{SHV-1_18}	1 (4.4)	0 (0.00)	1 (2)	--	0.273
Aminoglycoside	Gentamicin	*aac(3)-IIa*	0 (0.00)	1 (3.70)	1 (2)	--	0.351
	Streptomycin	*strA_1*	4 (17.4)	4 (14.8)	8 (16)	1.21(0.196-7.418)	0.804
		strA_4	7 (30)	10 (37)	17 (34)	0.74(0.190-2.815)	0.623
		strB_1	13 (56.5)	13 (48.2)	26 (52)	1.40(0.398-4.956)	0.554
Phenicol	Chloramphenicol	*CatA1_1*	8 (35)	9 (33)	17 (34)	1.07 (0.279-4.024)	0.914
Sulfonamide	Sulfonamide	*Sul1_1*	1 (4.35)	0 (0.00)	1 (2)	--	0.273
		Sul1_2	7 (30)	7 (26)	14 (28)	1.25 (0.301-5.156)	0.723
		Sul2_2	14 (61)	11 (37)	25 (50)	2.64 (0.732-9.718)	0.092
		Sul2_3	3 (13)	3 (11)	6 (12)	1.20 (0.144-9.947)	0.834
Trimethoprim	Trimethoprim	*dfrA12_1*	1 (4)	2 7)	3 (6)	0.57 (0.009-11.746)	0.649
		dfrA14_1	5 (22)	0 (00)	5 (10)	--	0.010
		dfrA17_1	0 (00)	2 (7)	2 (4)	--	0.182
		dfrA1_1	2 (9)	6 (22)	8 (16)	0.33 (0.030-2.191)	0.193
		dfrA1_30	4 (17)	2 (7)	6 (12)	2.63 (0.329-31.369)	0.278
		dfrA7_1	3 (13)	1 (4)	4 (8)	3.90 (0.281-212.447)	0.225
		dfrA8_1	3 (13)	4 (15)	7 (14)	0.86 (0.112-5.801)	0.857
Tetracycline	Tetracycline	*tet(A)_4*	8 (35)	5 (18)	13 (26)	2.34 (0.540-10.855)	0.191
		tet(B)_4	7 (30)	6 (22)	13 (26)	1.53 (0.357-6.674)	0.509
		tet(D)_1	6 (26)	3 (11)	9 (18)	2.82 (0.504-19.501)	0.169

Bonfire is 0.05/22 = 0.002 cut off

Table 8.3: Comparison between Genome (WGS) and Phenotypic (Disk diffusion) based predictive Antimicrobial Susceptibility Test (AST)

Class of Antimicrobial agent	**Antimicrobial agent**	**Resistance gene**	**Frequency of Resistance gene**	**Frequency Phenotypic resistance (Disk diffu)**	**Rate of Concordance between phenotypic and genotypic detection of resistance**	
Beta-lactams	Ampicillin	*bla*$_{TEM-1B_1}$	28	28	100%	100%
		bla$_{TEM-1C_5}$	1	1	100%	
		bla$_{SHV-1_18}$	1	1	100%	
Aminoglycoside	Gentamicin	*aac(3)-IIa*	1	1	100%	100%
Phenicol	Chloramphenicol	*CatA1_1*	17	13	76%	76%
Sulphamethoxazol-Trimethoprim	Sulfonamide	*Sul1_1*	1	1	100%	97%
		Sul1_2	14	13	93%	
		Sul2_2	25	25	100%	
		Sul2_3	6	6	100%	
	Trimethoprim	*dfrA12_1*	3	3	100%	
		dfrA14_1	5	5	100%	
		dfrA17_1	2	2	100%	
		dfrA1_1	8	7	88%	
		dfrA1_30	6	5	84%	
		dfrA7_1	4	4	100%	
		dfrA8_1	7	7	100%	
Tetracycline	Tetracycline	*tet(A)_4*	13	13	100%	95%
		tet(B)_4	13	11	85%	
		tet(D)_1	9	9	100%	

To obtain concordance for each resistance gene we applied formular;

Percentage concordance = Phenotypic resistance number/Genotypic-resistance genes number X 100

For example, in the case of Sul1_2 resistance gene; 13/14 X100 = 92.86% = 93%   OR

100/14 = 7, (100 – 7 = 93).

Table 8.4: showing MLST detail that results in STs and ST-complex

sid	*age	gender	ST	*adk*	*fumC*	*gyrB*	*icd*	*mdh*	*purA*	*recA*	ST Complex
100020	6	F	10	10	11	4	8	8	8	2	10
100096		F	2162	31	58	54	10	11	35	47	none
100119	7	M	14	14	14	10	14	17	7	10	14
100125	5	M	2707	9	23	64	18	11	8	219	278
100127	5	F	131	53	40	47	13	36	28	29	131
100138	5	M	200	6	4	5	26	7	8	14	40
100182	5	F	73	36	24	9	13	17	11	25	73
100191	6	M	58	6	4	4	16	24	8	14	155
100342	2	M	131	53	40	47	13	36	28	29	131
100382	15	F	394	21	35	61	52	5	5	4	394
100404	17	M	2622	13	363	10	97	17	94	93	None
100415	9	F	38	4	26	2	25	5	5	19	38
100427	15	F	222	9	6	15	56	11	8	6	None
100503	21	F	2067	6	95	3	18	11	122	2	none
100569	20	F	10	10	11	4	8	8	8	2	10
100590	22	M	3444	76	24	9	13	17	28	25	73
100715	9	M	73	36	24	9	13	17	11	25	73
100722	19	F	6907	10	929	4	8	8	2	2	None
100796	9	F	3032	54	22	211	342	40	16	4	None
102031	2	F	38	4	26	2	25	5	5	19	38
102098	8	F	38	4	26	2	25	5	5	19	38
102106	6	M	38	4	26	2	25	5	5	19	38
102191	6	M	2704	53	~400	47	13	36	28	29	131
102274	21	F	517	109	65	5	1	9	13	14	269
102296	4	F	38	4	26	2	25	5	5	19	38
102375	3	M	394	21	~35	~61	52	5	5	4	394
102425	6	F	3018	12	58	54	344	1	2	47	None
102469	22	F	3018	12	58	54	344	1	2	47	None
102602	21	M	222	9	6	~15	56	11	8	6	None
102705	6	M	1312	6	11	4	8	8	78	2	None
102742	8	F	111	6	29	14	16	24	8	2	None
102806	18	F	3032	54	22	211	342	40	16	4	None
102820	18	M	111	6	29	14	16	24	8	2	none
102871	9	F	58	6	4	4	16	24	8	14	155
102906	33	M	678	6	6	5	136	9	7	7	None
102951	24	M	31	18	22	17	6	5	5	4	31
103016	19	F	678	6	6	5	136	9	7	7	None
103047	56	M	349	34	36	39	87	67	16	4	349
103069	56	M	1291	10	11	4	8	8	5	2	None
103070	58	M	1291	10	11	4	8	8	5	2	None
103076	3	F	2141	101	88	~262	281	59	215	196	None
103275	5	M	156	6	29	32	16	11	8	44	156

Continue											
sid	***age**	**gender**	**ST**	**adk**	**fumC**	**gyrB**	**icd**	**mdh**	**purA**	**recA**	**ST Complex**
103276	43	M	156	6	29	32	16	11	8	44	156
103278	6	M	10	10	4	577	8	8	8	2	10
103400	28	F	38	4	26	2	25	5	5	19	38
103446	28	F	196	6	19	3	16	9	8	6	None
103530	9	M	38	4	26	2	25	5	5	19	38
103691	11	F	349	34	36	39	87	67	16	4	349
103693	7	M	2178	9	6	15	56	11	26	6	None
103709	16	F	38	4	26	2	25	5	5	19	38

Key: *age in month, gender: F – Female, M – Male

Following reanalysis of the whole genome sequencing data to obtain sequence types (STs), it became clear that allel fumC with ~400 and ~35 of SID 102191 and 102375 respectively, and allel gyrB with ~61 of SID 102375 were generated due to poor sequencing so they are sequencing artifacts. However, allel gyrB with ~15 and ~262 of SID 102602 and 103076 respectively were generated as new allels.

Table 8.5: Distribution of ST among Cases and Controls

ST (n = 29)	Case (n = 23)	Control (n = 27)	Total (n = 50) (%)
10	1	2	3 (6)
14	1	0	1 (2)
31	0	1	1 (2)
38	4	4	8 (16)
58	1	1	2 (4)
73	0	2	2 (4)
111	1	1	2 (4)
131	0	2	2 (4)
156	2	0	2 (4)
196	1	0	1(2)
200	0	1	1(2)
222	1	1	2 (4)
349	1	1	2 (4)
394	1	1	2(4)
517	0	1	1(2)
678	1	1	2 (4)
1291	0	2	2 (4)
1312	0	1	1(2)
2067	1	0	1 (2)
2141	1	0	1 (2)
2162	0	1	1 (2)
2178	1	0	1(2)
2622	0	1	1(2)
2704	1	0	1 (2)
2707	1	0	1(2)
3018	1	1	2(4)
3032	1	1	2 (4)
3444	1	0	1 (2)
6907	0	1	1(2)

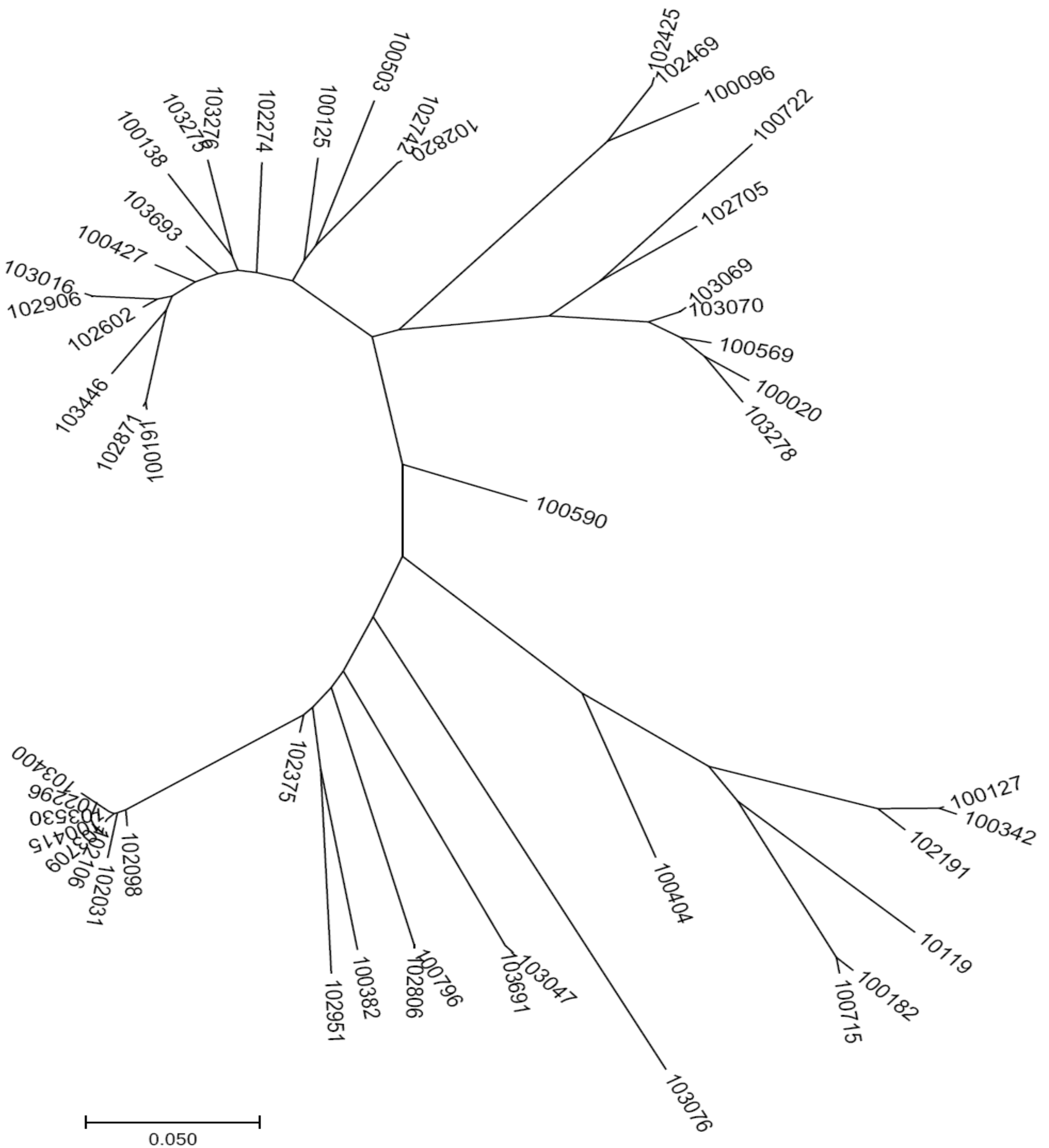


Figure 8.2: Core SNP Phylogeny showing the evolutionary relationships of taxa. The evolutionary history was inferred using the Neighbour-joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length = 2.09214678 is shown. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura, Peterson et al. 2011) and are in the units of the number of base substitutions per site. The analysis involved 50 nucleotide sequences. All positions containing gaps and missing data were eliminated. These were a total of 1159 positions in the final dataset. Evolutionary analyses were conducted in (MEGA7) Molecular Evolutionary Genetic Analysis (Kumar, Stecher et al. 2016).

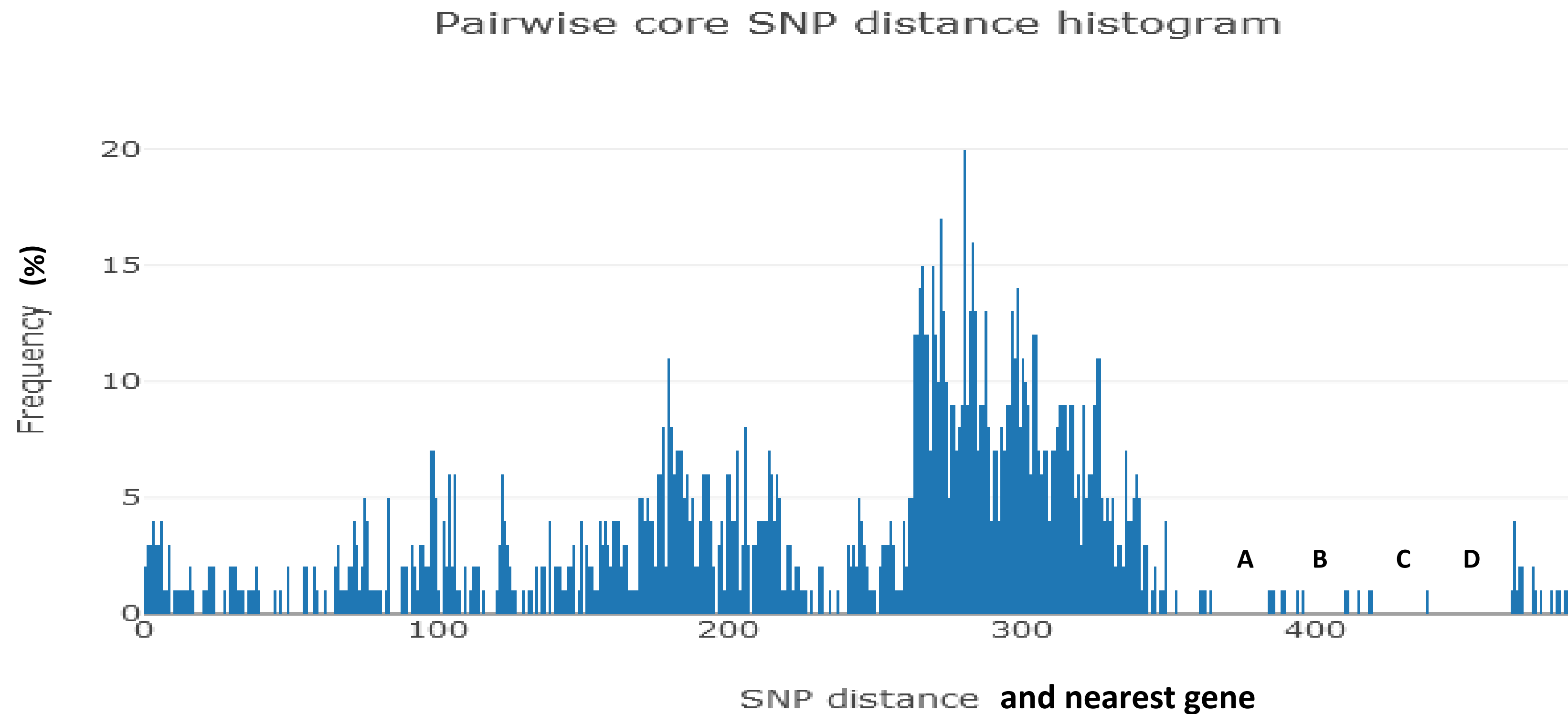


Figure 8.3: Showed distribution of SNPs according to their proximity to the nearest gene. Bars represent percentage of SNPs that have a gene within a certain distance from each other. The most distanced SNPs (2% SNPs) are at position 487. Similarly, 0 the closest distance interestingly has 2% SNPs are at this 0 position. The highest number of SNPs (20 SNPs) are at distance position 280 away from 0 distance. There are four wide gaps in the distances without SNPs and the wide gaps were observed at the distance position 365.5-383.5, 397.5-409.5, 420.5-437.5 and 439.5-466.5 of **A**, **B**, **C** and **D** respectively with **D** been the widest.

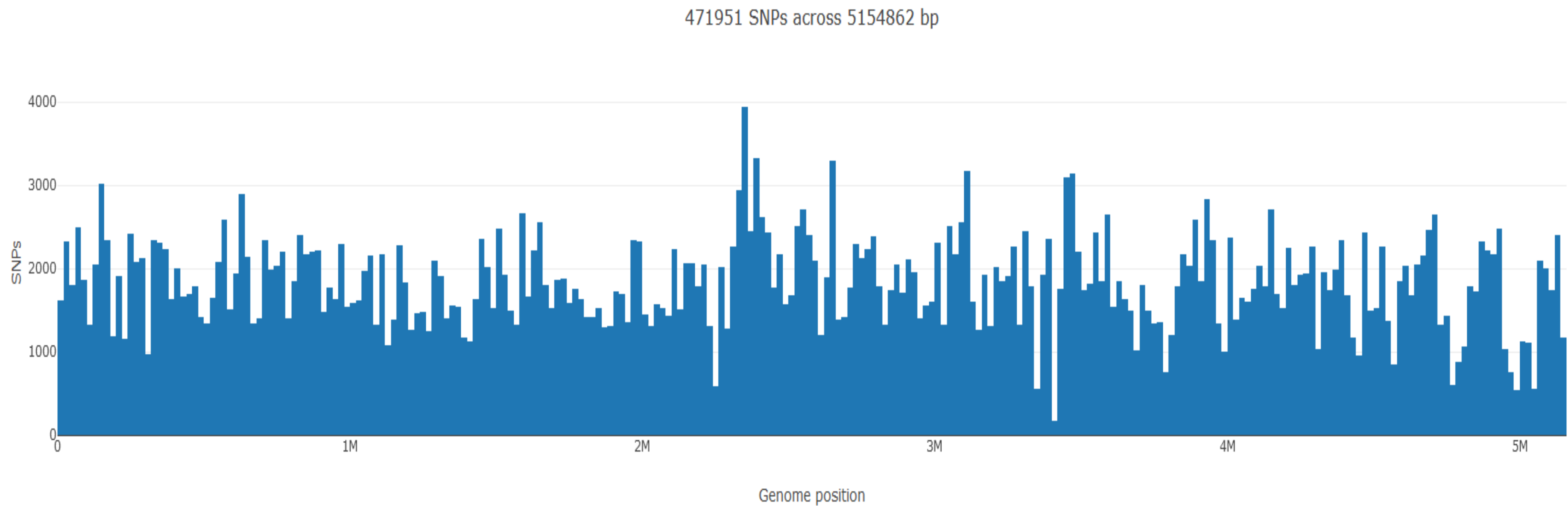
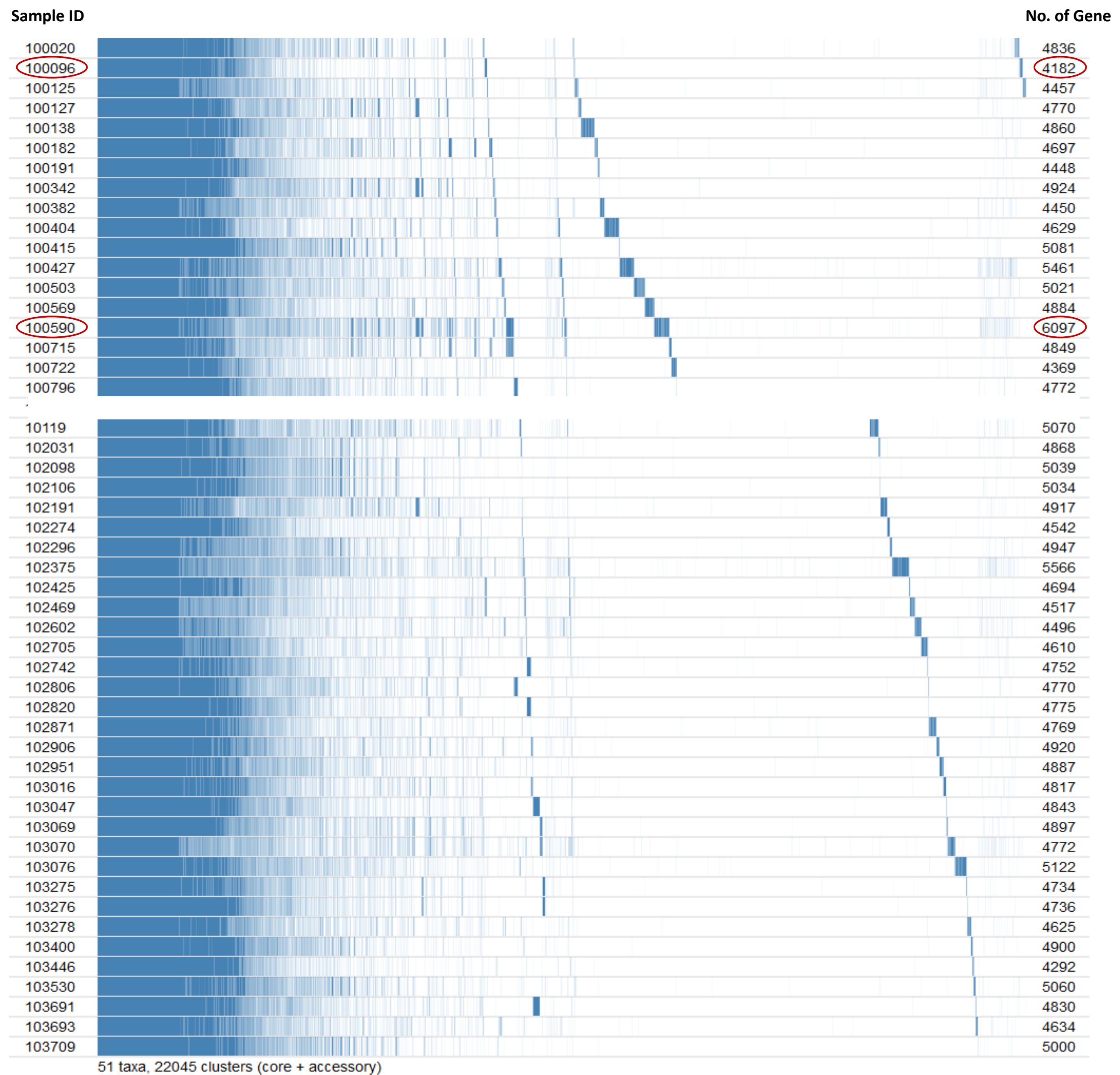


Figure 8.4: Core SNP density revealed 471951 SNPs across the approximate 5.2 million base pair. Each bar represents 20 thousand base pair (20k bp) at their respective genomic position containing different number of single nucleotide polymorphism (SNPs). The first bar at genome position 0-20k has 1616 SNPs, whilst the last bar at genome position 5.14M-5.16M has 1173 SNPs. The bar at position 2.34M-2.36M has the highest number of SNPs which is 3948 whilst the bar at position 3.4m-3.42m has the lowest number of SNPs which is 166.

Figure 8.5: Pan genome heat mapping showing the core genes (common to all isolates) and the accessory genes that are shared between isolates, determined by a pairwise comparison.



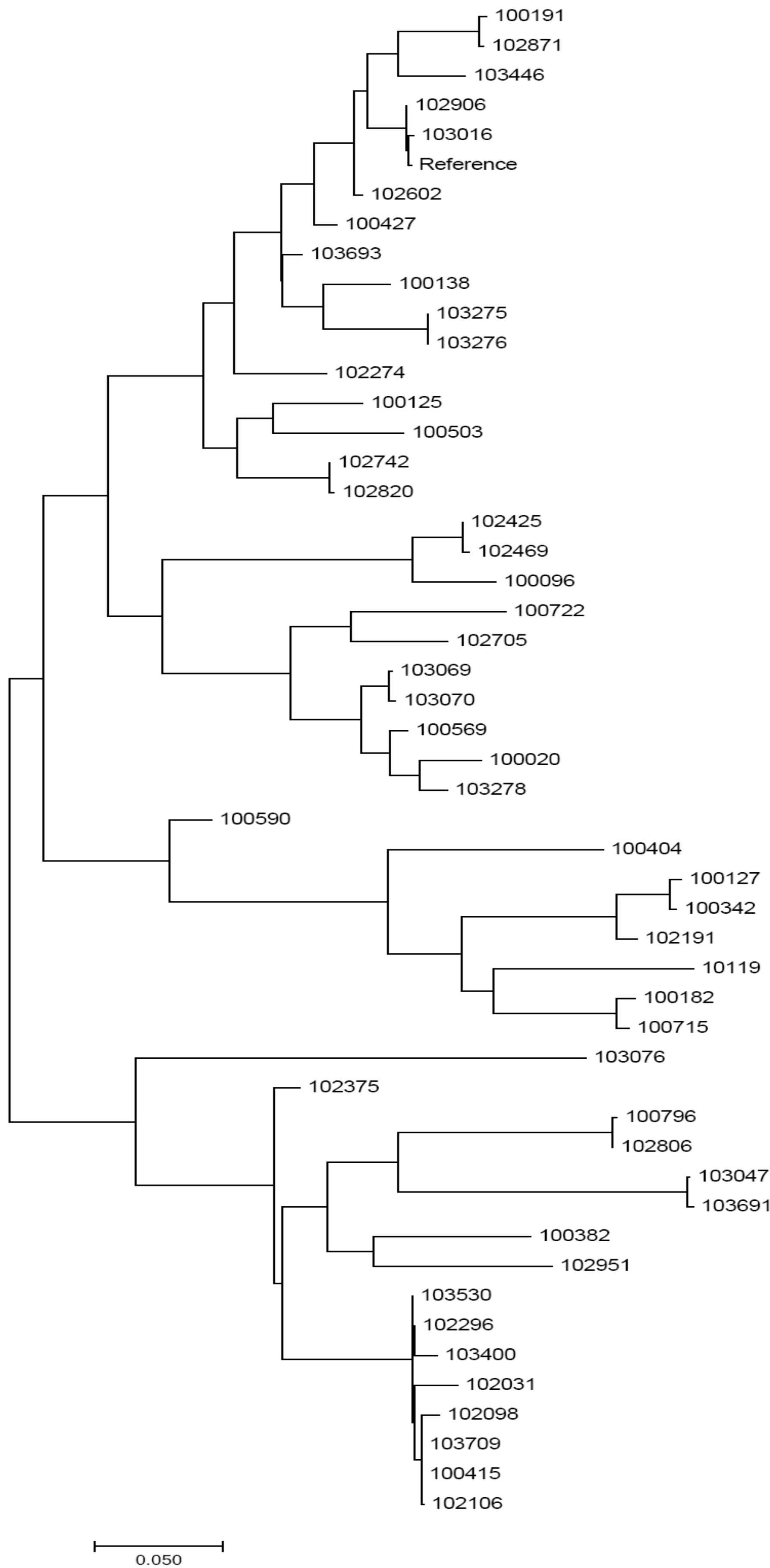


Figure 8.6: Molecular Phylogenetic analysis by Maximum Likelihood method

The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei 1993). The tree with the highest log likelihood (-13441.40) is shown. Initial tree(s) for the heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach, and then selecting the topology with superior log likelihood value. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 51 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 1159 positions in the final dataset. Evolutionary analyses were conducted in MEGA7 (Kumar, Stecher et al. 2016).

Reference strain used was NC_011748 *Escherichia coli* 55989 chromosome, complete genome of 5154862 bp

Table 8.6: Distribution of Resistance genes against the Sequence Types (STs)

ST	Case-Control	Ampicilli		Chl or	Genta	Sulfonamide			Trimethoprim						Tetracyline			Drug Resis-tant
		bla OXA	bla TEM	CatA 1	aac(3)-IIa	Sul 1_2	Sul 2_2	Sul 2_3	dfrA12	dfrA14	dfrA17	dfrA 1_1	dfrA 1_30	dfrA8_1	Tet-A	Tet-B	Tet-D	
10	Case	0	0	0	0	1	0	0	0	0	0	0	1	0	1	1	0	**2**
10	Control	1	1	1	0	0	0	0	0	0	0	0	0	1	1	1	0	**4**
10	Control	1	0	1	0	0	1	0	0	1	0	0	0	0	0	1	0	**4**
14	Case	0	0	0	0	0	1	0	0	1	0	0	0	0	0	1	0	**2**
31	Control	0	1	1	1	1	0	0	0	0	1	0	0	0	1	0	0	**4**
38	Case	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	**3**
38	Case	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	**5**
38	Case	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	**3**
38	Case	0	1	0	0	0	1	0	0	0	0	1	0	0	0	0	1	**3**
38	Control	0	1	1	0	0	1	0	0	0	0	1	0	0	0	0	1	**4**
38	Control	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	**1**
38	Control	0	1	1	0	0	1	0	0	0	0	0	0	0	0	0	1	**3**
38	Control	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	1	**4**
58	Case	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**2**
58	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
73	Control	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0	**2**
73	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
111	Case	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	**4**
111	Control	0	1	1	0	0	0	1	0	0	0	0	0	0	1	0	0	**4**
131	Control	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	**2**
131	Control	0	1	1	0	1	1	0	0	0	0	0	0	1	0	1	0	**4**
156	Case	0	1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	**4**
156	Case	0	1	0	0	1	0	0	0	0	0	0	1	0	1	0	0	**2**
196	Case	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
200	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**

Continuation of Table 8.6

ST	CaseCon	Ampicilli		Chlor	Genta	Sulfonamide			Trimethoprim						Tetracyline			Drug Resis-tant
		bla OXA	bla TEM	Cat A1	aac(3)-IIa	Sul 1_2	Sul 2_2	Sul 2_3	dfrA12	dfrA14	dfrA17	dfrA 1_1	dfrA 1_30	dfrA8_1	Tet-A	Tet-B	Tet-D	
222	Control	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	**2**
222	Case	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
349	Case	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	**3**
349	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	**2**
394	Control	0	1	0	0	1	0	1	0	0	0	0	1	0	0	0	0	**3**
394	Case	0	1	1	0	0	1	0	0	0	0	0	0		0	1	1	**3**
517	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
678	Control	0	1	0	0	0	1	0	0	0	0	0	0	1	0	1	0	**3**
678	Case	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	**1**
1291	Control	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	**0**
1291	Control	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	**0**
1312	Control	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	0	**3**
2067	Case	0	0	0	0	1	1	0	0	1	0	0	1	0	0	0	1	**3**
2141	Case	0	1	1	0	1	1	0	0	0	0	0	0	1	0	1	0	**5**
2162	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
2178	Case	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	**3**
2622	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**
2704	Case	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	**3**
2707	Case	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	**2**
3018	Control	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	**3**
3018	Case	0	0	0	0	0	1	0	0	0	0	1	0	0	1	0	0	**2**
3032	Case	0	1	1	0	1	1	0	1	0	0	0	0	0	1	0	0	**4**
3032	Control	0	1	1	0	1	1	0	1	0	0	0	0	0	1	0	0	**5**
3444	Case	0	1	1	0	1	1	0	0	1	0	0	0	1	0	1	0	**4**
6907	Control	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	**0**

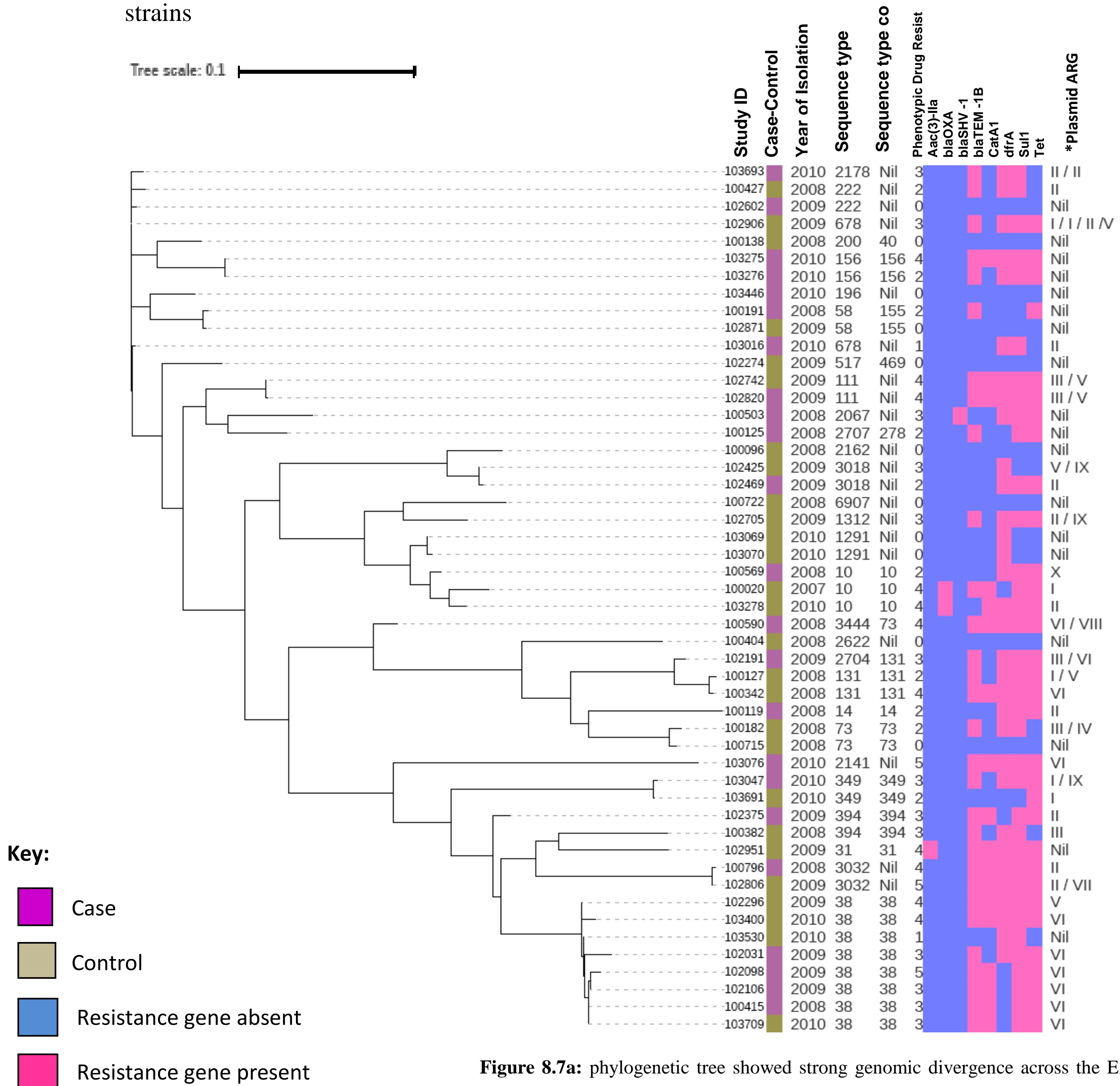
Key: 1 – Present; 0 – Absent

Table 8.7: Distribution of plasmid Associated Resistance Gene in EAEC isolates among cases and controls (n=32)

Plasmid-Associated-Resistance-Genes	Cases (n=16)	Controls (n=16)	Total (n=32)	P-value
IncFIC(FII)_1	1 (6)	5 (31)	6 (19)	1.070
ColRNAI_1	6 (38)	6 (38)	12 (38)	1.000
IncQ1_1	2 (13)	3 (19)	5 (16)	0.626
IncFIA_1	0 (0)	1 (6)	1 (3)	0.309
IncFIB(AP001918)_1	1 (6)	5 (33)	6 (19)	0.056
IncFIB(pB171)_1_pB171	7 (44)	3 (19)	10 (31)	0.127
IncFII(pCoo)_1_pCoo	0 (0)	1 (6)	1 (3)	0.309
IncI2_1	1 (6)	0 (0)	1 (3)	0.309
IncFII(pRSB107)_1_pRSB107	1 (6)	2 (13)	3 (9)	0.544
IncFII(pHN7A8)_1_pHN7A8	1 (6)	0 (0)	1 (3)	0.309

None of the ten distinct plasmids resistance genes is significantly associated with diarrhoeal children. The most prevalent plasmid associated resistance gene detected were incompatibility types (Inc) account for 90% (9/10) while the only compatible type (ColRNAI_1) account for 10% (1/10). Among the plasmids incompatibility type, the IncF plasmids preponderate 78% (7/9). Of the 28 IncF; IncFII, IncFIA, IncFIB and IncFIC account for 18% (5/28), 4% (1/28), 57% (16/28) and 21% (6/28) respectively.

Figure 8.7a: Phylogenetic tree showing relatedness of EAEC strains



<u>*Plasmid Associated Resistance Genes;</u>

I - IncFIC(FII)_1 **II** - ColRNAI_1

III - IncQ1_1 **IV** - IncFIA_1

V - IncFIB(AP001918)_1 **Nil** – No PlasmidARG

VI - IncFIB(pB171)_1_pB171

VII - IncFII(pCoo)_1_pCoo **VIII** - IncI2_1

IX - IncFII(pRSB107)_1_pRSB107

X - IncFII(pHN7A8)_1_pHN7A8

Figure 8.7a: phylogenetic tree showed strong genomic divergence across the EAEC isolates genome as few of the isolates are closely related. For example, isolates of study ID 103275 and 103276 are very close share the same clone, same distance, possess similar resistance genes and having similar ST/ST-complex 156/156. Similarly, study ID 100415, 102106 and 103709 share the same clone, distance, ST/ST-complex 38/38 and resistance-genes in common. Other than the 8 isolates that have ST38 and share the same clone, the rest of the isolates are highly diverse. Furthermore, 9 (82%) of the 11 isolates that had phenotypic zero (0) resistance to the eight antibiotics also had the corresponding resistance-genes absent in their genome following whole-genome-sequencing.

Plasmid: Thirty-two of the 50 sequenced EAEC isolates harboured antibiotic resistance genes associated with the detected plasmids. The most common plasmid associated resistance gene detected were incompatibility types (Inc) account for 90% (9/10) while the only compatibility type (CoIRNAI_1) account for 10% (1/10). Also, the figure above showed strong association of ST38 with plasmid IncFIB and strongest with IncFIB(pB171)_1_pB171 which accounts for 75% (6/8).

Figure 8.7b: Phylogenetic tree showing relatedness of EAEC strains in circular format

Key:

Case

Control

Resistance gane absent

Resistance gene present

8.3 Discussion

EAEC is globally recognised as a cause of paediatric persistent diarrhoea and travellers diarrhoea among adult arriving from EAEC endemic region (Knutton, Shaw et al. 2001; Cohen, Nataro et al. 2005; Kaur, Chakraborti et al. 2010; Estrada-Garcia, Perez-Martinez et al. 2014). We applied WGS to study EAEC in order to understand genetic factors that are likely responsiblefor the pathogen heterogeneity characteristics. In the past, very few studies have adopted WGS to demonstrate genomic components and phylogenetic diversity of EAEC isolates particularly from West-Africa region. We have shown in our previous study the contribution of EAEC virulence genes in infant diarrhoea (Ikumapayi, Boisen et al. 2017). It is interesting to report that 238 virulence genes identified by WGS were mostly evenly distributed among the diarrhoeal and non-diarrhoeal children, and some virulence genes are found more in controls than in cases. Thus, the virulence genes result in this study corroborate reports from past studies regarding heterogeneous nature of EAEC.

Following the discovery of the EAEC pathovar many clinical isolates of EAEC developed multiple antibiotic resistance (Bangar and Mamatha 2008), and high levels of resistance to tetracycline, Spectinomycin, Streptomycin, Sulphamethoxazole-trimethoprim and Ampicillin (Sang, Oundo et al. 1997; Vila, Vargas et al. 1999; Nguyen, Le et al. 2005; Mendez Arancibia, Pitart et al. 2009; Aslani, Alikhani et al. 2011; Ali, Ahmed et al. 2014). Similar, outcome is recorded in this study as our both clinical and environmental EAEC isolates showed over 80% resistance to Ampicillin, Sulphamethoxazol-Trimethoprim and Tetracycline, whilst resistant to Chlorampenicol was 23%. Although resistance to third generation antibiotics that include fluoroquinolone and cefolosporin were negligible even though variety of studies from geographically distinct areas have reported high rate of EAEC resistance to quinolone (Vila, Vargas et al. 2001; Khoshvaght, Haghi et al. 2014).

Also, it is amazing to discover that majority of the detected resistance genes in this study were found to be more evenly distributed among cases and control children, although resistant gene *dfrA14_1* was exclusively found in diarrhoeal children alone with p-value 0.010. A study in the past corroborates prevalence of the dfrA14_1 among bacterial infected individuals (Park 2018). Our study compared genetic (presence or absence of resistance genes) with resistant scores obtained from phenotypic disk diffusion technique that determined bacterial resistance and found that the concordance between the two methods was 94% combined. In the resent time this is one of the few studies in West-Africa that have looked at the concordance of the presence of resistance genes and compared to phenotypic disk diffusion method. Among the class of antibiotics that include bectalactams, aminoglycoside, phenicol, Sulphamethoxazol-Trimethoprim and tetracycline the concordance rate was 100%, 100%, 76%, 97% and 95% respectively (table 8.3). The possible explanation for the antibiotics that did not yield 100% concordance is that the organism did not express the resistance genes present thereby resulting in organism showing sensitivity in the phenotypic disk diffusion method. Also it was interesting to see nine isolates that showed zero resistance score to eight antibiotics in the disk diffusion method,also WGS showed absence of the corresponding resistance genes (table 8.6 and figure 8.7a). Limited studies have shown similar report among bacterial enterobacteriacae family and little is known of any such report regarding EAEC pathotype. However, the high concordance rate cannot be enough to justify use of whole genome sequence revealing presence of antibiotic resistance gene(s) to predict or determine bacterial isolates as resistant to an antibiotic until a multisite study is considered to conduct standard and thorough investigation. Moreover, in 2017, the outcome of European Committee on Antimicrobial Susceptibility Testing (EUCAST) subcommittee scientific meeting on the role of WGS in antimicrobial susceptibility (AST) of

bacteria issued ten recommendations that should be considered prior adopting use of WGS-inferred susceptibility to guide clinical decision making (Ellington, Ekelund et al. 2017).

In order to determine the linkage between diarrhoeal and non-diarrhoeal EAEC strains we obtained ST result from MLST data generated by WGS. A total of 29 different Sequence types (STs) were identified among the 50 EAEC isolates sequenced, and these 29 STs were heterogeneously disseminated among 23 diarrhoeal and 27 non-diarrhoeal children. Fifteen of the EAEC isolates have distinct ST types, whilst twelve STs that include ST58, ST73, ST111, ST131, ST156, ST222, ST349, ST394, ST678, ST1291, ST3018 and 3032 contained in 2 isolates, ST10 identified in 3 isolates and ST38 was identified in eight isolates. The commonest ST in the study was ST38 which constitute 16% of the STs and belongs to phylogroup A. All the ST38 produced ESBL *bla*$_{TEM}$ except one from healthy child, and 50% of the ST38 are linked with diarrhoea in children less than 10 month old. Two of the ST38 linked with diarrhoea are from the same peri-urban community and were isolated the same month and year and resistant to at least 3 drugs which are Ampicillin, Chloramphenicol, Sulphamthoxazol-trimethoprim and Tetracycline (figure 8.7a). ST38 has previously been described in some studies. For example, a study investigating EAEC STs among Nigerian diarrhoeal and healthy children showed ST38 and other clonal complexes with predicted ancestors ST10, ST23 and ST31 possesses both pathogenic and non-pathogenic EAEC strains (Okeke, Wallace-Gadsden et al. 2010). Similarly report obtained from a study in China emphasised the heterogeneous distribution of ST10, ST38 and ST131 among clinical and environmental samples with EAEC isolates (Zhang, Gu et al. 2016). However, the majority of Nigeria isolates in the Okeke's study possessed ST10-complex EAEC that were associated with diarrhoea in children older than 1 year. The ST10 in our study belong to phylogenetic group-A, this finding corroborates findings from the Nigeria study. Futher analysis of cluster ST38 showed strong association of ST38 with plasmid incompatibility types, particularly

with plasmid IncFIB and strongest with IncFIB(pB171)_1_pB171 and were mostly link to diseased children. This is a clear demonstration of virulence clone of ST38. The only ST38 (SID103530) lacking plasmid associated resistance gene was not link to disease. Although there are three ST38 that have IncFIB which were not link to disease during the study but the children with these three ST38 that harbours IncFIB were confirmed to have developed diarrhoea following 60 days followup.

Following reanalysis of the WGS data using galaxy pipeline instead of nullabo pipeline we realised that two EAEC isolates of SID 102191 and 102375 generated allels that are artefact due to poor sequencing so because of this finding the ST2704 and ST394 is cancelled for the two SIDs respectively. However, the SID102602 and SID103076 yielded new allels that require futher confirmation by the custodian of MLST database.

The production of extended-spectrum β-lactamases (ESBL) is one of the common causes of resistance to the oxyimino-cephalosporin such as ceftazidime, cefotaxime and ceftriaxone (Pitout and Laupland 2008). The predominant types of ESBLs in 1980s and early 1990s that belong to TEM and SHV families such as TEM-1, TEM-2, and SHV-1 β-lactamases, which are regarded as derivatives of non-ESBLs (Paterson and Bonomo 2005), are what we found in our investigation. It is interesting to note that all of our EAEC isolates did not produce ESBL CTX-M complex or CTX-M15 or CTX-M-14 which is contrary to many studies that have shown prevalence of ESBL CTX-M-15 producing *E. coli* (Fam, Leflon-Guibout et al. 2011; Aibinu, Odugbemi et al. 2012; Peirano, van der Bij et al. 2012; Imuta, Ooka et al. 2016). The reason for this is unclear, it may be that ESBLs CTX-M clone is not in circulation in the Gambia rural regions although it should be noted that our samples are mainly from children under 5 years old who are probably not expose to ESBL CTX-M producing drugs either through their case contacts or animals. This may explain why the majority of the EAEC

strains are carriage as they lack CTX-M resistant clone. More studies may be required to resolve our speculation.

8.4 Conclusion

This study shows that EAEC strains are diversely distributed among diarrhoeal and non-diarrhoeal children using variables such as phenotypic resistant score and genotypic typing that include sequence type and presence or absence of resistance-gene. Our study further showed that ESBL CTX-M clone associated with bacterial virulence and widespread in other geographic regions in the globe is not in circulation among children in the rural Gambia. Additionally, this is one of the few studies from sub-Sahara Africa that has shown higher rate of concordance of phenotypic resistant score with WGS resistant genes.

Chapter 9: Discussion of Hypotheses, Aims and Objectives

The primary aim of this study was to explore molecular approaches in the diagnosis of diarrhoea caused by infectious EAEC among children from rural Gambia. This has been fulfilled following characterisation and evaluation of the putative virulent factors of EAEC strains from children under five year old in relation to diarrhoeal outcome, measure of bacterial load and screening for production of biofilm factor.

9.1 Putative virulent factors such as *pet*, *sepA* and *aggA* influence disease outcome in The Gambia

The association of EAEC virulent genes such as *pet* and *agg*A and diarrhoeal disease have been shown in some studies (Eslava, Navarro-Garcia et al. 1998; Lima, Boisen et al. 2013; Bafandeh, Haghi et al. 2015; Jensen 2017). However, many studies from different regions of the world have not detected a strong association of *pet* and *agg*A genes among diarrhoeal infants (Samie, Obi et al. 2007; Boisen, Scheutz et al. 2012) as this study indicated. Therefore this hypothesis was accepted in respect of *pet* and *agg*A genes causing diarrhoea among infants. Additionally, the hypothesis was accepted for *sep*A, *ast*A and *cap*U that were associated to diarrhoeal among children less than five years old.

9.2 Measure of Bacterial load assay to ascertain the relevance of TaqMan-qPCR to diagnose EAEC

Globally, limited studies have utilised TaqMan-qPCR that target *aat*A gene to diagnose EAEC (Chattaway, Harris et al. 2013; Lima, Quetz Jda et al. 2013). This was the first study in sub-Sahara Africa to explore TaqMan-qPCR diagnostic tool to investigate infectious EAEC. This diagnostic tool is highly discriminatory, although the result obtained showed that EAEC was not associated with diarrhoea. However, presence of *pet* gene in the EAEC with higher bacterial load showed association with diarrhoeal illness among children.

9.3 Screening for biofilm production among EAEC strains to diagnose infectious EAEC diarrhoea

Many studies have used production of biofilm to detect infectious EAEC strains (Wakimoto, Nishi et al. 2004; Bangar and Mamatha 2007; Dadawala 2010). In this study, three biofilm screening methods were adopted (TT, CRA and TCP) of which TCP method found to be more reliable and specific and was used in the data analysis. Despite TCP reliability, result showed that its specificity and sensitivity will increase if the test is done on EAEC isolates known to have harboured *aat*A and *agg*R genes as well as other genes under the regulatory of *agg*R.

9.4 Investigating the prevalence of multidrug resistant EAEC

The treatment of EAEC and its eradication continue to be a challenge majorly in low income countries despite been sensitive to many antibiotics particularly third generation antibiotics (Nguyen, Le et al. 2005; Aslani, Alikhani et al. 2011). The problem associated with the eradication of EAEC includes increase in resistance to commonly use antibiotic likely due largely to indiscriminate use of antibiotics in human and animal. Several studies have reported EAEC strains to be resistant to multiple antibiotics (Nguyen, Le et al. 2005; Hebbelstrup Jensen, Stensvold et al. 2016) but majority of these studies have not shown multi-resistant strains been associated to a particular age stratum as shown in this study

9.5 Employ WGS to investigate Association of EAEC with diarrhoeal disease

Whole Genome Sequencing was employed to identify EAEC strains that caused diarrhoeal outbreak in developed countries (Scavia, Staffolani et al. 2008). In this study, WGS unveiled the resistant gene associated with diarrhoea in children. Also, the method has emphasised the concordance of phenotypic resistance with genotypic presence of resistance genes. WGS further displayed the divergence of EAEC which make it more difficult to attribute a particular sequence type (ST) and other genetic components to infectious EAEC.

Chapter 10: Concluding remarks

10. 1 Limitations of the study

10.1.1 Inability to perform adherence assay

Adherence assay is regarded as gold standard for diagnosing EAEC, the assay should have been considered in this study in order to classically measure the strength of the three diagnostic approaches used. Although an attempt was made to perform the adherence assay in the collaborator's laboratory in the USA but due to gross contamination of the EAEC isolates during shipment thus it became impossible to perform the assay.

10.1.2 Unable to conduct conventional serotyping on the study strains

I was unable to do serotyping on the EAEC isolates due to cost. If the serotyping was done, the study would have been strengthened by linking distinct serotypes to the cluster of virulence genes and possibly bacterial load and biofilm production. However, I was able to utilise whole genome sequencing to do MLST and obtained sequence type (ST). WGS also revealed resistance genes associated with the EAEC used in this study. Thus, WGS is considered to be more discriminatory compared to the conventional serotyping.

10.1.3 Inability to perform pathogenicity assay

Again, I was unable to perform pathogenicity assay using animal (mouse) model to investigate virulence activities of the identified EAEC virulent genes such as *pet*, *agg*A and

*cap*U due to lack of facility and skill. Although, this aspect is among the investigations planned for future study.

10.1.4 Data

The EAEC isolates used in this study were obtained from among the children living in the rural Gambia. It was recommended to have considered sampling adult particularly, mothers of children or adult case-contacts in order to extend possible source of infection. Animals in the compound where the children reside were recorded but samples were not taken.

10.2 Future studies

10.2.1 Adherence and pathogenicity assay

Given the fact that adherence (HEp-2 cell) assay is the gold standard for the identification of EAEC, it is important to factor it into the future studies relating to diagnosis of diarrhoea even though the technique is unable to differentiate typical-EAEC (presence of *agg*R gene) from atypical-EAEC (absence of *agg*R gene). Also, pathogenicity assay is an important experiment to consider as it can demonstrate causality and clear doubt on the importance of EAEC in diarrhoeal disease.

10.2.2 Conduct bacterial load that target *agg*R, *pet*, *sep*A, *orf*3 and *cap*U

Measure of bacterial load that target true and essential genes listed above must be considered in the future study, so that, EAEC genes that were found to be associated with diarrhoea can be well established.

10.2.3 Perform Whole Genome Sequence on identified EAEC isolates from human and animal to establish transmission route

Whole genome sequencing of EAEC that were concurrently isolated from human and animal will be ideal future investigation to identify source of transmission. This will unmask relatedness of the strains as well as reveal several virulence factors that include resistance genes and biofilm producing genes that are contributory to EAEC infectivity.

10.2.4 Investigate how pet may allow successful competition with commensal members of the gut microbiome

The mechanisms that regulate the ability of microbiota to restrain pathogen growth are complex and include competitive metabolic interactions, localization to intestinal niches, induction of host immune responses. Pathogens such as EAEC having pet virulent factor might, in turn, have evolved strategies to escape from commensal-mediated colonisation resistance. Investigating into these strategies will broaden our understanding of pathogen-commensal interactions which may lead to new therapeutic approaches that are critical for controlling infection and disease

10.2.5 Investigate into how EAEC infection predisposes children to malnutrition Vis versa

Many studies outside West-Africa have incriminated EAEC infection as a cause of malnutrition in children but no such studies have been demonstrated in the region where malnutrition is endemic particularly in the rural Gambia. Therefore, this type of study will be considered as an area of future investigation.

10.3 Closing remarks

Despite all the limitations, this study provides detailed initial description of EAEC virulence factors circulating among children in the rural Gambia and it is directly related to a geographically defined West-African population setting.

Moreover, the study unmasked the magnitude of multi-drug resistant EAEC strains among children and this was never done in the West-Africa before now.

This study has at least answered some relevant questions relating to characteristics of EAC in association to diarrhoea among younger children from Gambia. Although, more investigations are required in order to explain why many children in the rural Gambia are carrier of EAEC and did not fall ill of the infection to result to diarrhoea. However, this study has hinted that EAEC infection makes Gambian children susceptible to malnutrition.

Also, it gladdens me that all of the laboratory investigations and laboratory procedures were performed at MRCG@LSHTM The Gambia, providing basis for future studies into areas requiring exploring more in great detail in the sub-regions in future.

Chapter 11: Reference

Adachi, J. A., C. D. Ericsson, et al. (2002). "Natural history of enteroaggregative and enterotoxigenic Escherichia coli infection among US travelers to Guadalajara, Mexico." J Infect Dis **185**(11): 1681-1683.

Adachi, J. A., Z. D. Jiang, et al. (2001). "Enteroaggregative Escherichia coli as a major etiologic agent in traveler's diarrhea in 3 regions of the world." Clin Infect Dis **32**(12): 1706-1709.

Adamus-Bialek, W., A. Kubiak, et al. (2015). "Analysis of uropathogenic Escherichia coli biofilm formation under different growth conditions." Acta Biochim Pol **62**(4): 765-771.

Aibinu, I., T. Odugbemi, et al. (2012). "Sequence type ST131 and ST10 complex (ST617) predominant among CTX-M-15-producing Escherichia coli isolates from Nigeria." Clin Microbiol Infect **18**(3): E49-51.

Akter, S., M. Islam, et al. (2013). "Prevalence and distribution of different diarrhoeagenic Escherichia coli virulotypes in major water bodies in Bangladesh." Epidemiol Infect **141**(12): 2516-2525.

Al-Hasani, K., F. Navarro-Garcia, et al. (2009). "The immunogenic SigA enterotoxin of Shigella flexneri 2a binds to HEp-2 cells and induces fodrin redistribution in intoxicated epithelial cells." PLoS One **4**(12): e8223.

Ali, M. M., S. F. Ahmed, et al. (2014). "Enteroaggregative Escherichia coli in diarrheic children in Egypt: molecular characterization and antimicrobial susceptibility." J Infect Dev Ctries **8**(5): 589-596.

Andrade, J. A., E. Freymuller, et al. (2011). "Adherence of enteroaggregative Escherichia coli to the ileal and colonic mucosa: an in vitro study utilizing the scanning electron microscopy." Arq Gastroenterol **48**(3): 199-204.

AppliedBiosystem, A. B. (2003). "Creating Standard Curves with Genomic DNA or Plasmid DNA Templates for Use in Quantitative PCR." www.appliedbiosystems.com 1-9.

Araujo, J. M., G. F. Tabarelli, et al. (2007). "Typical enteroaggregative and atypical enteropathogenic types of Escherichia coli are the most prevalent diarrhea-associated pathotypes among Brazilian children." J Clin Microbiol **45**(10): 3396-3399.

Arenas-Hernandez, M. M., Y. Martinez-Laguna, et al. (2012). "Clinical implications of enteroadherent Escherichia coli." Curr Gastroenterol Rep **14**(5): 386-394.

Aschtgen, M. S., C. S. Bernard, et al. (2008). "SciN is an outer membrane lipoprotein required for type VI secretion in enteroaggregative Escherichia coli." J Bacteriol **190**(22): 7523-7531.

Aslani, M. M., M. Y. Alikhani, et al. (2011). "Characterization of enteroaggregative Escherichia coli (EAEC) clinical isolates and their antibiotic resistance pattern." Int J Infect Dis **15**(2): e136-139.

Bafandeh, S., F. Haghi, et al. (2015). "Prevalence and virulence characteristics of enteroaggregative Escherichia coli in a case-control study among patients from Iran." J Med Microbiol **64**(Pt 5): 519-524.

Baker, K. K., Debasish S, Adeyemi M, Ikumapayi UN, Antonio M., Diliruba N, Blackwelder W, Wu Y, Panchalingam S,Biswas K, Kotloff K, Nataro JP and Myron M. (2011). "Transmission of diarrhoeal Pathogens trough Drinking Water in Rural Gambia." Proceedings of the Water Environment Federation, Disinfection **11**: 53-63.

Baker, K. K., C. E. O'Reilly, et al. (2016). "Sanitation and Hygiene-Specific Risk Factors for Moderate-to-Severe Diarrhea in Young Children in the Global Enteric Multicenter Study, 2007-2011: Case-Control Study." PLoS Med **13**(5): e1002010.

Bangar, R. and B. Mamatha (2007). "Biochemical Identification of Enteroaggregative Escherichia Coli among Infants with Acute Diarrhea from Manipal, India." Int J Health Sci (Qassim) **1**(2): 237-241.

Bangar, R. and B. Mamatha (2008). "Identification of enteroaggregative Escherichia coli in infants with acute diarrhea based on biofilm production in Manipal, south India." Indian J Med Sci **62**(1): 8-12.
Bankevich, A., S. Nurk, et al. (2012). "SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing." J Comput Biol **19**(5): 455-477.
Barrell, R. A. and M. G. Rowland (1979). "Infant foods as a potential source of diarrhoeal illness in rural West Africa." Trans R Soc Trop Med Hyg **73**(1): 85-90.
Bartlett, J. G., D. N. Gilbert, et al. (2013). "Seven ways to preserve the miracle of antibiotics." Clin Infect Dis **56**(10): 1445-1450.
Baudry, B., S. J. Savarino, et al. (1990). "A sensitive and specific DNA probe to identify enteroaggregative Escherichia coli, a recently discovered diarrheal pathogen." J Infect Dis **161**(6): 1249-1251.
Bellini, E. M., W. P. Elias, et al. (2005). "Antibody response against plasmid-encoded toxin (Pet) and the protein involved in intestinal colonization (Pic) in children with diarrhea produced by enteroaggregative Escherichia coli." FEMS Immunol Med Microbiol **43**(2): 259-264.
Benjelloun-Touimi, Z., P. J. Sansonetti, et al. (1995). "SepA, the major extracellular protein of Shigella flexneri: autonomous secretion and involvement in tissue invasion." Mol Microbiol **17**(1): 123-135.
Bernier, C., P. Gounon, et al. (2002). "Identification of an aggregative adhesion fimbria (AAF) type III-encoding operon in enteroaggregative Escherichia coli as a sensitive probe for detecting the AAF-encoding operon family." Infect Immun **70**(8): 4302-4311.
Bhan, M. K., V. Khoshoo, et al. (1989). "Enteroaggregative Escherichia coli and Salmonella associated with nondysenteric persistent diarrhea." Pediatr Infect Dis J **8**(8): 499-502.
Bhargava, S., B. B. Johnson, et al. (2009). "Heat-resistant agglutinin 1 is an accessory enteroaggregative Escherichia coli colonization factor." J Bacteriol **191**(15): 4934-4942.
Bhatnagar, S., M. K. Bhan, et al. (1993). "Enteroaggregative Escherichia coli may be a new pathogen causing acute and persistent diarrhea." Scand J Infect Dis **25**(5): 579-583.
Bibbal, D., M. Kerouredan, et al. (2014). "Slaughterhouse effluent discharges into rivers not responsible for environmental occurrence of enteroaggregative Escherichia coli." Vet Microbiol **168**(2-4): 451-454.
Bielaszewska, M., E. A. Idelevich, et al. (2012). "Effects of antibiotics on Shiga toxin 2 production and bacteriophage induction by epidemic Escherichia coli O104:H4 strain." Antimicrob Agents Chemother **56**(6): 3277-3282.
Bielaszewska, M., A. Mellmann, et al. (2011). "Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study." Lancet Infect Dis **11**(9): 671-676.
Black, R. E., L. H. Allen, et al. (2008). "Maternal and child undernutrition: global and regional exposures and health consequences." Lancet **371**(9608): 243-260.
Black, R. E., S. S. Morris, et al. (2003). "Where and why are 10 million children dying every year?" Lancet **361**(9376): 2226-2234.
Boisen, N., A. R. Melton-Celsa, et al. (2015). "Shiga toxin 2a and Enteroaggregative Escherichia coli--a deadly combination." Gut Microbes **6**(4): 272-278.
Boisen, N., F. Ruiz-Perez, et al. (2009). "Short report: high prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative Escherichia coli." Am J Trop Med Hyg **80**(2): 294-301.
Boisen, N., F. Scheutz, et al. (2012). "Genomic characterization of enteroaggregative Escherichia coli from children in Mali." J Infect Dis **205**(3): 431-444.
Boisen, N., C. Struve, et al. (2008). "New adhesin of enteroaggregative Escherichia coli related to the Afa/Dr/AAF family." Infect Immun **76**(7): 3281-3292.
Bolger, A. M., M. Lohse, et al. (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data." Bioinformatics **30**(15): 2114-2120.

Boll, E. J., C. Struve, et al. (2013). "Role of enteroaggregative Escherichia coli virulence factors in uropathogenesis." Infect Immun **81**(4): 1164-1171.
Boll, E. J., C. Struve, et al. (2012). "Enteroaggregative Escherichia coli promotes transepithelial migration of neutrophils through a conserved 12-lipoxygenase pathway." Cell Microbiol **14**(1): 120-132.
Boudailliez, B., P. Berquin, et al. (1997). "Possible person-to-person transmission of Escherichia coli O111--associated hemolytic uremic syndrome." Pediatr Nephrol **11**(1): 36-39.
Brzuszkiewicz, E., A. Thurmer, et al. (2011). "Genome sequence analyses of two isolates from the recent Escherichia coli outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic Escherichia coli (EAHEC)." Arch Microbiol **193**(12): 883-891.
Buchholz, U., H. Bernard, et al. (2011). "German outbreak of Escherichia coli O104:H4 associated with sprouts." N Engl J Med **365**(19): 1763-1770.
Buermans, H. P. and J. T. den Dunnen (2014). "Next generation sequencing technology: Advances and applications." Biochim Biophys Acta **1842**(10): 1932-1941.
Bush, K. (2013). "Proliferation and significance of clinically relevant beta-lactamases." Ann N Y Acad Sci **1277**: 84-90.
Busscher, H. J., R. Bos, et al. (1995). "Initial microbial adhesion is a determinant for the strength of biofilm adhesion." FEMS Microbiol Lett **128**(3): 229-234.
Cassar, C. A., M. Ottaway, et al. (2004). "Absence of enteroaggregative Escherichia coli in farmed animals in Great Britain." Vet Rec **154**(8): 237-239.
CDC (2013). "Antibiotic resistance threats in the United States ".
Cennimo, D., A. Abbas, et al. (2009). "The prevalence and virulence characteristics of enteroaggregative Escherichia coli at an urgent-care clinic in the USA: a case-control study." J Med Microbiol **58**(Pt 4): 403-407.
Cerna, J. F., J. P. Nataro, et al. (2003). "Multiplex PCR for detection of three plasmid-borne genes of enteroaggregative Escherichia coli strains." J Clin Microbiol **41**(5): 2138-2140.
Chan, K. N., A. D. Phillips, et al. (1994). "Enteroaggregative Escherichia coli: another cause of acute and chronic diarrhoea in England?" J Pediatr Gastroenterol Nutr **18**(1): 87-91.
Chattaway, M. A., R. Harris, et al. (2013). "Investigating the link between the presence of enteroaggregative Escherichia coli and infectious intestinal disease in the United Kingdom, 1993 to 1996 and 2008 to 2009." Euro Surveill **18**(37).
Chaudhuri, R. R., M. Sebaihia, et al. (2010). "Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative Escherichia coli strain 042." PLoS One **5**(1): e8801.
Cho, S. H., Y. S. Lim, et al. (2011). "Prevalence of Antibiotic Resistance in Escherichia coli Fecal Isolates From Healthy Persons and Patients With Diarrhea." Osong Public Health Res Perspect **2**(1): 41-45.
Christensen, G. D., W. A. Simpson, et al. (1985). "Adherence of coagulase-negative staphylococci to plastic tissue culture plates: a quantitative model for the adherence of staphylococci to medical devices." J Clin Microbiol **22**(6): 996-1006.
Chung, H., S. J. Pamp, et al. (2012). "Gut immune maturation depends on colonization with a host-specific microbiota." Cell **149**(7): 1578-1593.
Clarke, S. C. (2001). "Diarrhoeagenic Escherichia coli--an emerging problem?" Diagn Microbiol Infect Dis **41**(3): 93-98.
CLSI (2012). "Clinical and Laboratory Standards Institute guidelines." CLSI version M100-S22 **32**(3): 1-50.
Cobeljic, M., B. Miljkovic-Selimovic, et al. (1996). "Enteroaggregative Escherichia coli associated with an outbreak of diarrhoea in a neonatal nursery ward." Epidemiol Infect **117**(1): 11-16.
Cohen, M. B., J. P. Nataro, et al. (2005). "Prevalence of diarrheagenic Escherichia coli in acute childhood enteritis: a prospective controlled study." J Pediatr **146**(1): 54-61.

Cravioto, A., A. Tello, et al. (1991). "Association of Escherichia coli HEp-2 adherence patterns with type and duration of diarrhoea." Lancet **337**(8736): 262-264.
Cravioto, J. and R. Arrieta (1979). "Stimulation and mental development of malnourished infants." Lancet **2**(8148): 899.
Croxen, M. A. and B. B. Finlay (2010). "Molecular mechanisms of Escherichia coli pathogenicity." Nat Rev Microbiol **8**(1): 26-38.
Croxen, M. A., R. J. Law, et al. (2013). "Recent advances in understanding enteric pathogenic Escherichia coli." Clin Microbiol Rev **26**(4): 822-880.
CRSR (2005). "Life expectancy in the United States. ." http://www.cnie.org/nle/crsreports/05mar/RL32792.pdf. **March 2005**.
Cuting, W. A. M. (1998). "Diarrhoea pathophysiology: Mechanisms of diarrhoea and why they matter.". Dialogue on Diarrhoea (35): : 4-5.
Czeczulin, J. R., S. Balepur, et al. (1997). "Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative Escherichia coli." Infect Immun **65**(10): 4135-4145.
Czeczulin, J. R., T. S. Whittam, et al. (1999). "Phylogenetic analysis of enteroaggregative and diffusely adherent Escherichia coli." Infect Immun **67**(6): 2692-2699.
Dadawala, A. I. (2010). "Assessment of Escherichia coli isolates for In vitro biofilm production." Veterinary world **3**(8): 364-366.
Dallman, T., G. P. Smith, et al. (2012). "Characterization of a verocytotoxin-producing enteroaggregative Escherichia coli serogroup O111:H21 strain associated with a household outbreak in Northern Ireland." J Clin Microbiol **50**(12): 4116-4119.
Danese, P. N., L. A. Pratt, et al. (2000). "Exopolysaccharide production is required for development of Escherichia coli K-12 biofilm architecture." J Bacteriol **182**(12): 3593-3596.
Davies, D. G., A. M. Chakrabarty, et al. (1993). "Exopolysaccharide production in biofilms: substratum activation of alginate gene expression by Pseudomonas aeruginosa." Appl Environ Microbiol **59**(4): 1181-1186.
Davoodabadi, A., M. Abbaszadeh, et al. (2015). "Phenotypic and Genotypic Characterization of Enteroaggregative Escherichia coli Strains Isolated From Diarrheic Children in Iran." Jundishapur J Microbiol **8**(9): e22295.
de la Cabada Bauche, J. and H. L. Dupont (2011). "New Developments in Traveler's Diarrhea." Gastroenterol Hepatol (N Y) **7**(2): 88-95.
Denno, D. M., N. Shaikh, et al. (2012). "Diarrhea etiology in a pediatric emergency department: a case control study." Clin Infect Dis **55**(7): 897-904.
Deurenberg, R. H., E. Bathoorn, et al. (2017). "Application of next generation sequencing in clinical microbiology and infection prevention." J Biotechnol **243**: 16-24.
Dobrindt, U. (2005). "(Patho-)Genomics of Escherichia coli." Int J Med Microbiol **295**(6-7): 357-371.
Dudley, E. G., C. Abe, et al. (2006). "An IncI1 plasmid contributes to the adherence of the atypical enteroaggregative Escherichia coli strain C1096 to cultured cells and abiotic surfaces." Infect Immun **74**(4): 2102-2114.
Dudley, E. G., N. R. Thomson, et al. (2006). "Proteomic and microarray characterization of the AggR regulon identifies a pheU pathogenicity island in enteroaggregative Escherichia coli." Mol Microbiol **61**(5): 1267-1282.
DuPont, H. L. (2007). "Therapy for and prevention of traveler's diarrhea." Clin Infect Dis **45 Suppl 1**: S78-84.
Durack, J. and S. V. Lynch (2019). "The gut microbiome: Relationships with disease and opportunities for therapy." J Exp Med **216**(1): 20-40.
Dutta, P. R., R. Cappello, et al. (2002). "Functional comparison of serine protease autotransporters of enterobacteriaceae." Infect Immun **70**(12): 7105-7113.
Dutta, S., S. Guin, et al. (2013). "Trends in the prevalence of diarrheagenic Escherichia coli among hospitalized diarrheal patients in Kolkata, India." PLoS One **8**(2): e56068.

Dutta, S., S. Pal, et al. (1999). "Use of PCR to identify enteroaggregative Escherichia coli as an important cause of acute diarrhoea among children living in Calcutta, India." J Med Microbiol **48**(11): 1011-1016.
Eckmann, L. and M. F. Kagnoff (2005). "Intestinal mucosal responses to microbial infection." Springer Semin Immunopathol **27**(2): 181-196.
Eddy, S. F. and K. B. Storey (2007). "p38 MAPK regulation of transcription factor targets in muscle and heart of the hibernating bat, Myotis lucifugus." Cell Biochem Funct **25**(6): 759-765.
EFSA (2015). "Emerging food borne pathogen." Eropean Food Safety Autority - www.efsa.europa.eu/efsajournal.
Ellington, M. J., O. Ekelund, et al. (2017). "The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee." Clin Microbiol Infect **23**(1): 2-22.
Ericsson, C. D., H. L. DuPont, et al. (2007). "Loperamide plus azithromycin more effectively treats travelers' diarrhea in Mexico than azithromycin alone." J Travel Med **14**(5): 312-319.
Eslava, C., F. Navarro-Garcia, et al. (1998). "Pet, an autotransporter enterotoxin from enteroaggregative Escherichia coli." Infect Immun **66**(7): 3155-3163.
Eslava, C., Villaseca, J., Morales, R., Navarro, A., and Cravioto, A. (1993). Identification of a protein with toxigenic activity produced by enteroaggregative Escherichia coli. abstr B-105 of the 93rd General Meeting of the American Society for Microbiology
Estrada-Garcia, T., J. F. Cerna, et al. (2005). "Drug-resistant diarrheogenic Escherichia coli, Mexico." Emerg Infect Dis **11**(8): 1306-1308.
Estrada-Garcia, T., C. Lopez-Saucedo, et al. (2009). "Association of diarrheagenic Escherichia coli Pathotypes with infection and diarrhea among Mexican children and association of atypical Enteropathogenic E. coli with acute diarrhea." J Clin Microbiol **47**(1): 93-98.
Estrada-Garcia, T. and F. Navarro-Garcia (2012). "Enteroaggregative Escherichia coli pathotype: a genetically heterogeneous emerging foodborne enteropathogen." FEMS Immunol Med Microbiol **66**(3): 281-298.
Estrada-Garcia, T., I. Perez-Martinez, et al. (2014). "Enteroaggregative coli: A Pathogen Bridging the North and South." Curr Trop Med Rep **1**(2): 88-96.
Ewing, W. H. D., B.R. and Montague, T.S. (1963). "Studies on the Occurence of Escherichia coli Serotypes Associated with Diarrhoeal Disease " UD Department of Health, Education and Welfare, Public Health Services, CDC 1-38.
Fam, N., V. Leflon-Guibout, et al. (2011). "CTX-M-15-producing Escherichia coli clinical isolates in Cairo (Egypt), including isolates of clonal complex ST10 and clones ST131, ST73, and ST405 in both community and hospital settings." Microb Drug Resist **17**(1): 67-73.
Farag, T. H., D. Nasrin, et al. (2012). "Some epidemiologic, clinical, microbiologic, and organizational assumptions that influenced the design and performance of the Global Enteric Multicenter Study (GEMS)." Clin Infect Dis **55 Suppl 4**: S225-231.
Farfan, M. J., K. G. Inman, et al. (2008). "The major pilin subunit of the AAF/II fimbriae from enteroaggregative Escherichia coli mediates binding to extracellular matrix proteins." Infect Immun **76**(10): 4378-4384.
Fasano, A., F. R. Noriega, et al. (1995). "Shigella enterotoxin 1: an enterotoxin of Shigella flexneri 2a active in rabbit small intestine in vivo and in vitro." J Clin Invest **95**(6): 2853-2861.
Feachem, R. G. (1984). "Interventions for the control of diarrhoeal diseases among young children: promotion of personal and domestic hygiene." Bull World Health Organ **62**(3): 467-476.
Flores, J. and P. C. Okhuysen (2009). "Enteroaggregative Escherichia coli infection." Curr Opin Gastroenterol **25**(1): 8-11.
Frank, C., D. Werber, et al. (2011). "Epidemic profile of Shiga-toxin-producing Escherichia coli O104:H4 outbreak in Germany." N Engl J Med **365**(19): 1771-1780.

Frech, S. A., H. L. Dupont, et al. (2008). "Use of a patch containing heat-labile toxin from Escherichia coli against travellers' diarrhoea: a phase II, randomised, double-blind, placebo-controlled field trial." Lancet **371**(9629): 2019-2025.

Freeman, D. J., F. R. Falkiner, et al. (1989). "New method for detecting slime production by coagulase negative staphylococci." J Clin Pathol **42**(8): 872-874.

Fujiyama, R., J. Nishi, et al. (2008). "The shf gene of a Shigella flexneri homologue on the virulent plasmid pAA2 of enteroaggregative Escherichia coli 042 is required for firm biofilm formation." Curr Microbiol **56**(5): 474-480.

Gambia (2016). "Economy of The Gambia."

Gambia, T. (2015). "The Gambia Statistics." The Gambia Central Statistics Department.

Garnett, J. A. and S. Matthews (2012). "Interactions in bacterial biofilm development: a structural perspective." Curr Protein Pept Sci **13**(8): 739-755.

Glandt, M., J. A. Adachi, et al. (1999). "Enteroaggregative Escherichia coli as a cause of traveler's diarrhea: clinical response to ciprofloxacin." Clin Infect Dis **29**(2): 335-338.

Goldberg, J. (2002). "Biofilms and antibiotic resistance: a genetic linkage." Trends Microbiol **10**(264).

Goldstone, R. J. and D. G. E. Smith (2017). "A population genomics approach to exploiting the accessory 'resistome' of Escherichia coli." Microb Genom **3**(4): e000108.

Gonzalez, R., C. Diaz, et al. (1997). "Age-specific prevalence of Escherichia coli with localized and aggregative adherence in Venezuelan infants with acute diarrhea." J Clin Microbiol **35**(5): 1103-1107.

Gouyer, V., F. Gottrand, et al. (2011). "The extraordinarily complex but highly structured organization of intestinal mucus-gel unveiled in multicolor images." PLoS One **6**(4): e18761.

Goyal, A., M. Konar, et al. (2010). "Galactose specific adhesin of enteroaggregative E. coli induces IL-8 secretion via activation of MAPK and STAT-3 in INT-407 cells." Biochim Biophys Acta **1800**(6): 574-579.

Grad, Y. H., M. Lipsitch, et al. (2012). "Genomic epidemiology of the Escherichia coli O104:H4 outbreaks in Europe, 2011." Proc Natl Acad Sci U S A **109**(8): 3065-3070.

Grada, A. and K. Weinbrecht (2013). "Next-generation sequencing: methodology and application." J Invest Dermatol **133**(8): e11.

Greenberg, D. E., Z. D. Jiang, et al. (2002). "Markers of inflammation in bacterial diarrhea among travelers, with a focus on enteroaggregative Escherichia coli pathogenicity." J Infect Dis **185**(7): 944-949.

Guerrant, R. L., R. B. Oria, et al. (2008). "Malnutrition as an enteric infectious disease with long-term effects on child development." Nutr Rev **66**(9): 487-505.

Guerrant, R. L., T. Van Gilder, et al. (2001). "Practice guidelines for the management of infectious diarrhea." Clin Infect Dis **32**(3): 331-351.

Guignot, J., C. Chaplais, et al. (2007). "The secreted autotransporter toxin, Sat, functions as a virulence factor in Afa/Dr diffusely adhering Escherichia coli by promoting lesions in tight junction of polarized epithelial cells." Cell Microbiol **9**(1): 204-221.

Guion, C. E., T. J. Ochoa, et al. (2008). "Detection of diarrheagenic Escherichia coli by use of melting-curve analysis and real-time multiplex PCR." J Clin Microbiol **46**(5): 1752-1757.

Gutierrez-Jimenez, J., I. Arciniega, et al. (2008). "The serine protease motif of Pic mediates a dose-dependent mucolytic activity after binding to sugar constituents of the mucin substrate." Microb Pathog **45**(2): 115-123.

Guyer, D. M., S. Radulovic, et al. (2002). "Sat, the secreted autotransporter toxin of uropathogenic Escherichia coli, is a vacuolating cytotoxin for bladder and kidney epithelial cells." Infect Immun **70**(8): 4539-4546.

Harada, T., M. Hiroi, et al. (2007). "A food poisoning diarrhea outbreak caused by enteroaggregative Escherichia coli serogroup O126:H27 in Shizuoka, Japan." Jpn J Infect Dis **60**(2-3): 154-155.

Harrington, S. M., E. G. Dudley, et al. (2006). "Pathogenesis of enteroaggregative Escherichia coli infection." FEMS Microbiol Lett **254**(1): 12-18.

Harrington, S. M., J. Sheikh, et al. (2009). "The Pic protease of enteroaggregative Escherichia coli promotes intestinal colonization and growth in the presence of mucin." Infect Immun **77**(6): 2465-2473.
Harrington, S. M., M. C. Strauman, et al. (2005). "Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative Escherichia coli." Cell Microbiol **7**(11): 1565-1578.
Harris, V. C., B. W. Haak, et al. (2017). "The Intestinal Microbiome in Infectious Diseases: The Clinical Relevance of a Rapidly Emerging Field." Open Forum Infect Dis **4**(3): ofx144.
Hassan, A., J. Usman, et al. (2011). "Evaluation of different detection methods of biofilm formation in the clinical isolates." Braz J Infect Dis **15**(4): 305-311.
Hauswaldt, S., M. Nitschke, et al. (2013). "Lessons Learned From Outbreaks of Shiga Toxin Producing Escherichia coli." Curr Infect Dis Rep **15**(1): 4-9.
Hebbelstrup Jensen, B., K. E. Olsen, et al. (2014). "Epidemiology and clinical manifestations of enteroaggregative Escherichia coli." Clin Microbiol Rev **27**(3): 614-630.
Hebbelstrup Jensen, B., D. Roser, et al. (2016). "Childhood diarrhoea in Danish day care centres could be associated with infant colic, low birthweight and antibiotics." Acta Paediatr **105**(1): 90-95.
Hebbelstrup Jensen, B., C. R. Stensvold, et al. (2016). "Enteroaggregative Escherichia coli in Daycare-A 1-Year Dynamic Cohort Study." Front Cell Infect Microbiol **6**: 75.
Heilmann, C., O. Schweitzer, et al. (1996). "Molecular basis of intercellular adhesion in the biofilm-forming Staphylococcus epidermidis." Mol Microbiol **20**(5): 1083-1091.
Henderson, I. R., J. Czeczulin, et al. (1999). "Characterization of pic, a secreted protease of Shigella flexneri and enteroaggregative Escherichia coli." Infect Immun **67**(11): 5587-5596.
Herzog, K., J. Engeler Dusel, et al. (2014). "Diarrheagenic enteroaggregative Escherichia coli causing urinary tract infection and bacteremia leading to sepsis." Infection **42**(2): 441-444.
Hicks, S., D. C. Candy, et al. (1996). "Adhesion of enteroaggregative Escherichia coli to pediatric intestinal mucosa in vitro." Infect Immun **64**(11): 4751-4760.
Huang, D. B., H. Koo, et al. (2004). "Enteroaggregative Escherichia coli: An Emerging Pathogen." Curr Infect Dis Rep **6**(2): 83-86.
Huang, D. B., J. A. Mohamed, et al. (2007). "Virulence characteristics and the molecular epidemiology of enteroaggregative Escherichia coli isolates from travellers to developing countries." J Med Microbiol **56**(Pt 10): 1386-1392.
Huppertz, H. I., S. Rutkowski, et al. (1997). "Acute and chronic diarrhoea and abdominal colic associated with enteroaggregative Escherichia coli in young children living in western Europe." Lancet **349**(9066): 1660-1662.
Ikumapayi (2014). "Trend in the Epidemiology of Diarrhoegenic Escherichia coli among under 5 year old children in te Gambia." Unpublished.
Ikumapayi, U. N. (2016). "A quantitative assessment of clinical significance of EAEC in diarrhoea using TaqMan-QPCR." Unpublished.
Ikumapayi, U. N., M. Antonio, et al. (2007). "Molecular epidemiology of community-acquired invasive non-typhoidal Salmonella among children aged 2 29 months in rural Gambia and discovery of a new serovar, Salmonella enterica Dingiri." J Med Microbiol **56**(Pt 11): 1479-1484.
Ikumapayi, U. N., N. Boisen, et al. (2017). "Identification of Subsets of Enteroaggregative Escherichia coli Associated with Diarrheal Disease among Under 5 Years of Age Children from Rural Gambia." Am J Trop Med Hyg **97**(4): 997-1004.
Imuta, N., J. Nishi, et al. (2008). "The Escherichia coli efflux pump TolC promotes aggregation of enteroaggregative E. coli 042." Infect Immun **76**(3): 1247-1256.
Imuta, N., T. Ooka, et al. (2016). "Phylogenetic Analysis of Enteroaggregative Escherichia coli (EAEC) Isolates from Japan Reveals Emergence of CTX-M-14-Producing EAEC O25:H4 Clones Related to Sequence Type 131." J Clin Microbiol **54**(8): 2128-2134.

Infante, R. M., C. D. Ericsson, et al. (2004). "Enteroaggregative Escherichia coli diarrhea in travelers: response to rifaximin therapy." Clin Gastroenterol Hepatol **2**(2): 135-138.

Itoh, Y., I. Nagano, et al. (1997). "Laboratory investigation of enteroaggregative Escherichia coli O untypeable:H10 associated with a massive outbreak of gastrointestinal illness." J Clin Microbiol **35**(10): 2546-2550.

Iwanaga, M., T. Song, et al. (2002). "Enteroaggregative Escherichia coli: incidence in Japan and usefulness of the clump-formation test." J Infect Chemother **8**(4): 345-348.

Iyoda, S., K. Tamura, et al. (2000). "Inducible stx2 phages are lysogenized in the enteroaggregative and other phenotypic Escherichia coli O86:HNM isolated from patients." FEMS Microbiol Lett **191**(1): 7-10.

Izquierdo, M., F. Navarro-Garcia, et al. (2014). "Identification of cell surface-exposed proteins involved in the fimbria-mediated adherence of enteroaggregative Escherichia coli to intestinal cells." Infect Immun **82**(4): 1719-1724.

Januszkiewicz, A., J. Szych, et al. (2012). "Molecular epidemiology of a household outbreak of Shiga-toxin-producing Escherichia coli in Poland due to secondary transmission of STEC O104:H4 from Germany." J Med Microbiol **61**(Pt 4): 552-558.

Jenkins, C., H. Chart, et al. (2006). "Genotyping of enteroaggregative Escherichia coli and identification of target genes for the detection of both typical and atypical strains." Diagn Microbiol Infect Dis **55**(1): 13-19.

Jenkins, C., M. Tembo, et al. (2006). "Detection of enteroaggregative Escherichia coli in faecal samples from patients in the community with diarrhoea." J Med Microbiol **55**(Pt 11): 1493-1497.

Jensen, B. e. a. (2017). "Genetic Virulence Profile of Enteroaggregative Escherichia coli Strains Isolated from Danish Children with Either Acute or Persistent Diarrhoea " Front Cell Infect Microbiol **7**(230).

Jiang, Z. D., D. Greenberg, et al. (2002). "Rate of occurrence and pathogenic effect of enteroaggregative Escherichia coli virulence factors in international travelers." J Clin Microbiol **40**(11): 4185-4190.

Jiang, Z. D., P. C. Okhuysen, et al. (2003). "Genetic susceptibility to enteroaggregative Escherichia coli diarrhea: polymorphism in the interleukin-8 promotor region." J Infect Dis **188**(4): 506-511.

Jonsson, R., C. Struve, et al. (2015). "Novel aggregative adherence fimbria variant of enteroaggregative Escherichia coli." Infect Immun **83**(4): 1396-1405.

Kagambega, A., O. Martikainen, et al. (2012). "Prevalence of diarrheagenic Escherichia coli virulence genes in the feces of slaughtered cattle, chickens, and pigs in Burkina Faso." Microbiologyopen **1**(3): 276-284.

Kahali, S., B. Sarkar, et al. (2004). "Virulence characteristics and molecular epidemiology of enteroaggregative Escherichia coli isolates from hospitalized diarrheal patients in Kolkata, India." J Clin Microbiol **42**(9): 4111-4120.

Kang, G., B. S. Ramakrishna, et al. (2001). "Epidemiological and laboratory investigations of outbreaks of diarrhoea in rural South India: implications for control of disease." Epidemiol Infect **127**(1): 107-112.

Kaper, J. B., J. P. Nataro, et al. (2004). "Pathogenic Escherichia coli." Nat Rev Microbiol **2**(2): 123-140.

Karch, H., E. Denamur, et al. (2012). "The enemy within us: lessons from the 2011 European Escherichia coli O104:H4 outbreak." EMBO Mol Med **4**(9): 841-848.

Kaur, P., A. Chakraborti, et al. (2010). "Enteroaggregative Escherichia coli: An Emerging Enteric Food Borne Pathogen." Interdiscip Perspect Infect Dis **2010**: 254159.

Khan, K., M. Konar, et al. (2010). "Enteroaggregative Escherichia coli infection induces IL-8 production via activation of mitogen-activated protein kinases and the transcription factors NF-kappaB and AP-1 in INT-407 cells." Mol Cell Biochem **337**(1-2): 17-24.

Khoshvaght, H., F. Haghi, et al. (2014). "Extended spectrum betalactamase producing Enteroaggregative Escherichia coli from young children in Iran." Gastroenterol Hepatol Bed Bench **7**(2): 131-136.
Kingsmore, S. F. and C. J. Saunders (2011). "Deep sequencing of patient genomes for disease diagnosis: when will it become routine?" Sci Transl Med **3**(87): 87ps23.
Knutton, S., R. Shaw, et al. (2001). "Phenotypic and genetic analysis of diarrhea-associated Escherichia coli isolated from children in the United Kingdom." J Pediatr Gastroenterol Nutr **33**(1): 32-40.
Koo, H. L., Z. D. Jiang, et al. (2008). "Coliform contamination of vegetables obtained from popular restaurants in Guadalajara, Mexico, and Houston, Texas." Clin Infect Dis **47**(2): 218-221.
Kotloff, K. L., J. P. Nataro, et al. (2013). "Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study." Lancet **382**(9888): 209-222.
Kubomura, A., T. Misaki, et al. (2017). "Phenotypic and Molecular Characterization of Enteroaggregative Escherichia coli Isolated in Kawasaki, Japan." Jpn J Infect Dis **70**(5): 507-512.
Kucharzik, T., J. T. Hudson, 3rd, et al. (2005). "Acute induction of human IL-8 production by intestinal epithelium triggers neutrophil infiltration without mucosal injury." Gut **54**(11): 1565-1572.
Kumar, S., G. Stecher, et al. (2016). "MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets." Mol Biol Evol **33**(7): 1870-1874.
Laing, C. R., Y. Zhang, et al. (2012). "A comparison of Shiga-toxin 2 bacteriophage from classical enterohemorrhagic Escherichia coli serotypes and the German E. coli O104:H4 outbreak strain." PLoS One **7**(5): e37362.
Lefterova, M. I., C. J. Suarez, et al. (2015). "Next-Generation Sequencing for Infectious Disease Diagnosis and Management: A Report of the Association for Molecular Pathology." J Mol Diagn **17**(6): 623-634.
Letunic, I. and P. Bork (2016). "Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees." Nucleic Acids Res **44**(W1): W242-245.
Lewis, K. (2001). "Riddle of biofilm resistance." Antimicrob Agents Chemother **45**(4): 999-1007.
Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows-Wheeler transform." Bioinformatics **25**(14): 1754-1760.
Lima, A. A., G. Fang, et al. (1992). "Persistent diarrhea in northeast Brazil: etiologies and interactions with malnutrition." Acta Paediatr Suppl **381**: 39-44.
Lima, A. A. and R. L. Guerrant (1992). "Persistent diarrhea in children: epidemiology, risk factors, pathophysiology, nutritional impact, and management." Epidemiol Rev **14**: 222-242.
Lima, I. F., N. Boisen, et al. (2013). "Prevalence of enteroaggregative Escherichia coli and its virulence-related genes in a case-control study among children from north-eastern Brazil." J Med Microbiol **62**(Pt 5): 683-693.
Lima, I. F., S. Quetz Jda, et al. (2013). "Enteroaggregative Escherichia coli quantification in children stool samples using quantitative PCR." APMIS **121**(7): 643-651.
Lingwood, C. A., B. Binnington, et al. (2010). "Globotriaosyl ceramide receptor function - where membrane structure and pathology intersect." FEBS Lett **584**(9): 1879-1886.
Linnarsson, S. (2010). "Recent advances in DNA sequencing methods - general principles of sample preparation." Exp Cell Res **316**(8): 1339-1343.
Littman, D. R. and E. G. Pamer (2011). "Role of the commensal microbiota in normal and pathogenic host immune responses." Cell Host Microbe **10**(4): 311-323.
Liu, J., F. Kabir, et al. (2014). "Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: a multicentre study." Lancet Infect Dis **14**(8): 716-724.
Maps (2015). "Map of The Gambia." https://www.ezilon.com/maps/africa/gambia-maps.html.

Mathewson, J. J., P. C. Johnson, et al. (1985). "A newly recognized cause of travelers' diarrhea: enteroadherent Escherichia coli." J Infect Dis **151**(3): 471-475.
Mathewson, J. J., R. A. Oberhelman, et al. (1987). "Enteroadherent Escherichia coli as a cause of diarrhea among children in Mexico." J Clin Microbiol **25**(10): 1917-1919.
Menard, L. P. and J. D. Dubreuil (2002). "Enteroaggregative Escherichia coli heat-stable enterotoxin 1 (EAST1): a new toxin with an old twist." Crit Rev Microbiol **28**(1): 43-60.
Mendez-Arancibia, E., M. Vargas, et al. (2008). "Prevalence of different virulence factors and biofilm production in enteroaggregative Escherichia coli isolates causing diarrhea in children in Ifakara (Tanzania)." Am J Trop Med Hyg **78**(6): 985-989.
Mendez Arancibia, E., C. Pitart, et al. (2009). "Evolution of antimicrobial resistance in enteroaggregative Escherichia coli and enterotoxigenic Escherichia coli causing traveller's diarrhoea." J Antimicrob Chemother **64**(2): 343-347.
Menne, J., J. T. Kielstein, et al. (2012). "[Treatment of typical hemolytic-uremic syndrome. Knowledge gained from analyses of the 2011 E. coli outbreak]." Internist (Berl) **53**(12): 1420-1430.
Mohamed, J. A., H. L. DuPont, et al. (2011). "Single nucleotide polymorphisms in the promoter of the gene encoding the lipopolysaccharide receptor CD14 are associated with bacterial diarrhea in US and Canadian travelers to Mexico." Clin Infect Dis **52**(11): 1332-1341.
Mohamed, J. A., D. B. Huang, et al. (2007). "Association of putative enteroaggregative Escherichia coli virulence genes and biofilm production in isolates from travelers to developing countries." J Clin Microbiol **45**(1): 121-126.
Mondal, D., R. Haque, et al. (2009). "Attribution of malnutrition to cause-specific diarrheal illness: evidence from a prospective study of preschool children in Mirpur, Dhaka, Bangladesh." Am J Trop Med Hyg **80**(5): 824-826.
Monteiro-Neto, V., S. Y. Bando, et al. (2003). "Characterization of an outer membrane protein associated with haemagglutination and adhesive properties of enteroaggregative Escherichia coli O111:H12." Cell Microbiol **5**(8): 533-547.
Mora, A., C. Lopez, et al. (2012). "Emerging avian pathogenic Escherichia coli strains belonging to clonal groups O111:H4-D-ST2085 and O111:H4-D-ST117 with high virulence-gene content and zoonotic potential." Vet Microbiol **156**(3-4): 347-352.
Morabito, S., H. Karch, et al. (1998). "Enteroaggregative, Shiga toxin-producing Escherichia coli O111:H2 associated with an outbreak of hemolytic-uremic syndrome." J Clin Microbiol **36**(3): 840-842.
Morais, T. B., M. B. Morais, et al. (1998). "Bacterial contamination of the lacteal contents of feeding bottles in metropolitan Sao Paulo, Brazil." Bull World Health Organ **76**(2): 173-181.
Morin, N., A. E. Santiago, et al. (2013). "Characterization of the AggR regulon in enteroaggregative Escherichia coli." Infect Immun **81**(1): 122-132.
Mortensen, N. P., J. D. Fowlkes, et al. (2011). "Effects of sub-minimum inhibitory concentrations of ciprofloxacin on enteroaggregative Escherichia coli and the role of the surface protein dispersin." Int J Antimicrob Agents **38**(1): 27-34.
Mossoro, C., P. Glaziou, et al. (2002). "Chronic diarrhea, hemorrhagic colitis, and hemolytic-uremic syndrome associated with HEp-2 adherent Escherichia coli in adults infected with human immunodeficiency virus in Bangui, Central African Republic." J Clin Microbiol **40**(8): 3086-3088.
Moyo, S. J., S. Y. Maselle, et al. (2007). "Identification of diarrheagenic Escherichia coli isolated from infants and children in Dar es Salaam, Tanzania." BMC Infect Dis **7**: 92.
Murray, C. J. L. (2001). "The global burden of disease 2000 project: aims, methods, and data sources." Geneva: World Heath Organization.
NaNA (2015). "The Gambia World Breastfeeding Trends Initiative." WBTi.
Nataro, J. P. (2005). "Enteroaggregative Escherichia coli pathogenesis." Curr Opin Gastroenterol **21**(1): 4-8.

Nataro, J. P. (2006). "Atypical enteropathogenic Escherichia coli: typical pathogens?" Emerg Infect Dis **12**(4): 696.
Nataro, J. P. (2011). "Outbreak of hemolytic-uremic syndrome linked to Shiga toxin-producing enteroaggregative Escherichia coli O104:H4." Pediatr Res **70**(3): 221.
Nataro, J. P., Y. Deng, et al. (1995). "Heterogeneity of enteroaggregative Escherichia coli virulence demonstrated in volunteers." J Infect Dis **171**(2): 465-468.
Nataro, J. P., S. Hicks, et al. (1996). "T84 cells in culture as a model for enteroaggregative Escherichia coli pathogenesis." Infect Immun **64**(11): 4761-4768.
Nataro, J. P. and J. B. Kaper (1998). "Diarrheagenic Escherichia coli." Clin Microbiol Rev **11**(1): 142-201.
Nataro, J. P., J. B. Kaper, et al. (1987). "Patterns of adherence of diarrheagenic Escherichia coli to HEp-2 cells." Pediatr Infect Dis J **6**(9): 829-831.
Nataro, J. P., V. Mai, et al. (2006). "Diarrheagenic Escherichia coli infection in Baltimore, Maryland, and New Haven, Connecticut." Clin Infect Dis **43**(4): 402-407.
Nataro, J. P. and J. Martinez (1998). "Diagnosisand Investigation of Diarrheagenic Escherichia coli." Methods Mol Med **15**: 387-406.
Nataro, J. P., T. Steiner, et al. (1998). "Enteroaggregative Escherichia coli." Emerg Infect Dis **4**(2): 251-261.
Nataro, J. P., D. Yikang, et al. (1994). "AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative Escherichia coli." J Bacteriol **176**(15): 4691-4699.
Navarro-Garcia, F. (2010). "Enteroaggregative Escherichia coli plasmid-encoded toxin." Future Microbiol **5**(7): 1005-1013.
Navarro-Garcia, F., A. Canizalez-Roman, et al. (2001). "Plasmid-encoded toxin of enteroaggregative Escherichia coli is internalized by epithelial cells." Infect Immun **69**(2): 1053-1060.
Navarro-Garcia, F. and W. P. Elias (2011). "Autotransporters and virulence of enteroaggregative E. coli." Gut Microbes **2**(1): 13-24.
Navarro-Garcia, F., C. Sears, et al. (1999). "Cytoskeletal effects induced by pet, the serine protease enterotoxin of enteroaggregative Escherichia coli." Infect Immun **67**(5): 2184-2192.
Nguyen, T. V., P. V. Le, et al. (2005). "Antibiotic resistance in diarrheagenic Escherichia coli and Shigella strains isolated from children in Hanoi, Vietnam." Antimicrob Agents Chemother **49**(2): 816-819.
Nguyen, T. V., P. Le Van, et al. (2005). "Detection and characterization of diarrheagenic Escherichia coli from young children in Hanoi, Vietnam." J Clin Microbiol **43**(2): 755-760.
Nguyen, V. T. (2005). "Children's diarrhea in Hanoi, Vietnam: Importance of Enteric pathogens." Karolinska University Press.
Nitschke, M., F. Sayk, et al. (2012). "Association between azithromycin therapy and duration of bacterial shedding among patients with Shiga toxin-producing enteroaggregative Escherichia coli O104:H4." JAMA **307**(10): 1046-1052.
Niveditha, S., S. Pramodhini, et al. (2012). "The Isolation and the Biofilm Formation of Uropathogens in the Patients with Catheter Associated Urinary Tract Infections (UTIs)." J Clin Diagn Res **6**(9): 1478-1482.
O'Ryan, M., V. Prado, et al. (2005). "A millennium update on pediatric diarrheal illness in the developing world." Semin Pediatr Infect Dis **16**(2): 125-136.
Ochoa, T. J. and T. G. Cleary (2009). "Effect of lactoferrin on enteric pathogens." Biochimie **91**(1): 30-34.
Okeke, I. N. (2009). "Diarrheagenic Escherichia coli in sub-Saharan Africa: status, uncertainties and necessities." J Infect Dev Ctries **3**(11): 817-842.
Okeke, I. N., A. Lamikanra, et al. (2000). "Heterogeneous virulence of enteroaggregative Escherichia coli strains isolated from children in Southwest Nigeria." J Infect Dis **181**(1): 252-260.
Okeke, I. N. and J. P. Nataro (2001). "Enteroaggregative Escherichia coli." Lancet Infect Dis **1**(5): 304-313.

Okeke, I. N., O. Ojo, et al. (2003). "Etiology of acute diarrhea in adults in southwestern Nigeria." J Clin Microbiol **41**(10): 4525-4530.
Okeke, I. N., F. Wallace-Gadsden, et al. (2010). "Multi-locus sequence typing of enteroaggregative Escherichia coli isolates from Nigerian children uncovers multiple lineages." PLoS One **5**(11): e14093.
Okhuysen, P. C. and H. L. Dupont (2010). "Enteroaggregative Escherichia coli (EAEC): a cause of acute and persistent diarrhea of worldwide importance." J Infect Dis **202**(4): 503-505.
Olesen, B., F. Scheutz, et al. (2012). "Enteroaggregative Escherichia coli O78:H10, the cause of an outbreak of urinary tract infection." J Clin Microbiol **50**(11): 3703-3711.
Oliveira, A. and L. Cunha Mde (2010). "Comparison of methods for the detection of biofilm production in coagulase-negative staphylococci." BMC Res Notes **3**: 260.
Opintan, J. A., M. J. Newman, et al. (2010). "Pediatric diarrhea in southern Ghana: etiology and association with intestinal inflammation and malnutrition." Am J Trop Med Hyg **83**(4): 936-943.
Orden A. Jose, G. D.-B., Pilar Horeajo, Jose A. Ruiz-Santa-Quiteria, Angel Garcia, Javier Carrion (2017). "Ruminant are not a reservoir of enteroaggregative Escherichia coli." Austral Journal of Vetrinary Science **1,**(1,): 25-26.
Overdevest, I. T., I. Willemsen, et al. (2011). "Laboratory detection of extended-spectrum-beta-lactamase-producing Enterobacteriaceae: evaluation of two screening agar plates and two confirmation techniques." J Clin Microbiol **49**(2): 519-522.
Pabst, W. L., M. Altwegg, et al. (2003). "Prevalence of enteroaggregative Escherichia coli among children with and without diarrhea in Switzerland." J Clin Microbiol **41**(6): 2289-2293.
Pacheco, A. R. and V. Sperandio (2012). "Shiga toxin in enterohemorrhagic E.coli: regulation and novel anti-virulence strategies." Front Cell Infect Microbiol **2**: 81.
Page, A. J., C. A. Cummins, et al. (2015). "Roary: rapid large-scale prokaryote pan genome analysis." Bioinformatics **31**(22): 3691-3693.
Pai, M., G. Kang, et al. (1997). "An epidemic of diarrhoea in south India caused by enteroaggregative Escherichia coli." Indian J Med Res **106**: 7-12.
Paiva de Sousa, C. and J. D. Dubreuil (2001). "Distribution and expression of the astA gene (EAST1 toxin) in Escherichia coli and Salmonella." Int J Med Microbiol **291**(1): 15-20.
Panchalingam, S., M. Antonio, et al. (2012). "Diagnostic microbiologic methods in the GEMS-1 case/control study." Clin Infect Dis **55 Suppl 4**: S294-302.
Park, S. E. (2018). "The phylogeography and incidence of multi-drug resistant typhoid fever in sub-Saharan Africa." to be published.
Paterson, D. L. and R. A. Bonomo (2005). "Extended-spectrum beta-lactamases: a clinical update." Clin Microbiol Rev **18**(4): 657-686.
Paul, M., T. Tsukamoto, et al. (1994). "The significance of enteroaggregative Escherichia coli in the etiology of hospitalized diarrhoea in Calcutta, India and the demonstration of a new honey-combed pattern of aggregative adherence." FEMS Microbiol Lett **117**(3): 319-325.
Peirano, G., A. K. van der Bij, et al. (2012). "Molecular epidemiology over an 11-year period (2000 to 2010) of extended-spectrum beta-lactamase-producing Escherichia coli causing bacteremia in a centralized Canadian region." J Clin Microbiol **50**(2): 294-299.
Pereira, A. L., L. R. Ferraz, et al. (2007). "Enteroaggregative Escherichia coli virulence markers: positive association with distinct clinical characteristics and segregation into 3 enteropathogenic E. coli serogroups." J Infect Dis **195**(3): 366-374.
Philipson, C. W., J. Bassaganya-Riera, et al. (2013). "Animal models of enteroaggregative Escherichia coli infection." Gut Microbes **4**(4): 281-291.
Piddock, L. J. (2006). "Multidrug-resistance efflux pumps - not just for resistance." Nat Rev Microbiol **4**(8): 629-636.
Pinto, B. (2009). "The Global Burden of Childhood Diarrhoea." International Maternal and Child Health.

Pitout, J. D. and K. B. Laupland (2008). "Extended-spectrum beta-lactamase-producing Enterobacteriaceae: an emerging public-health concern." Lancet Infect Dis **8**(3): 159-166.
Ponnusamy, P., V. Natarajan, et al. (2012). "In vitro biofilm formation by uropathogenic Escherichia coli and their antimicrobial susceptibility pattern." Asian Pac J Trop Med **5**(3): 210-213.
Pratt, L. A. and R. Kolter (1998). "Genetic analysis of Escherichia coli biofilm formation: roles of flagella, motility, chemotaxis and type I pili." Mol Microbiol **30**(2): 285-293.
Presterl, E., R. Nadrchal, et al. (1999). "Enteroaggregative and enterotoxigenic Escherichia coli among isolates from patients with diarrhea in Austria." Eur J Clin Microbiol Infect Dis **18**(3): 209-212.
Price, M. N., P. S. Dehal, et al. (2010). "FastTree 2--approximately maximum-likelihood trees for large alignments." PLoS One **5**(3): e9490.
Prigent-Combaret, C., G. Prensier, et al. (2000). "Developmental pathway for biofilm formation in curli-producing Escherichia coli strains: role of flagella, curli and colanic acid." Environ Microbiol **2**(4): 450-464.
Puno-Sarmiento, J., L. Medeiros, et al. (2013). "Detection of diarrheagenic Escherichia coli strains isolated from dogs and cats in Brazil." Vet Microbiol **166**(3-4): 676-680.
Rajendran, P., S. S. Ajjampur, et al. (2010). "Pathotypes of diarrheagenic Escherichia coli in children attending a tertiary care hospital in South India." Diagn Microbiol Infect Dis **68**(2): 117-122.
Raju, B. and M. Ballal (2009). "Multidrug resistant enteroaggregative Escherichia coli diarrhoea in rural southern Indian population." Scand J Infect Dis **41**(2): 105-108.
Rappelli, P., E. Folgosa, et al. (2005). "Pathogenic enteric Escherichia coli in children with and without diarrhea in Maputo, Mozambique." FEMS Immunol Med Microbiol **43**(1): 67-72.
Rasko, D. A., D. R. Webster, et al. (2011). "Origins of the E. coli strain causing an outbreak of hemolytic-uremic syndrome in Germany." N Engl J Med **365**(8): 709-717.
Reither, K., R. Ignatius, et al. (2007). "Acute childhood diarrhoea in northern Ghana: epidemiological, clinical and microbiological characteristics." BMC Infect Dis **7**: 104.
Ren, C. P., R. R. Chaudhuri, et al. (2004). "The ETT2 gene cluster, encoding a second type III secretion system from Escherichia coli, is present in the majority of strains but has undergone widespread mutational attrition." J Bacteriol **186**(11): 3547-3560.
Rizzo, J. M. and M. J. Buck (2012). "Key principles and clinical applications of "next-generation" DNA sequencing." Cancer Prev Res (Phila) **5**(7): 887-900.
Robins-Browne, R. M. and E. L. Hartland (2002). "Escherichia coli as a cause of diarrhea." J Gastroenterol Hepatol **17**(4): 467-475.
Roca, A., P. C. Hill, et al. (2011). "Effects of community-wide vaccination with PCV-7 on pneumococcal nasopharyngeal carriage in the Gambia: a cluster-randomized trial." PLoS Med **8**(10): e1001107.
Rocha-Ramirez, L. M., U. Hernandez-Chinas, et al. (2016). "Pet serine protease from enteroaggregative Escherichia coli stimulates the inflammatory response activating human macrophages." BMC Microbiol **16**(1): 158.
Roche, J. K., A. Cabel, et al. (2010). "Enteroaggregative Escherichia coli (EAEC) impairs growth while malnutrition worsens EAEC infection: a novel murine model of the infection malnutrition cycle." J Infect Dis **202**(4): 506-514.
Rohde, H., J. Qin, et al. (2011). "Open-source genomic analysis of Shiga-toxin-producing E. coli O104:H4." N Engl J Med **365**(8): 718-724.
Rossen, J. W. A., A. W. Friedrich, et al. (2018). "Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology." Clin Microbiol Infect **24**(4): 355-360.
Ruiz-Perez, F., R. Wahid, et al. (2011). "Serine protease autotransporters from Shigella flexneri and pathogenic Escherichia coli target a broad range of leukocyte glycoproteins." Proc Natl Acad Sci U S A **108**(31): 12881-12886.
Saitou, N. and M. Nei (1987). "The neighbor-joining method: a new method for reconstructing phylogenetic trees." Mol Biol Evol **4**(4): 406-425.

Samie, A., C. L. Obi, et al. (2007). "Enteroaggregative Escherichia coli in Venda, South Africa: distribution of virulence-related genes by multiplex polymerase chain reaction in stool samples of human immunodeficiency virus (HIV)-positive and HIV-negative individuals and primary school children." Am J Trop Med Hyg **77**(1): 142-150.
Sang, W. K., J. O. Oundo, et al. (1997). "Multidrug-resistant enteroaggregative Escherichia coli associated with persistent diarrhea in Kenyan children." Emerg Infect Dis **3**(3): 373-374.
Santosham, M., E. M. Keenan, et al. (1997). "Oral rehydration therapy for diarrhea: an example of reverse transfer of technology." Pediatrics **100**(5): E10.
Sarantuya, J., J. Nishi, et al. (2004). "Typical enteroaggregative Escherichia coli is the most prevalent pathotype among E. coli strains causing diarrhea in Mongolian children." J Clin Microbiol **42**(1): 133-139.
Sauer, K. (2003). "The genomics and proteomics of biofilm formation." Genome Biol **4**(6): 219.
Savarino, S. J., A. Fasano, et al. (1991). "Enteroaggregative Escherichia coli elaborate a heat-stable enterotoxin demonstrable in an in vitro rabbit intestinal model." J Clin Invest **87**(4): 1450-1455.
Savarino, S. J., A. McVeigh, et al. (1996). "Enteroaggregative Escherichia coli heat-stable enterotoxin is not restricted to enteroaggregative E. coli." J Infect Dis **173**(4): 1019-1022.
Scalesky, I. e. a. (2001). "Diffuse and Enteroaggregative Patterns of Adherence of Escherichia coli isolated from stool of children in Northeastern Brazil." Brazilian Journal of Microbiology **32**: 313-319.
Scavia, G., M. Staffolani, et al. (2008). "Enteroaggregative Escherichia coli associated with a foodborne outbreak of gastroenteritis." J Med Microbiol **57**(Pt 9): 1141-1146.
Scheutz, F., E. M. Nielsen, et al. (2011). "Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing Escherichia coli O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011." Euro Surveill **16**(24).
Schmidt, H., C. Knop, et al. (1995). "Development of PCR for screening of enteroaggregative Escherichia coli." J Clin Microbiol **33**(3): 701-705.
Schubert, S., A. Rakin, et al. (1998). "Prevalence of the "high-pathogenicity island" of Yersinia species among Escherichia coli strains that are pathogenic to humans." Infect Immun **66**(2): 480-485.
Scott, S., A. Odutola, et al. (2014). "Coverage and timing of children's vaccination: an evaluation of the expanded programme on immunisation in The Gambia." PLoS One **9**(9): e107280.
Seeman, T., Goncalves da Silva A, Bulach DM, Schultz MB, Kwong JC, Howden BP. (2016). "Nullarbor Github " https://github.com/tseemann/nullarbor.
Seemann, T. (2014). "Prokka: rapid prokaryotic genome annotation." Bioinformatics **30**(14): 2068-2069.
Seifert, M. E. and P. I. Tarr (2012). "Azithromycin decolonization of STEC--a new risk emerges." Nat Rev Nephrol **8**(7): 429.
Sheikh, J., J. R. Czeczulin, et al. (2002). "A novel dispersin protein in enteroaggregative Escherichia coli." J Clin Invest **110**(9): 1329-1337.
Sheikh, J., E. G. Dudley, et al. (2006). "EilA, a HilA-like regulator in enteroaggregative Escherichia coli." Mol Microbiol **61**(2): 338-350.
Sheikh, J., S. Hicks, et al. (2001). "Roles for Fis and YafK in biofilm formation by enteroaggregative Escherichia coli." Mol Microbiol **41**(5): 983-997.
Smith, H. R., T. Cheasty, et al. (1997). "Enteroaggregative Escherichia coll and outbreaks of gastroenteritis in UK." Lancet **350**(9080): 814-815.
Spencer, J., H. R. Smith, et al. (1999). "Characterization of enteroaggregative Escherichia coli isolated from outbreaks of diarrhoeal disease in England." Epidemiol Infect **123**(3): 413-421.
Steiner, T. S., A. A. Lima, et al. (1998). "Enteroaggregative Escherichia coli produce intestinal inflammation and growth impairment and cause interleukin-8 release from intestinal epithelial cells." J Infect Dis **177**(1): 88-96.

Steiner, T. S., J. P. Nataro, et al. (2000). "Enteroaggregative Escherichia coli expresses a novel flagellin that causes IL-8 release from intestinal epithelial cells." J Clin Invest **105**(12): 1769-1777.
Stepanovic, S., D. Vukovic, et al. (2000). "A modified microtiter-plate test for quantification of staphylococcal biofilm formation." J Microbiol Methods **40**(2): 175-179.
Strauman, M. C., J. M. Harper, et al. (2010). "Enteroaggregative Escherichia coli disrupts epithelial cell tight junctions." Infect Immun **78**(11): 4958-4964.
Sutjita, M., A. R. Bouckenooghe, et al. (2000). "Intestinal secretory immunoglobulin A response to enteroaggregative Escherichia coli in travelers with diarrhea." Clin Diagn Lab Immunol **7**(3): 501-503.
Taddei, C. R., A. Fasano, et al. (2005). "Secreted autotransporter toxin produced by a diffusely adhering Escherichia coli strain causes intestinal damage in animal model assays." FEMS Microbiol Lett **250**(2): 263-269.
Tamura, K. and M. Nei (1993). "Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees." Mol Biol Evol **10**(3): 512-526.
Tamura, K., D. Peterson, et al. (2011). "MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods." Mol Biol Evol **28**(10): 2731-2739.
Telli, M., E. Guiral, et al. (2010). "Prevalence of enterotoxins among Escherichia coli isolates causing bacteraemia." FEMS Microbiol Lett **306**(2): 117-121.
Tokuda, K., J. Nishi, et al. (2010). "Characterization of typical and atypical enteroaggregative escherichia coli in Kagoshima, Japan: biofilm formation and acid resistance." Microbiol Immunol **54**(6): 320-329.
Tompkins, D. S., M. J. Hudson, et al. (1999). "A study of infectious intestinal disease in England: microbiological findings in cases and controls." Commun Dis Public Health **2**(2): 108-113.
Toole, M. J. and R. J. Waldman (1997). "The public health aspects of complex emergencies and refugee situations." Annu Rev Public Health **18**: 283-312.
Torres, A. G., X. Zhou, et al. (2005). "Adherence of diarrheagenic Escherichia coli strains to epithelial cells." Infect Immun **73**(1): 18-29.
Toshima, H., E. Uenaka, et al. (2004). "Detection and isolation of Escherichia coli with a coding gene for enteroaggregative Escherichia coli heat-stable enterotoxin 1 from food and comparison with fecal isolates." J Food Prot **67**(10): 2117-2122.
Unicef (2013). "Joint UNICEF/WHO Statement at the Launching of teh Rotavirus Vaccine." https://www.unicef.org/gambia/media_8418.html.
Vasudevan, P., T. Annamalai, et al. (2003). "Behavior of enteroaggregative Escherichia coli in bottled spring and mineral water." J Food Prot **66**(3): 497-500.
Veilleux, S., N. Holt, et al. (2008). "Escherichia coli EAST1 toxin toxicity of variants 17-2 and O 42." Comp Immunol Microbiol Infect Dis **31**(6): 567-578.
Velarde, J. J., K. M. Varney, et al. (2007). "Solution structure of the novel dispersin protein of enteroaggregative Escherichia coli." Mol Microbiol **66**(5): 1123-1135.
Vila, J., A. Gene, et al. (1998). "A case-control study of diarrhoea in children caused by Escherichia coli producing heat-stable enterotoxin (EAST-1)." J Med Microbiol **47**(10): 889-891.
Vila, J., M. Vargas, et al. (1999). "Antimicrobial resistance of diarrheagenic Escherichia coli isolated from children under the age of 5 years from Ifakara, Tanzania." Antimicrob Agents Chemother **43**(12): 3022-3024.
Vila, J., M. Vargas, et al. (2001). "Susceptibility patterns of enteroaggregative Escherichia coli associated with traveller's diarrhoea: emergence of quinolone resistance." J Med Microbiol **50**(11): 996-1000.
Villaseca, J. M., F. Navarro-Garcia, et al. (2000). "Pet toxin from enteroaggregative Escherichia coli produces cellular damage associated with fodrin disruption." Infect Immun **68**(10): 5920-5927.

Vossenkamper A., M. T. T., Marches O. (2011). "Always one step ahead: How pathogenic bacteria use the type III secretion system to manipulate the intestinal mucosal immunne system " Journal of Inflammation **8**(11): PMID:21539730; http://dx.doi.org/21539710.21531186/21531476-21539255-21539738-21539711

Wakimoto, N., J. Nishi, et al. (2004). "Quantitative biofilm assay using a microtiter plate to screen for enteroaggregative Escherichia coli." Am J Trop Med Hyg **71**(5): 687-690.

Wani, S. A., I. Hussain, et al. (2013). "Diarrhoeagenic Escherichia coli and salmonellae in calves and lambs in Kashmir absence, prevalence and antibiogram." Rev Sci Tech **32**(3): 833-840.

Wani, S. A., I. Hussain, et al. (2012). "Putative Virulence Genes and Biofilm Production Among Typical Enteroaggregative Escherichia coli Isolates from Diarrhoeic Children in Kashmir and Andhra Pradesh." Indian J Microbiol **52**(4): 587-592.

Warren, W. E. (1983). "Diarrhea." CAN. FAM. PHYSICIAN **29**(09): 1641-1645.

Weintraub, A. (2007). "Enteroaggregative Escherichia coli: epidemiology, virulence and detection." J Med Microbiol **56**(Pt 1): 4-8.

Wells, J. M., O. Rossi, et al. (2011). "Epithelial crosstalk at the microbiota-mucosal interface." Proc Natl Acad Sci U S A **108 Suppl 1**: 4607-4614.

WHO-document (1992). "Reading on diarrhoea: student manual." WHO Liberary Cataloguing in Publication Data.

WHO (1992). "The Epidemiology and Etiology of Diarrhoea." http://rehydrate.org/diarrhoea/tmsdd/1med.html.

WHO (2017). "Diarrhoeal disease." http://www.who.int/mediacentre/factssheets/fs330/en/.

Yatsunenko, T., F. E. Rey, et al. (2012). "Human gut microbiome viewed across age and geography." Nature **486**(7402): 222-227.

Zamboni, A., S. H. Fabbricotti, et al. (2004). "Enteroaggregative Escherichia coli virulence factors are found to be associated with infantile diarrhea in Brazil." J Clin Microbiol **42**(3): 1058-1063.

Zhang, R., D. X. Gu, et al. (2016). "Comparative genetic characterization of Enteroaggregative Escherichia coli strains recovered from clinical and non-clinical settings." Sci Rep **6**: 24321.

Zhou, K., M. Lokate, et al. (2016). "Use of whole-genome sequencing to trace, control and characterize the regional expansion of extended-spectrum beta-lactamase producing ST15 Klebsiella pneumoniae." Sci Rep **6**: 20840.

Zhou, Z., J. Ogasawara, et al. (2002). "An outbreak of gastroenteritis in Osaka, Japan due to Escherichia coli serogroup O166:H15 that had a coding gene for enteroaggregative E. coli heat-stable enterotoxin 1 (EAST1)." Epidemiol Infect **128**(3): 363-371.

Chapter 12: Appendix

Appendix A – Study approvals

Scientific Coordinating Committee
MRC Unit: The Gambia, Fajara
PO Box 273 Banjul, The Gambia
West Africa
Switchboard (+220) 4495442/6 Ext 2308
Fax (+220) 4495919/4496513
E-mail: scc@mrc.gm
Website (intranet): \\open.mrc.gm\home.asp

17 October 2011

Dr Martin Antonio
Vaccinology Research Theme
MRC Unit, The Gambia
Fajara

Dear Dr Antonio

SCC 1254v3, Molecular approaches in the diagnosis of diarrhoeal disease in children from the rural Gambia

Thank you for resubmitting your revised proposal dated 16 October 2011 addressing the issues raised by the SCC at its meeting held on 3 October 2011.

I am pleased to approve your proposal which will be forwarded to the Ethics Committee for consideration at its meeting on 28 October 2011.

With best wishes

Yours sincerely

Professor Beate Kampmann
Chair, Scientific Coordinating Committee

Cc: Mr Usman Nurudeen Ikumapayi

Additional documents submitted for review:-

- Consent Form for SCC 1054(Cases), Version 6.0 – 14 April 2011
- Consent Form for SCC 1054 (controls), Version 21August 2006

The Gambia Government / MRC Joint
ETHICS COMMITTEE

C/o MRC Unit: The Gambia, Fajara
P. O. Box 273, Banjul
The Gambia, West Africa
Fax: +220 – 4495919 or 4496 513
Tel: +220 – 4495442-6 Ext. 2308

1 November 2011

Dr Martin Antonio
Vaccinology Research Theme
MRC Unit, The Gambia
Fajara

Dear Dr Antonio

SCC 1254v3, Molecular approaches in the diagnosis of diarrhoeal disease in children from the rural Gambia

Thank you for submitting your proposal dated 16 October 2011 for consideration by the Gambia Government/MRC Joint Ethics Committee at its meeting held on 28 October 2011.

The Committee carefully reviewed the proposed PhD project and was pleased to give its approval.

Best wishes

Yours sincerely

Mr Malcolm Clarke
Chairman, Gambia Government/MRC Joint Ethics Committee

Cc: Mr Usman Nurudeen Ikumapayi

Additional documents submitted for review:-

- Consent Form for SCC 1054 (Cases), Version 6.0 – 14 April 2011
- Consent Form for SCC 1054 (controls), 21 August 2006

The Gambia Government / MRC Joint Ethics Committee:

Mr Malcolm Clarke, Chairman
Mrs Kathy Hill, Secretary
Ms Naffie Jobe, 2nd Secretary
Professor Ousman Nyan, Scientific Advisor
Mr Dawda Jagne
Mrs Bertha Mboge
Mr Modou Phall
Professor Tumani Corrah
Dr Stephen Howie
Dr Mamady Cham
Dr Lamin Sidibeh
Dr Kalifa Bojang
Mr Malamin Sonko

Appendix B – Good Clinical Practice / Good Clinical Laboratory Practice training & Certificates

Hereby Certifies that

USMAN NURUDEEN IKUMAPAYI

has completed the e-learning course

ICH GOOD CLINICAL PRACTICE

with a score of

83%

on

31/10/2016

This e-learning course has been formally recognised for its quality and content by the following organisations and institutions

This ICH E6 GCP Investigator Site Training meets the Minimum Criteria for ICH GCP Investigator Site Personnel Training identified by ***TransCelerate BioPharma*** *as necessary to enable mutual recognition of GCP training among trial sponsors.*

Global Health Training Centre
globalhealthtrainingcentre.org/elearning

Certificate Number 167814

Appendix C – Creating standard curves with genomic DNA

(Adapted from Applied Biosystem 2003)

1. Extract total DNA from sample (stool) using appropriate extraction kit
2. Measure/Quantify the concentration of the total DNA (EAEC 042 – 55ng)
3. Identify the genome size of the organism (EAEC 042 – 5,355323bp)
4. Identify the mass of the genome by inserting the bacterial genome size value in the following formula

M = [n][1.096e-21 g/bp]
Where: n = genome size (bp),
m = mass,
e-21 = $x10^{21}$

5. The mass of the EAEC 042 genome is calculated and convert to picogram units

   M = [5.0e6 ~~bp~~][1.096e-21g/~~bp~~] = 5.5e-15g

6. Convert the mass to picogram as below

   [5.5e-15~~g~~][1e12pg/~~g~~] = 0.0055pg

7. Divide the mass of the genome by the copy number of the gene of interest (*aaiC* & *aatA*) per haploid genome

   0.0055 pg/genome ÷ 1 copy gene T /genome = [0.0055pg/genome][genome/1 copy]
   = 0.0055p/1 copy gene T

Therefore, 0.0055 pg of EAEC 42 genome contains one of the aatA gene and aaiC gene

8. Calculate the mass of gDNA containing the copy number of interest, that is 5,000000 to 5 copies (5×10^6 to 5×10^{-1})

Copy # of interest x mass of genome = mass of gDNA required

Copy #		Mass of gDNA needed (pg)
5000000		27500
500000		2750
50000		275
5000	X 0.0055pg	27.5
500		2.75
50		0.275
5		0.0275
0.5		0.00275

9. Calculate the concentrations of the gDNA needed to achieve the copy numbers of interest Divide the mass required (step 4) by the volume to pipette into each reaction

10. 2 µL or 5uL of gDNA solution is pipetted into each PCR reaction

11. Calculate the concentration of gDNA needed to achieve the required masses of gDNA

Copy #	Mass of gDNA needed (pg)		Final Conc. (pg/µl) of gDNA
5000000	27500		13750
500000	2750		1375
50000	27.5		137.5
5000	2.75	÷ 2µL	13.75
500	0.275		1.375
50	0.0275		0.1375
5	0.00275		0.01375
0.5			0.001375

12. Prepare a serial dilution of the gDNA using the formula C1V1 = C2V2

13. Determine the stock concentration using spectrophotometric analysis (picogreen) given 55ng or 0.055 µg/µL (55000pg/µL) as C1. Each dilution has a final volume (V2) of 100µL
14. [55000 pg/ µL][V1] = [13750 pg/µL][100 µL] = 25 µL
15. V1 = 25 µL
16. Volume of diluents = 100 µL – 25 µL = 75 µL

17. To achieve the final volume of 100 µL, add 25 µL of stock gDNA to 75 µL of diluents (Nuclease free water)

Table below is representation of the calculated volumes of gDNA and diluent for all the 7 dilutions

Dilution	Source gDNA diluents	Initial Conc (pg/ µL)	Vol of gDNA (µL)	Vol of Diluent (µL)	Final Vol. (µL)	Final conc. Of dilution (pg/ µL)	Resulting copy DNAse aaiC/aatA gene / 1 µL

Appendix D – Media preparations and reagents

Congo red agar (CRA)

- Weigh 37g of Brain Heart Infusion (BHI) broth
- Weigh 50g of Sucrose
- Weigh 10g of Nutrient agar or Agar No.1 + optional 0.1g glucose
- Dissolve in 1 L Distilled/H_2O and sterilise at 121°C for 15 minutes
- Separately, weigh 0.8g/L Congo-red and dissolve in d/H_2O, sterilise appropriately
- Add the sterilised Congo-red to the other agar medium constituents after cooling to 55°C
- Set up QC using appropriate control strains

MacConkey agar – Oxoid CM0007

- Suspend 52g in 1 litre of distilled-water and bring to boil to dissolve completely
- Sterilise by autoclaving at 121°C for 15 minutes
- Allow to cool to 55°C in cooling water bath
- Distribute in 20mls in 30mm Petri dishes
- Flame the agar surface and allow to solidify
- Dry the gel surface before use
- Perform quality control (QC)
- Label and store in plastic bag at +4°C to +8°C

Muller-Hinton agar – Oxoid CM337

- Suspend 38g in 1 litre of distilled-water and bring to boil to dissolve completely
- Sterilise by autoclaving at 121°C for 15 minutes
- Allow to cool to 55°C
- Distribute in 20mls in 30mm petri dishes
- Flame the agar surface and allow to solidify
- Dry the gel surface before use

- Perform quality control (QC)
- Label and store in plastic bag at +4°C to +8°C

Nutrient agar – Oxoid CM0003

- Suspend 28g in 1 litre of distilled-water and bring to boil to dissolve completely
- Sterilise by autoclaving at 121°C for 15 minutes
- Allow to cool to 55°C
- Distribute in 20mls in 30mm petri dishes
- Flame the agar surface and allow to solidify
- Dry the gel surface before use
- Perform quality control (QC)
- Label and store in plastic bag at +4°C to +8°C

Preparation of Tryptic soy broth – (SIGMA-ALDRICH T8907) for Slime production

- Weigh 30g of Tryptic soy broth
- Add 10g of glucose
- Mix to dissolve in 1 litre of distilled-water
- Distribute in bijou bottles or borosilicate tubes
- Sterilise by autoclaving at 121°C for 15 minutes
- Perform quality control (QC)
- Label and store in plastic bag at +4°C to +8°C

Appendix E: Peer-reviewed journal from the thesis

Am. J. Trop. Med. Hyg., 97(4), 2017, pp. 997–1004
doi:10.4269/ajtmh.16-0705
Copyright © 2017 by The American Society of Tropical Medicine and Hygiene

Identification of Subsets of Enteroaggregative *Escherichia coli* Associated with Diarrheal Disease among Under 5 Years of Age Children from Rural Gambia

Usman N. Ikumapayi,[1,2] Nadia Boisen,[3] Mohammad J. Hossain,[2] Modupeh Betts,[1] Modou Lamin,[2] Debasish Saha,[1] Brenda Kwambana-Adams,[1] Michel Dione,[1] Richard A. Adegbola,[1] Anna Roca,[2,4] James P. Nataro,[5,6] and Martin Antonio[1,7,8]*

[1]*Vaccines and Immunity Theme, Medical Research Council Unit The Gambia, Banjul, The Gambia;* [2]*Disease Control and Elimination Theme, Medical Research Council Unit The Gambia, Banjul, The Gambia;* [3]*Department of Microbiological Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark;* [4]*Faculty of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London, United Kingdom;* [5]*Centre for Vaccine Development, University of Maryland School of Medicine, Baltimore, Maryland;* [6]*Department of Pediatrics, University of Virginia School of Medicine, Charlottesville, Virginia;* [7]*Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom;* [8]*Microbiology and Infection Unit, Warwick Medical School, University of Warwick, Coventry, United Kingdom*

Abstract. Enteroaggregative *Escherichia coli* (EAEC) cause acute and persistent diarrhea, mostly in children worldwide. Outbreaks of diarrhea caused by EAEC have been described, including a large outbreak caused by a Shiga toxin expressing strain. This study investigated the association of EAEC virulence factors with diarrhea in children less than 5 years. We characterized 428 EAEC strains isolated from stool samples obtained from moderate-to-severe diarrhea cases (157) and healthy controls (217) children aged 0–59 months recruited over 3 years as part of the Global Enteric Multicenter Study (GEMS) in The Gambia. Four sets of multiplex polymerase chain reaction were applied to detect 21 EAEC-virulence genes from confirmed EAEC strains that target pCVD432 (aatA) and AAIC (aaiC). In addition, Kirby-Bauer disc diffusion antimicrobial susceptibility testing was performed on 88 EAEC strains following Clinical Laboratory Standard Institute guidelines. We observed that the plasmid-encoded enterotoxin [odds ratio (OR): 6.9, 95% confidence interval (CI): 2.06–29.20, $P < 0.001$], aggregative adherence fimbriae/I fimbriae (aggA) [OR: 2.2, 95% CI: 1.16–4.29, $P = 0.008$], and hexosyltransferase (capU) [OR: 1.9, 95% CI 1.02–3.51, $P = 0.028$] were associated with moderate-to-severe diarrhea among children < 12 months old but not in the older age strata (> 12 months). Our data suggest that some EAEC-virulent factors have age-specific associations with moderate-to-severe diarrhea in infants. Furthermore, our study showed that 85% and 72% of EAEC strains tested were resistant to sulphamethoxazole-trimethoprim and ampicillin, respectively. Sulphamethoxazole-trimethoprim and ampicillin are among the first-line antibiotics used for the treatment of diarrhea in The Gambia.

INTRODUCTION

Enteroaggregative *Escherichia coli* (EAEC) is an important causative agent of both acute and persistent diarrhea among adults and children worldwide[1] and it has been among the most common *E. coli* pathotypes causing diarrhea among children less than 5 years of age in some developing countries.[2] Several outbreaks of EAEC diarrhea have been reported in both developed and developing nations and infants are the most affected.[3–6] EAEC has been implicated in travelers' diarrhea[7,8] and persistent diarrhea among human immunodeficiency virus–infected individuals.[9] This pathotype was implicated in a massive outbreak of hemolytic uremic syndrome in Germany in 2011.[10] The clinical presentation of EAEC infection is characterized by watery and mucoid diarrhea with low-grade fever and insignificant vomiting.[11,12]

The pathogenesis of EAEC diarrhea is thought to comprise colonization of the intestinal mucosa, followed by elaboration of enterotoxins and cytotoxins and the release of proinflamatory cytokines from infected epithelial cells,[13,14] induced by the EAEC adherence factors called aggregative adherence fimbriae (AAF). In addition, EAEC strains characteristically enhance mucus secretion from the mucosa, potentially trapping the bacterium in a bacterium-mucus biofilm.[1] A distinctive feature of EAEC is its ability to elicit characteristic stacked brick-liked aggregative adherence to HEp-2 or HeLa cells, a test that remains the gold standard to identify this pathotype.[15] EAEC strains express several genes that may confer virulence and are highly heterogeneous regarding the combination of these virulence genes, which are encoded on the bacterial chromosome or on an EAEC-specific plasmid-designated pAA. The majority of EAEC strains harbor a transcriptional activator of the AraC/XylS called AggR, which control genes on both the plasmid and the chromosome. Among the genes under AggR control includes those that encode the AAF where at least five variants exist. These genes encoding the major structural pilin subunits are designated as *aggA* (AAF/I), *aafA* (AAF/II), *agg3A* (AAF/III), *agg4A* (AAF/IV), and *agg5A* (AAF/V).[16–18] Other plasmid-borne potential virulence factors include the EAEC heat-stable enterotoxin 1 (EAST1) (encoded by the *astA* gene),[19] an anti-aggregation protein called dispersin (encoded by the *aap* gene), and a transporter apparatus for dispersin called Aat (encoded by the *aat* genes). EAEC frequently harbor members of the serine protease autotransporters of Enterobacteriaceae (SPATEs), which have been described as enterotoxins and cytotoxins. The heat-labile enterotoxin/cytotoxin called Plasmid-encoded toxin (Pet)[20] has been implicated in causing cytotoxic effects on the human intestinal mucosa. Other SPATEs carried by EAEC strains include the cryptic protease called SepA, and the mucinase called Pic (protein involved in intestinal colonization),[21,22] which is encoded on the chromosome. Other important chromosomal gene that encodes virulence markers include 1) Irp2 (iron repressible high-molecular-weight protein 2) a protein responsible for yersiniabactin biosynthesis and 2) flagellin, which interacts with the epithelial cells, leading to the secretion of an intestinal interleukin-8.[23] The EAEC genome has been found to be markedly mosaic; thus, the various putative virulence

*Address correspondence to Martin Antonio, Vaccines and Immunity Theme, Medical Research Council Unit, Atlantic Road, Fajara, Banjul, The Gambia. E-mail: mantonio@mrc.gm

factors are found inconsistently among individual strains, suggesting that some strains considered EAEC may be truly virulent, and others not.[20]

Several studies have shown that EAEC is the most frequently detected *E. coli* pathotype in humans, particularly among children from both developed and developing countries.[24–26] The Global Enteric Multicenter Study (GEMS) comprised identical case–control studies of moderate-to-severe diarrhea (MSD) among children under 5 years of age at four sites in sub-Saharan Africa and three in south Asia showing high frequency of EAEC.[27] Although EAEC was not associated with MSD disease in the GEMs study, a subsequent analysis of the association of individual EAEC genes alone and in combination among EAEC isolates from MSD cases and controls of GEMS in Mali by Boisen found that SepA protease was associated with MSD.[28] In the study presented, we replicated the analysis by Boisen, scoring the presence of 21 putative EAEC virulence factors from 428 EAEC isolates randomly selected among 741 EAEC isolates obtained from diarrheal and nondiarrheal children enrolled in the GEMS study to characterize the virulence genes in children from The Gambia. We analyzed these EAEC virulence genes by age strata (0–11, 12–23, and 24–59 months). Furthermore, antimicrobial resistance patterns were investigated on a 20% of the EAEC strains selected at random.

MATERIALS AND METHODS

Patients and enteric pathogens tested. The study participants were enrolled as part of the 3-year (December 2007 to December 2010) prospective case–control GEMS in The Gambia. The clinical and microbiologic methods for the GEMS have been published.[29] In GEMS, children aged 0- to 59-month old from a defined census population attending to sentinel health center with sign of MSD were enrolled. One to three healthy controls matched by age, sex, and community were recruited within 14 days after the enrolment of cases.[30] As part of the main study, stool samples were processed for enteric bacterial pathogens (including *E. coli*, *Salmonella* spp., Shigella spp., Vibrio spp., *Campylobacter* spp., and Aeromonas spp.), Viruses (Sapovirus, Norovirus, and Rotavirus), and parasites (*Giardia lamblia*, Cryptosporidium species, and *Entamoeba histolytica*).[31] A total of 2,598 children were enrolled in The Gambia and 741 (28.5%) children had EAEC isolated (278 cases and 463 controls).

For this ancillary study, the matching design was not maintained. We randomly selected 428 (58%) of the participants with samples positive for EAEC (157 cases and 271 controls) (Table 1) to conduct genotypic characterization of the 21 virulence factors. Of the 157 cases, 94 (60%) and 63 (40%) account for male and female, respectively, whereas among the controls 151 (56%) account for male and 120 (44%) account for female.

Testing of 21 enteroaggregative *E. coli* virulence genes. From an overnight growth on MacConkey agar (Oxoid, Hampshire, United Kingdom), three suspected colonies of *E. coli* (typically lactose fermenting) for each patient were purified and identified as *E. coli* using the biochemical reagent kit Analytical Profile Index 20 E (BioMeriux Ltd, Hampshire, United Kingdom). The resultant confirmed *E. coli* were screened to detect EAEC, ETEC, and EPEC pathotypes using a multiplex polymerize chain reaction (PCR) GEMS protocol.[31] In this study, we performed monoplex PCR on each isolates that has initially showed presence of EAEC in GEMS *E. coli* multiplex PCR protocol. The target sought for EAEC are the EAEC plasmid-encoded gene *aat*A (primer CVD432F–sequence 5′-CTGGCGAAAGACTGTATCAT-3′and primer CVD432R–sequence 5′-CAATGTATAGAAATCCGCTGTT-3′) and the EAEC chromosomally encoded *aai*C (primer AAIC F–sequence 5′-ATTGTCCTCAGGCATTTCAC-3′ and primer AAIC R-sequence 5′-ACGACACCCCTGATAAACAA-3′), these two loci are known as virulence determinants. EAEC colonies were investigated for the presence of the 21 putative virulence genes using four multiplex PCRs as previously described.[28] The 21 genes were grouped into four. On each group, multiplex PCR was performed. On group 1 (*sat*, *sepA*, *pic*, *sigA*, plasmid-encoded enterotoxin [*pet*], and *astA*), multiplex PCR master mix was achieved using Qiagen kit (Catalogue no. 206143 [Manchester, United Kingdom]) following the manufacturer's instructions. Multiplex PCR assay was performed in a final reaction volume of 25 μL consists of 12.5 μL mastemix (MM), 2.5 μL Q-solution, 6 μL primer (MM), 2.5 μL H_2O, and 1.5 μL DNA template. PCR reaction cycles were as follows: 15 minutes preheating at 95°C at the start, 50 seconds denaturation at 94°C, annealing for 1.5 minutes, and extension at 72°C for 1.5 minutes with 35 cycles returning to step 2. The final extension was 10 minutes at 72°C.

On group 2 (*aatA*, *aggR*, *aaiC*, *aaP*, and ORF3), group 3 (*aafC*, *agg3/4C*, *agg3A*, *aafA*, *aggA*, and *agg4A*), and group 4 (*air*, *capU*, *ailA*, and ORF61) Fementers kit (Catalogue no. K0171 [Paisley, United Kingdom]) was used for the PCR master mix (2X) following the manufacturer's instructions. Multiplex PCR assay was achieved in a final reaction volume of 25 μL that compose of 12.5 μL (MM), 1 μL (25 mM magnesium chloride), 5 μL primer (MM), 5 μL H_2O, and 1.5 μL DNA template. PCR reaction cycles were achieved as follows: 2 minutes preheating at 95°C at the start, 50 seconds denaturation at 94°C, annealing at 57°C (58°C for groups 3 and 4) for 1.5 minutes, and extension at 72°C for 1.5 minutes with 35 cycles returning to step 2. The final extension was 10 minutes at 72°C.

Amplifications were performed using Thermocycler (TECHNE Flexigen, Model FFG02FSD, Serial 11733-1 [Paisley, United Kingdom]). Amplified PCR products were analyzed on a 2% agarose gel containing ethidium bromide (final concentration of 0.5 μg/mL in 1 × TBE buffer and visualized on the 2%

TABLE 1
Baseline information of study population

Demographic factors	Case (*N* = 157) no. (%)	Control (*N* = 271) no. (%)	Total (*N* = 428) no. (%)	OR (95% CI)	*P* value
Age (month)					
0–11	85 (54.1)	132 (48.7)	217 (50.7)	1.2 (0.82–1.88)	0.278
12–23	61 (38.9)	105 (38.8)	166 (38.8)	1.0 (0.65–1.53)	0.982
24–59	11 (7.0)	34 (12.6)	45 (10.5)	0.5 (0.23–1.10)	0.071

CI = confidence interval; OR = odds ratio.

[w/v] agarose gel under ultraviolet radiation. The gel images were captured digitally with a gel documentation system.

The *E. coli* strains used as controls for detection of the target genes are C1010-00 (*sat*, *sepA*, *agg3/4C*, and *agg4A*), JM221 (*sat* and *aggA*), 042 (*pic*, *pet*, *astA*, *aatA*, *aggR*, *aaiC*, *aap*, *ORF3*, *aafC*, *aaFA*, *air*, *capU*, and *eilA*), 55989 (*sigA*, *agg3A/4C*, and *agg3A*), 63 (*sigA*, *agg3/4C*, and *agg4A*), and 17-2 (*aggA*). GIBCO distilled water (DNase/RNase free, Catalogue no. 10977-035 [Paisley, United Kingdom]) was used as negative control.

Antimicrobial assay. For antimicrobial susceptibility testing, 88 EAEC isolates were randomly selected from the 428 EAEC isolates used for the virulence genes assay. Disk diffusion (Kirby Bauer) method for susceptibility testing that allows categorization of bacterial isolates as susceptible, resistant, or intermediate to eight commercially acquired antimicrobial agents which include ampicillin 10 μg, cotrimoxazole 25 μg, tetracycline 30 μg, ceftazidime 30 μg, ciprofloxacin 5 μg, ceftriaxone 30 μg, cefoxitin 30 μg, and nalidixic acid 30 μg (Oxoid) were used. The Clinical Laboratory Standard Institute version M100-S22, Vol. 32, No. 3, 2012, guidelines were followed. *E. coli* ATCC 25922 was used for quality control antibiotic susceptibility assay.

Ethical consideration. The GEMS obtained ethical clearance from The Gambian Government/Medical Research Council (MRC) Ethics committee following the scrutiny of the study proposal by the MRC Scientific Coordinating Committee.

Statistical analyses. Bivariate analysis was applied to compare prevalence of virulence factors between cases and controls in different age group using STATA 12 reporting odds ratios (ORs) with 95% confidence intervals (CIs). A two-sided P value < 0.05 was considered statistically significant.

In addition, we used Classification and Regression Tree (CART) pro-Version 6.0 (Salford Systems, San Diego, CA) software to input 21 factors of interest as binary (present/absent) independent variables. Although, case–control status was input as the binary dependent outcome variable.

Significance of combinations of EAEC genes. We generated a virulence factor score (VFS), representing the collective number of virulence loci present in each strain. To consider the combination factors, we used CART analysis (Supplemental Figures 1 and 2), which builds a model in stepwise fashion to yield the combination of factors most strongly associated with the queried outcome. Each branch of a CART output tree ends in a terminal "node"; each observation falls into exactly one terminal node; and each terminal node is uniquely defined by a set of rules, such as having or not having a certain factor.

We considered all genotypic and phenotypic assays performed and considered the association with case status versus control status (Supplemental Figures 1 and 2).

RESULTS

Among all EAEC strains in cases and controls (N = 428), the age and sex distribution were similar among cases and controls except for a lower prevalence of children above 23 months among cases (Table 1). Overall, orf61 (*aar*) was the most commonly detected gene, (69.6%). This was followed by the cryptic ORF3 (64%), *capU* (62%), *aggR* (60.1%), *astA* (51.4%), *eilA* (48.3%), and *aap* (46.3%); the rest of the genes were present in less than 40% of isolates (Table 2). Analysis of the EAEC virulence genes in all age groups together, showed that only four of the 21 genes assayed (*sepA*, *pet*, *astA*, and *capU*) were more prevalent among cases. Prevalence of AAF/I encoded by *aggA* gene was slightly higher in cases than controls (29.9% versus 22.9%) (OR: 1.4, 95% CI: 0.89–2.29, P = 0.106) (Table 2). The frequency of other AAF pilin genes, AAF/II (aafA), and AAF/III (agg3A) were low in both cases and controls but slightly high for AAF/IV (agg4A) in cases compared with controls. However, the AAF usher-encoding gene agg3/4C was similar in cases and controls (36.9% versus 35.4%, respectively). Of the five SPATE genes (*sat*, *pet*, *sigA*, *pic*, and *sepA*), prevalence of *sepA* (OR: 1.6, 95% CI:

Table 2
Distribution of EAEC virulence genes from case and control children (age 0–59 months)

Gene class		Virulence gene	Case (N = 157) no. (%)	Control (N = 271) no. (%)	Total (N = 428) no. (%)	OR (95% CI)	χ^2	P value
		aatA	51 (32.5)	82 (30.3)	133 (31.1)	1.1 (0.71–1.73)	0.2	0.631
		aggR	97 (61.8)	161 (59.4)	258 (60.3)	1.1 (0.72–1.69)	0.2	0.628
		aaP	77 (49.0)	121 (44.7)	198 (46.3)	1.2 (0.78–1.80)	0.8	0.379
		ORF3	106 (67.5)	168 (62.9)	274 (64.0)	1.3 (0.82–1.98)	1.3	0.251
		capU	108 (68.8)	158 (58.3)	266 (62.2)	1.6 (1.02–2.45)	4.7	0.031
		aar	110 (70.1)	188 (69.4)	298 (69.6)	1.0 (0.66–1.63)	0.1	0.880
pAA	A	***aafC***	7 (4.5)	16 (6.0)	23 (5.4)	0.7 (0.25–1.97)	0.4	0.522
P	D	***agg3/4C***	58 (36.9)	96 (35.4)	154 (36.8)	1.1 (0.69–1.64)	0.1	0.752
L	H	***agg3A***	10 (6.4)	28 (10.3)	38 (9.8)	0.6 (0.25–1.29)	1.9	0.164
A	E	***aafA***	3 (1.9)	15 (5.5)	18 (4.2)	0.3 (0.06–1.20)	3.2	0.071
S	S	***aggA***	47 (29.9)	62 (22.9)	109 (25.5)	1.4 (0.89–2.29)	2.6	0.106
M	I	***agg4A***	15 (9.6)	16 (6.0)	31 (7.2)	1.7 (0.74–3.75)	1.9	0.160
	N							
I	T	***astA***	91 (58.6)	129 (47.6)	220 (51.4)	1.5 (1.00–2.30)	4.3	0.038
D	O	***sat***	29 (18.5)	56 (20.7)	85 (19.9)	0.9 (0.51–1.47)	0.3	0.583
	X	***sepA***	50 (31.9)	62 (22.9)	112 (26.2)	1.6 (0.99–2.49)	4.1	0.041
	I	***pet***	24 (15.3)	24 (8.9)	48 (11.2)	1.9 (0.97–3.56)	4.1	0.042
CH	N	***pic***	55 (35.0)	88 (32.5)	143 (33.4)	1.1 (0.72–1.73)	0.3	0.588
RO	S	***sigA***	18 (11.5)	31 (11.4)	49 (11.5)	1.0 (0.50–1.93)	0.0	0.993
MO		***aaiC***	44 (28.0)	97 (35.8)	141 (32.9)	0.7 (0.44–1.09)	2.7	0.099
SO		***air***	41 (26.1)	57 (21.0)	98 (22.9)	1.3 (0.81–2.15)	1.5	0.227
ME		***eilA***	79 (50.3)	128 (47.2)	207 (48.4)	1.1 (0.75–1.71)	0.4	0.538

CI = confidence interval; EAEC = enteroaggregaive *Escherichia coli*; OR = odds ratio.

0.99–2.49, $P = 0.041$) and *pet* (OR: 1.9, 95% CI: 0.97–3.56, $P = 0.042$) genes were higher among diarrhea cases (Table 2).

The distribution of the characterized virulence genes varied across the age strata. In 0- to 11-month stratum, prevalence of *pet* (OR: 6.9, 95% CI: 2.06–29.20, $P < 0.001$), *aggA* (OR: 2.2, 95% CI: 1.16–4.29, $P = 0.008$), and *capU* (OR: 1.9, 95% CI: 1.02–3.51, $P = 0.028$) genes were more common in cases compared with controls (Table 3). Similar higher prevalence pattern was observed for *pet* (OR: 15.0, 95% CI: 1.35–750.0, $P = 0.003$) and *capU* (OR: 4.3, 95% CI: 1.27–18.54, $P = 0.009$) when the virulence factors were characterized among the sole EAEC pathogen from MSD children 0- to 11-month age in cases than controls.

Prevalence of virulence genes that were proportionately higher in cases compared with controls in children 0–11 months were *sepA* (36.5% versus 26.5%), *astA* (54.1% versus 41.7%), *aggR* (71.8% versus 62.1%), *aap* (56.5% versus 44.7%), and ORF3 (75.3% versus 63.6%). The *astA* gene was found more often in cases (67.2%) than in controls (49.5%) in the age stratum 12–23 months (OR: 2.1, 95% CI1.03–4.27, $P = 0.026$); none of the putative virulence factors were found to be significantly more common in MSD children ≥ 2 years of age (Table 3).

In addition to considering each factor individually, we considered the importance of combinations of potential EAEC virulence factors by employing CART analysis. The CART analysis builds a model in stepwise fashion to yield the combination of factors most strongly associated with the queried outcome, in this case the combination of factors most strongly associated with MSD. Each branch of a CART output tree ends in a terminal "node"; each observation falls into exactly one terminal node; and each terminal node is uniquely defined by a set of rules, such as having or not having a certain factor.

We examined all genotypic assays performed: *aatA*, *aggR*, *aaiC*, *aap*, ORF3, *sat*, *sepA*, *pic*, *sigA*, *pet*, *astA*, *aafC*, *agg3/4C*, *aafA*, *agg3A*, *aggA*, *agg4A*, *air*, *capU*, *eilA*, *aar*, as well as considering the collective number of virulence loci present (generating a VFS) (Supplemental Figures 1 and 2).

As noted, prevalence of the virulence were significantly higher in cases compared with controls in children 0–11 months and applying the CART analysis (Supplemental Figure 2) showed that the presence of *pet* (Node 1), regardless of the presence or absence of any other scored genotype among the *pet*-positive strains, provided a strong association with diarrhea. Among the *pet*-negative strains, CART analysis suggested two additional trait clusters that were associated with MSD: Node 2 includes those strains with a VFS ≤ 8 in combination with *sepA*, whereas Node 3 includes a VFS > 8, suggesting a combination of typical EAEC factors in addition to the toxin EAST-1 toxin.

Antibiotic susceptibility testing. The susceptibility patterns of the randomly selected 88 (20%) EAEC strains were similar among cases and controls. The data showed that all the randomly selected 88 (100%) EAEC strains tested were susceptible to ceftriaxone and cefoxitin, over 90% were susceptible to cefotazidime and ciprofloxacin, and more than 80% were susceptible to chloramphenicol and nalidixic acid. However, susceptibility to sulphamethoxazole-trimethoprim and ampicillin was low (15% and 28%, respectively) (Table 4).

DISCUSSION

EAEC is a common cause of diarrhea worldwide.[32] The assessment of the 21 genes in the 428 EAEC strains in this study showed that the frequency of most genes correlated well with similar studies, particularly the study from the GEMS neighboring site, in Bamako Mali.[28] In this study, more than half of participants were younger than 1 year of age, although there were no differences between cases and controls. Particularly, striking was the consistency of the association between SepA and MSD in Mali and this study. SepA is a SPATE protease that was initially found in *Shigella flexneri* strains,[33] but has subsequently been found commonly among EAEC.[34] The protease has been implicated in causing increased inflammation in Shigella strains but it may also have enterotoxic activity.

In this study, the virulence genes *aggA* encoding for AAF/1, capU, and pet, encoding a member of Class 1 SPATEs family, were statistically implicated as genes responsible for EAEC diarrhea in younger children < 12 months.

Our study highlights significant heterogeneity in gene profiles among the EAEC isolates. Of the 21 genes targeted, none of the EAEC isolates characterized genetically harbors more than 15 genes. The heterogeneous nature of EAEC enables it to display variation in causing clinical illness,[32] although factors responsible for its virulence are not well understood.

Several studies have shown possible genes that confer virulence on EAEC.[32,35] Our data show three virulence genes associated with diarrhea in infants. Interestingly, the three incriminated virulence genes are plasmid genes that include *pet*, AAF/1 fimbrial subunit (*aggA*), and hexosyltransferase homolog (*capU*). The Pet toxin is a 108-kDa protease implicated in cytoskeletal changes and epithelial-cell rounding by cleavage of the cytoskeletal protein spectrin.[36–39] In Mexico, the Pet gene was initially detected from EAEC strain 049766 implicated in a highly virulent outbreak of diarrhea in which some infants died.[40] Also, the reported enterotoxic activity of EAEC induced by Pet is consistent with the secretory diarrhea seen in most patients with EAEC enteritis.[41] A recent report from Iran alluded that pet gene is more prevalent among EAEC strains isolated from adult diarrheal patients.[42] Therefore, our findings support the role of Pet gene in EAEC causing diarrhea in infants (Supplemental Figure 1; Table 3). However, earlier EAEC virulence factor study conducted in southwest Nigeria over a decade ago showed that the Pet gene was equally distributed among EAEC strains isolated from children < 5 years with or without diarrheal.[43] Seemingly, our study also showed, no association of Pet with diarrheal disease in the children < 5 years but the effect is only seemed in EAEC strains isolated from children < 1 year and so the differences between our findings could be due to age stratification.

AAF/I was associated with diarrhea in the first year of life. The Shiga toxin producing EAEC strain implicated in the German outbreak expressed AAF/I.[44]

Hexosyltransferase homolog (*capU*), a plasmid-encoded protein was significantly high among the younger children. Its role in EAEC diarrhea is not clearly defined. Notably, the *capU* gene was the third most common gene found (62%) among genes investigated in this study. This probably highlights the importance of genes acting in concert.

astA encodes an EAST1 that is related to the heat-labile enterotoxin of enterotoxigenic *E. coli*. The relevance

TABLE 3

Distribution of EAEC virulence genes in case and control children in three age strata

Virulence genes	0–11 months (*N* = 217)				12–23 months (*N* = 166)				24–59 months (*N* = 45)			
	Case (*N* = 85) no. (%)	Control (*N* = 132) no. (%)	OR (95% CI)	*P* value	Case (*N* = 61) no. (%)	Control (*N* = 105) no. (%)	OR (95% CI)	*P* value	Case (*N* = 11) no. (%)	Control (*N* = 34) no. (%)	OR (95% CI)	*P* value
aatA	33 (38.8)	48 (36.4)	1.1 (0.60–2.02)	0.714	16 (26.2)	26 (24.8)	1.1 (0.48–2.34)	0.833	2 (18.2)	8 (23.5)	0.7 (0.06–4.71)	0.710
aggR	61 (71.8)	82 (62.1)	1.5 (0.82–2.93)	0.143	33 (54.1)	62 (59.1)	0.8 (0.41–1.62)	0.534	3 (27.3)	17 (50.0)	0.4 (0.05–1.95)	0.187
aaP	48 (56.5)	59 (44.7)	1.6 (0.89–2.88)	0.090	28 (45.9)	46 (43.8)	1.1 (0.54–2.15)	0.793	1 (9.1)	16 (47.1)	0.1 (0.00–0.98)	0.024
ORF3	64 (75.3)	84 (63.6)	1.7 (0.91–3.37)	0.071	37 (60.7)	63 (60.0)	1.0 (0.51–2.06)	0.933	5 (45.5)	21 (61.8)	0.5 (0.10–2.53)	0.341
capU	59 (69.4)	72 (54.6)	1.9 (1.02–3.51)	0.028	42 (68.9)	62 (59.1)	1.5 (0.75–3.18)	0.208	7 (63.6)	24 (70.6)	0.7 (0.14–4.20	0.665
aar	62 (72.9)	98 (74.2)	0.9 (0.48–1.82)	0.831	42 (68.9)	70. (66.7)	1.1 (0.53–2.32)	0.772	6 (54.6)	20 (58.8)	0.8 (0.17–4.24)	0.802
aafC	3 (3.5)	7 (5.3)	0.7 (0.10–2.96)	0.543	3 (4.9)	6 (5.7)	0.9 (0.13–4.18)	0.827	1 (9.1)	3 (8.8)	1.0 (0.01–14.6)	0.978
agg3/4C	34 (40.0)	51 (38.6)	1.1 (0.58–1.91)	0.840	20 (32.8)	33 (31.4)	1.1 (0.50–2.19)	0.856	4 (36.4)	12 (35.3)	1.0 (0.18–5.18)	0.948
agg3A	3 (3.5)	18 (13.6)	0.2 (0.04–0.83)	0.014	5 (8.2)	8 (7.6)	1.1 (0.26–3.96)	0.893	2 (18.2)	2 (5.9)	3.6 (0.22–53.6)	0.212
aafA	1 (1.2)	3 (2.3)	0.3 (0.01–6.51)	0.557	2 (3.3)	10 (9.5)	0.3 (0.03–1.59)	0.134	0 (0)	2 (5.9)	0.0 (0.00–16.8)	0.410
aggA	32 (37.7)	28 (21.2)	2.2 (1.16–4.29)	0.008	13 (21.3)	25 (23.8)	0.9 (0.37–1.95)	0.711	2 (18.2)	9 (26.5)	0.6 (0.05–3.94)	0.578
agg4A	10 (11.8)	12 (9.1)	1.3 (0.48–3.55)	0.524	4 (6.6)	2 (1.9)	3.6 (0.49–40.7)	0.121	1 (9.1)	2 (5.9)	1.6 (0.02–33.4)	0.710
astA	46 (54.1)	55 (41.7)	1.7 (0.91–2.96)	0.072	41 (67.2)	52 (49.5)	2.1 (1.03–4.27)	0.026	4 (36.4)	22 (64.7)	0.3 (0.05–1.56)	0.098
sat	20 (23.5)	24 (18.2)	1.4 (0.66–2.84)	0.338	9 (14.8)	24 (22.9)	0.6 (0.22–1.43)	0.207	0 (0)	8 (23.5)	0.0 (0.00–1.69)	0.076
sepA	31 (36.5)	35 (26.5)	1.6 (0.84–2.86)	0.119	16 (26.2)	23 (21.9)	1.3 (0.56–2.79)	0.526	3 (27.3)	4 (11.8)	2.8 (0.33–20.1)	0.217
pet	15 (17.7)	4 (3.0)	6.9 (2.06–29.20)	< 0.001	9 (14.8)	16 (15.2)	1.0 (0.34–2.51)	0.933	0 (0)	4 (11.8)	0.0 (0.00–4.78)	0.233
pic	24 (28.2)	34 (25.8)	1.1 (0.58–2.18)	0.687	28 (45.9)	41 (39.1)	1.3 (0.66–2.63)	0.387	3 (27.3)	13 (38.2)	0.6 (0.08–3.18)	0.509
sigA	4 (4.7)	5 (3.8)	1.3 (0.24–6.01)	0.740	10 (16.4)	20 (19.1)	0.8 (0.32–2.04)	0.668	4 (36.4)	6 (17.7)	2.7 (0.42–15.0)	0.194
aaiC	22 (25.9)	34 (25.8)	1.0 (0.51–1.95)	0.983	19 (31.2)	48 (45.7)	0.5 (0.25–1.09)	0.065	3 (27.3)	15 (44.1)	0.5 (0.07–2.47)	0.321
air	26 (30.6)	34 (25.8)	1.3 (0.66–2.32)	0.437	13 (21.3)	21 (20.0)	1.1 (0.45–2.50)	0.840	2 (18.2)	2 (5.9)	3.6 (0.22–53.6)	0.212
eilA	37 (43.5)	54 (40.9)	1.1 (0.61–2.00)	0.702	35 (57.4)	52 (49.5)	1.4 (0.69–2.72)	0.328	7 (63.6)	22 (64.7)	1.0 (0.19–5.38)	0.948

CI = confidence interval; EAEC = enteroaggregaive *Escherichia coli*; OR = odds ratio.

Table 4
Antimicrobial-resistant pattern of EAEC strains (N = 88)

Antibiotics (μg)	Resistant no. (%)	Susceptible no. (%)
Sulphamethoxazole-trimethoprim (25)	75 (85.0)	13 (15.0)
Ampicillin (10)	63 (72.0)	25 (28.0)
Chloramphenicol (30)	16 (18.2)	72 (81.8)
Nalidixic acid (30)	16 (18.2)	72 (81.8)
Ciprofloxacin (5)	8 (9.0)	80 (91.0)
Cefotazidime (30)	2 (2.3)	86 (97.7)
Ceftriaxone (30)	0 (0)	88 (100)
Cefoxitin (30)	0 (0)	88 (100)

EAEC = enteroaggregaive *Escherichia coli*.

of *astA* gene in EAEC diarrhea has been reported in several studies,[45–48] and EAST1 was found to be associated with diarrhea in combination with other genes in the Mali study.[28] *astA* is not restricted to EAEC but is widely distributed among other enteric pathogens[49,50] as well as commensal *E. coli*.

Globally, EAEC strains have shown a low to high level of resistance to antimicrobial agents.[51] Our data from the antimicrobial susceptibility investigation highlights resistant pattern of the EAEC strains against cotrimoxazole and ampicillin. The first line of antibiotics prescribed for patient management in our region is cotrimoxazole and ampicillin, which explains the high resistance against these antibiotics. An increase in resistance of EAEC strains to chloramphenicol, nalidixic acid, and quinolones was observed in this study compared with a similar study on a member of enterobacteriaceae family from the same region[52] and in eastern Asia.[53] Twenty percent of the EAEC strains tested showed multidrug resistance to three antimicrobial agents, whereas 6% showed resistance to more than three antimicrobial agents. This finding is in contrast to a similar study conducted in India, showing 75% of strains with multidrug resistance, that is, > 3 antimicrobial agents.[54]

The limitations of this study included exclusion of multiple comparisons such as malnutrition and other enteric coinfections. Hence, future studies can consider these essential confounders.

Our study has strengthened the role of *pet* and *aggA* genes of EAEC in the cause of MSD in African infants. The EAEC virulence gene profiles found in this study have also proven the heterogeneity of the genetic component of the EAEC isolates studied. However, further investigations are needed to establish the specific or combination of gene(s) that are associated with EAEC diarrheal in different age strata, particularly children from developing countries. In addition, the pattern of antimicrobial resistance against EAEC is worrisome and needs to be addressed.

Received August 29, 2016. Accepted for publication May 22, 2017.

Published online August 7, 2017.

Note: Supplemental figures appear at www.ajtmh.org.

Acknowledgments: We thank the study participants and their parents, members of field, and laboratory team at Medical Research Council Basse who actively participated in the successful completion of this study. Among these are Henry Badji, Edrisa Sabally, Sampa Sanneh, Omar T. Jallow, and Jarra Manneh.

Financial support: The financial support for this study was from The Medical Research Council, (United Kingdom), and the Bill & Melinda Gates Foundation via Professor James P. Nataro who was previously at CVD-University of Maryland, Baltimore, now at the University of Virginia, Virginia.

Authors' addresses: Usman N. Ikumapayi, Mohammad J. Hossain, Modupeh Betts, Modou Lamin, Brenda Kwambana-Adams, Michel Dione, Anna Roca, and Martin Antonio, Medical Research Council Unit, Banjul, The Gambia, E-mails: uikumapayi@mrc.gm, jhossain@mrc.gm, mbetts@mrc.gm, mlamin@mc.gm, bkwambana@mrc.gm, m.dione@cgiar.org, aroca@mrc.gm, and mantonio@mrc.gm. Nadia Boisen, Department of Microbiological Surveillance and Research, Statens Serum Institut, Copenhagen, Denmark, E-mail: nboboisen@gmail.com. Debasish Saha and Richard A. Adegbola, GlaxoSmithKline Vaccines, Wavre, Belgium, E-mails: debasish.x.saha@gsk.com and richard.a.adegbola@gsk.com. James P. Nataro, Department of Pediatrics, University of Virginia School of Medicine, Charlottesville, VA, E-mail: jpn2r@hscmail.mcc.virginia.edu.

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

REFERENCES

1. Nataro JP, Kaper JB, 1998. Diarrheagenic *Escherichia coli*. *Clin Microbiol Rev 11:* 142–201.
2. Moyo SJ, Maselle SY, Matee MI, Langeland N, Mylvaganam H, 2007. Identification of diarrheagenic *Escherichia coli* isolated from infants and children in Dar es Salaam, Tanzania. *BMC Infect Dis 7:* 92.
3. Cobeljic M, Miljkovic-Selimovic B, Paunovic-Todosijevic D, Velickovic Z, Lepsanovic Z, Zec N, Savic D, Ilic R, Konstantinovic S, Jovanovic B, Kostic V, 1996. Enteroaggregative *Escherichia coli* associated with an outbreak of diarrhoea in a neonatal nursery ward. *Epidemiol Infect 117:* 11–16.
4. Itoh Y, Nagano I, Kunishima M, Ezaki T, 1997. Laboratory investigation of enteroaggregative *Escherichia coli* O untypeable: H10 associated with a massive outbreak of gastrointestinal illness. *J Clin Microbiol 35:* 2546–2550.
5. Pai M, Kang G, Ramakrishna BS, Venkataraman A, Muliyil J, 1997. An epidemic of diarrhoea in south India caused by enteroaggregative *Escherichia coli*. *Indian J Med Res 106:* 7–12.
6. Smith HR, Cheasty T, Rowe B, 1997. Enteroaggregative *Escherichia coli* and outbreaks of gastroenteritis in UK. *Lancet 350:* 814–815.
7. Adachi JA, Jiang ZD, Mathewson JJ, Verenkar MP, Thompson S, Martinez-Sandoval F, Steffen R, Ericsson CD, DuPont HL, 2001. Enteroaggregative *Escherichia coli* as a major etiologic agent in traveler's diarrhea in 3 regions of the world. *Clin Infect Dis 32:* 1706–1709.
8. Januszkiewicz A, Szych J, Rastawicki W, Wolkowicz T, Chrost A, Leszczynska B, Kuzma E, Roszkowska-Blaim M, Gierczynski R, 2012. Molecular epidemiology of a household outbreak of Shiga-toxin-producing *Escherichia coli* in Poland due to secondary transmission of STEC O104:H4 from Germany. *J Med Microbiol 61:* 552–558.
9. Samie A, Obi CL, Dillingham R, Pinkerton RC, Guerrant RL, 2007. Enteroaggregative *Escherichia coli* in Venda, South Africa: distribution of virulence-related genes by multiplex polymerase chain reaction in stool samples of human immunodeficiency virus (HIV)-positive and HIV-negative individuals and primary school children. *Am J Trop Med Hyg 77:* 142–150.
10. Bielaszewska M, Mellmann A, Zhang W, Kock R, Fruth A, Bauwens A, Peters G, Karch H, 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. *Lancet Infect Dis 11:* 671–676.
11. Bhan MK, Raj P, Levine MM, Kaper JB, Bhandari N, Srivastava R, Kumar R, Sazawal S, 1989. Enteroaggregative *Escherichia coli* associated with persistent diarrhea in a cohort of rural children in India. *J Infect Dis 159:* 1061–1064.
12. Paul M, Tsukamoto T, Ghosh AR, Bhattacharya SK, Manna B, Chakrabarti S, Nair GB, Sack DA, Sen D, Takeda Y, 1994. The significance of enteroaggregative *Escherichia coli* in the etiology of hospitalized diarrhoea in Calcutta, India and the

demonstration of a new honey-combed pattern of aggregative adherence. *FEMS Microbiol Lett 117:* 319–325.
13. Harrington SM, Strauman MC, Abe CM, Nataro JP, 2005. Aggregative adherence fimbriae contribute to the inflammatory response of epithelial cells infected with enteroaggregative *Escherichia coli*. *Cell Microbiol 7:* 1565–1578.
14. Harrington SM, Dudley EG, Nataro JP, 2006. Pathogenesis of enteroaggregative *Escherichia coli* infection. *FEMS Microbiol Lett 254:* 12–18.
15. Nataro JP, Martinez J, 1998. Diagnosisand investigation of diarrheagenic *Escherichia coli*. *Methods Mol Med 15:* 387–406.
16. Nataro JP, Yikang D, Yingkang D, Walker K, 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. *J Bacteriol 176:* 4691–4699.
17. Czeczulin JR, Balepur S, Hicks S, Phillips A, Hall R, Kothary MH, Navarro-Garcia F, Nataro JP, 1997. Aggregative adherence fimbria II, a second fimbrial antigen mediating aggregative adherence in enteroaggregative *Escherichia coli*. *Infect Immun 65:* 4135–4145.
18. Jonsson R, Struve C, Boisen N, Mateiu RV, Santiago AE, Jenssen H, Nataro JP, Krogfelt KA, 2015. Novel aggregative adherence fimbria variant of enteroaggregative *Escherichia coli*. *Infect Immun 83:* 1396–1405.
19. Savarino SJ, Fasano A, Robertson DC, Levine MM, 1991. Enteroaggregative *Escherichia coli* elaborate a heat-stable enterotoxin demonstrable in an in vitro rabbit intestinal model. *J Clin Invest 87:* 1450–1455.
20. Nataro JP, Steiner T, Guerrant RL, 1998. Enteroaggregative *Escherichia coli*. *Emerg Infect Dis 4:* 251–261.
21. Henderson IR, Czeczulin J, Eslava C, Noriega F, Nataro JP, 1999. Characterization of pic, a secreted protease of *Shigella flexneri* and enteroaggregative *Escherichia coli*. *Infect Immun 67:* 5587–5596.
22. Kahali S, Sarkar B, Rajendran K, Khanam J, Yamasaki S, Nandy RK, Bhattacharya SK, Ramamurthy T, 2004. Virulence characteristics and molecular epidemiology of enteroaggregative *Escherichia coli* isolates from hospitalized diarrheal patients in Kolkata, India. *J Clin Microbiol 42:* 4111–4120.
23. Steiner TS, Nataro JP, Poteet-Smith CE, Smith JA, Guerrant RL, 2000. Enteroaggregative *Escherichia coli* expresses a novel flagellin that causes IL-8 release from intestinal epithelial cells. *J Clin Invest 105:* 1769–1777.
24. Presterl E, Nadrchal R, Wolf D, Rotter M, Hirschl AM, 1999. Enteroaggregative and enterotoxigenic *Escherichia coli* among isolates from patients with diarrhea in Austria. *Eur J Clin Microbiol Infect Dis 18:* 209–212.
25. Knutton S, Shaw R, Phillips AD, Smith HR, Willshaw GA, Watson P, Price E, 2001. Phenotypic and genetic analysis of diarrhea-associated *Escherichia coli* isolated from children in the United Kingdom. *J Pediatr Gastroenterol Nutr 33:* 32–40.
26. Cohen MB, Nataro JP, Bernstein DI, Hawkins J, Roberts N, Staat MA, 2005. Prevalence of diarrheagenic *Escherichia coli* in acute childhood enteritis: a prospective controlled study. *J Pediatr 146:* 54–61.
27. Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D, Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL, Tamboura B, Sanogo D, Onwuchekwa U, Manna B, Ramamurthy T, Kanungo S, Ochieng JB, Omore R, Oundo JO, Hossain A, Das SK, Ahmed S, Qureshi S, Quadri F, Adegbola RA, Antonio M, Hossain MJ, Akinsola A, Mandomando I, Nhampossa T, Acacio S, Biswas K, O'Reilly CE, Mintz ED, Berkeley LY, Muhsen K, Sommerfelt H, Robins-Browne RM, Levine MM, 2013. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet 382:* 209–222.
28. Boisen N, Scheutz F, Rasko DA, Redman JC, Persson S, Simon J, Kotloff KL, Levine MM, Sow S, Tamboura B, Toure A, Malle D, Panchalingam S, Krogfelt KA, Nataro JP, 2012. Genomic characterization of enteroaggregative *Escherichia coli* from children in Mali. *J Infect Dis 205:* 431–444.
29. Farag TH, Nasrin D, Wu Y, Muhsen K, Blackwelder WC, Sommerfelt H, Panchalingam S, Nataro JP, Kotloff KL, Levine MM, 2012. Some epidemiologic, clinical, microbiologic, and organizational assumptions that influenced the design and performance of the Global Enteric Multicenter Study (GEMS). *Clin Infect Dis 55 (Suppl 4):* S225–S231.
30. Kotloff KL, Blackwelder WC, Nasrin D, Nataro JP, Farag TH, van Eijk A, Adegbola RA, Alonso PL, Breiman RF, Faruque AS, Saha D, Sow SO, Sur D, Zaidi AK, Biswas K, Panchalingam S, Clemens JD, Cohen D, Glass RI, Mintz ED, Sommerfelt H, Levine MM, 2012. The Global Enteric Multicenter Study (GEMS) of diarrheal disease in infants and young children in developing countries: epidemiologic and clinical methods of the case/control study. *Clin Infect Dis 55 (Suppl 4):* S232–S245.
31. Panchalingam S, Antonio M, Hossain A, Mandomando I, Ochieng B, Oundo J, Ramamurthy T, Tamboura B, Zaidi AK, Petri W, Houpt E, Murray P, Prado V, Vidal R, Steele D, Strockbine N, Sansonetti P, Glass RI, Robins-Browne RM, Tauschek M, Svennerholm AM, Berkeley LY, Kotloff K, Levine MM, Nataro JP, 2012. Diagnostic microbiologic methods in the GEMS-1 case/control study. *Clin Infect Dis 55 (Suppl 4):* S294–S302.
32. Cennimo D, Abbas A, Huang DB, Chiang T, 2009. The prevalence and virulence characteristics of enteroaggregative *Escherichia coli* at an urgent-care clinic in the USA: a case-control study. *J Med Microbiol 58:* 403–407.
33. Benjelloun-Touimi Z, Sansonetti PJ, Parsot C, 1995. SepA, the major extracellular protein of *Shigella flexneri*: autonomous secretion and involvement in tissue invasion. *Mol Microbiol 17:* 123–135.
34. Boisen N, Ruiz-Perez F, Scheutz F, Krogfelt KA, Nataro JP, 2009. Short report: high prevalence of serine protease autotransporter cytotoxins among strains of enteroaggregative *Escherichia coli*. *Am J Trop Med Hyg 80:* 294–301.
35. Opintan JA, Bishar RA, Newman MJ, Okeke IN, 2010. Carriage of diarrhoeagenic *Escherichia coli* by older children and adults in Accra, Ghana. *Trans R Soc Trop Med Hyg 104:* 504–506.
36. Navarro-Garcia F, Sears C, Eslava C, Cravioto A, Nataro JP, 1999. Cytoskeletal effects induced by pet, the serine protease enterotoxin of enteroaggregative *Escherichia coli*. *Infect Immun 67:* 2184–2192.
37. Villaseca JM, Navarro-Garcia F, Mendoza-Hernandez G, Nataro JP, Cravioto A, Eslava C, 2000. Pet toxin from enteroaggregative *Escherichia coli* produces cellular damage associated with fodrin disruption. *Infect Immun 68:* 5920–5927.
38. Navarro-Garcia F, Canizalez-Roman A, Luna J, Sears C, Nataro JP, 2001. Plasmid-encoded toxin of enteroaggregative *Escherichia coli* is internalized by epithelial cells. *Infect Immun 69:* 1053–1060.
39. Dutta PR, Cappello R, Navarro-Garcia F, Nataro JP, 2002. Functional comparison of serine protease autotransporters of enterobacteriaceae. *Infect Immun 70:* 7105–7113.
40. Eslava CVJ, Morales R, Navarro A, Cravioto A, 1993. Identification of a protein with toxigenic activity produced by enteroaggregative *Escherichia coli*. *Abstract of the 93rd General Meeting of the American Society for Microbiology*. Washington, DC: American Society for Microbiology, 44.
41. Eslava C, Navarro-Garcia F, Czeczulin JR, Henderson IR, Cravioto A, Nataro JP, 1998. Pet, an autotransporter enterotoxin from enteroaggregative *Escherichia coli*. *Infect Immun 66:* 3155–3163.
42. Bafandeh S, Haghi F, Zeighami H, 2015. Prevalence and virulence characteristics of enteroaggregative *Escherichia coli* in a case-control study among patients from Iran. *J Med Microbiol 64:* 519–524.
43. Okeke IN, Lamikanra A, Czeczulin J, Dubovsky F, Kaper JB, Nataro JP, 2000. Heterogeneous virulence of enteroaggregative *Escherichia coli* strains isolated from children in southwest Nigeria. *J Infect Dis 181:* 252–260.
44. Scheutz F, Nielsen EM, Frimodt-Moller J, Boisen N, Morabito S, Tozzoli R, Nataro JP, Caprioli A, 2011. Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104:H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. *Euro Surveill 16:* pii: 19889.
45. Vila J, Gene A, Vargas M, Gascon J, Latorre C, Jimenez de Anta MT, 1998. A case-control study of diarrhoea in children caused by *Escherichia coli* producing heat-stable enterotoxin (EAST-1). *J Med Microbiol 47:* 889–891.

46. Paiva de Sousa C, Dubreuil JD, 2001. Distribution and expression of the astA gene (EAST1 toxin) in *Escherichia coli* and *Salmonella*. *Int J Med Microbiol 291:* 15–20.
47. Toshima H, Uenaka E, Bi Y, Nakamura H, Ogasawara J, Hase A, Kamata Y, Nishikawa Y, 2004. Detection and isolation of *Escherichia coli* with a coding gene for enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 from food and comparison with fecal isolates. *J Food Prot 67:* 2117–2122.
48. Veilleux S, Holt N, Schultz BD, Dubreuil JD, 2008. *Escherichia coli* EAST1 toxin toxicity of variants 17-2 and O 42. *Comp Immunol Microbiol Infect Dis 31:* 567–578.
49. Menard LP, Dubreuil JD, 2002. Enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 (EAST1): a new toxin with an old twist. *Crit Rev Microbiol 28:* 43–60.
50. Zhou Z, Ogasawara J, Nishikawa Y, Seto Y, Helander A, Hase A, Iritani N, Nakamura H, Arikawa K, Kai A, Kamata Y, Hoshi H, Haruki K, 2002. An outbreak of gastroenteritis in Osaka, Japan due to *Escherichia coli* serogroup O166:H15 that had a coding gene for enteroaggregative *E. coli* heat-stable enterotoxin 1 (EAST1). *Epidemiol Infect 128:* 363–371.
51. Mendez Arancibia E, Pitart C, Ruiz J, Marco F, Gascon J, Vila J, 2009. Evolution of antimicrobial resistance in enteroaggregative *Escherichia coli* and enterotoxigenic *Escherichia coli* causing traveller's diarrhoea. *J Antimicrob Chemother 64:* 343–347.
52. Ikumapayi UN, Antonio M, Sonne-Hansen J, Biney E, Enwere G, Okoko B, Oluwalana C, Vaughan A, Zaman SM, Greenwood BM, Cutts FT, Adegbola RA, 2007. Molecular epidemiology of community-acquired invasive non-typhoidal *Salmonella* among children aged 2 29 months in rural Gambia and discovery of a new serovar, *Salmonella enterica* Dingiri. *J Med Microbiol 56:* 1479–1484.
53. Aslani MM, Alikhani MY, Zavari A, Yousefi R, Zamani AR, 2011. Characterization of enteroaggregative *Escherichia coli* (EAEC) clinical isolates and their antibiotic resistance pattern. *Int J Infect Dis 15:* e136–e139.
54. Raju B, Ballal M, 2009. Multidrug resistant enteroaggregative *Escherichia coli* diarrhoea in rural southern Indian population. *Scand J Infect Dis 41:* 105–108.