The Role of the Gastrointestinal Tract Microbiota in Colonisation Resistance Against Common Enteric Pathogens in Malawian Children

Thesis submitted in accordance with the requirements of the University of Liverpool and the University of Malawi for the degree of

Doctor of Philosophy

By

Angeziwa Chunga

February 2022

DECLARATION OF WORK DONE

I declare that the material contained in “The Role Of Gastrointestinal Tract Microbiota in Colonisation Resistance Against Commonly Isolated Enteric Pathogens in Malawian Children” is my own work. In some cases, the work was done with the help of others. Details of the work and responsibilities are as indicated in **Table 1**. A. Chunga has been highlighted to indicate the works and responsibilities done by A. Chunga. The work presented in this thesis was conducted at the University of Liverpool (UoL), Institute of Infection and Global Health, Kamuzu University of Health Sciences (KUHeS), formerly known as University of Malawi (UNIMA), College of Medicine, Malawi Liverpool Wellcome Trust Clinical Research Programme (MLW), Quadram Institute Biosciences (QIB), Wellcome Trust Sanger Institute (WTSI), Ndirande Health Centre (NHC) and Zingwangwa Health Centre (ZHC). All the material in this thesis has not been presented, nor is currently being presented, wholly or in part, for any qualification.

Table 1: Attribution of work

Activity	Thesis reference	Site	Responsibility
Study designs and protocols	Chapter 3 and 4 Chapter 5 and 6		M. Gordon, T. Nyirenda, M. Iturriza-Gomara, A. Kamng'ona and **A. Chunga** L. Hall, M. Gordon, T. Nyirenda, M. Iturriza-Gomara, A. Kamng'ona and **A. Chunga**
Ethics application	Chapter 3 and 4 Chapter 5 and 6	UNIMA UNIMA and UoL	T. Nyirenda **A. Chunga**
Consent and recruitment of study participants	Chapter 3 and 4 Chapter 5 and 6	ZHC NHC	R. Nkhata R. Nkhata
Sample collection	Chapter 3 and 4 Chapter 5 and 6	ZHC NHC	R. Nkhata R. Nkhata
DNA extraction	Chapter 3 Chapter 4 Chapter 5	MLW MLW MLW & QIB	**A. Chunga** N. Jambo & T. Nyirenda **A. Chunga**
Taqman Array card assays	Chapter 3	MLW	**A. Chunga**
Stool culture	Chapter 3 Chapter 5	MLW MLW, UoL & QIB	MLW core laboratory **A. Chunga** H.C. Banda
Competition assays	Chapter 6	MLW, UoL & QIB	**A. Chunga**
Library preparation	Chapter 4	WTSI	WTSI

	Chapter 5	QIB & WTSI	**A. Chunga** & WTSI
16S rRNA sequencing	Chapter 4 Chapter 5	WTSI WTSI	WTSI WTSI
Whole-genome sequencing	Chapter 5	WTSI	WTSI
Sanger sequencing	Chapter 5	Eurofins	Eurofins
QIIME2 microbiome data analysis	Chapter 4 Chapter 5	MLW MLW & UoL	**A. Chunga** **A. Chunga**
WGS data analysis	Chapter 5	MLW & UoL	**A. Chunga**
Statistical analysis	Chapter 3 Chapter 4 Chapter 5 Chapter 6	MLW MLW MLW MLW	**A. Chunga** **A. Chunga** **A. Chunga** **A. Chunga**

ABSTRACT

Introduction: Gut microbiota disruption during early life has immediate and long-term impact on host health. Early feeding habits and pathogen exposure are important factors that can cause gut microbiota disruption. This study aimed to describe the role of gut microbiota in controlling enteric pathogens that Malawian children are exposed to, and to characterise and explore anti-infective properties of *Bifidobacterium* isolated from Malawian infants against *Salmonella Typhimurium.*

Methods: The study used an enteric Taqman Array Card to detect pathogens that Malawian children aged 6 to 18 months are exposed to, 16S rRNA sequencing for microbiota profiling, whole-genome sequencing to characterise *Bifidobacterium* and in vitro competitive assay to explore anti-infective properties of the Malawi *Bifidobacterium* against an invasive *Salmonella typhimurium* strain.

Results: The study has demonstrated that healthy Malawian children are exposed to multiple enteric pathogens that are mostly not associated with clinical symptoms. Apart from *Giardia*, EAEC and *B. fragilis*, most of these pathogens do not affect gut microbiota composition. In addition, the data demonstrate that *Bifidobacterium* is predominant in Malawian children, especially those that are exclusively breastfeeding. Malawian breastfeeding infants are predominantly colonised by *Bifidobacterium longum* possessing a wide range of human milk oligosaccharide digesting genes. Exploratory data show that Malawian *Bifidobacterium* strains possess anti-infective properties against *Salmonella Typhimurium.*

Conclusion: Malawian children are asymptomatically exposed to multiple enteric pathogens that may affect gut microbiota composition. *Bifidobacterium* from Malawian breastfeeding infants possess HMOs that may be important for colonisation resistance against important enteric pathogens.

ACKNOWLEDGEMENTS

With great pleasure, I would like to thank my supervisors, Prof Melita Gordon, Prof Miren Iturriza-Gomara, Prof Julian Parkhill Dr Tonney Nyirenda, Dr Arox Kamng'ona and Dr Lindsay Hall, for the tremendous support and guidance that they continuously provided throughout my study period. I would also like to thank Dr Benjamin Kumwenda, Dr Anmol Kiran, Dr Edward Parker and Dr Phillip Ashton for always being available and willing to help.

I am very grateful to the Malawi Liverpool Wellcome Trust Clinical Research Program – Training grants for awarding me a PhD grant to conduct this study. Commonwealth Scholarship Commission awarded me a PhD split-site Scholarship, which allowed me to spend one year of my study period in the United Kingdom and conduct most of my experimental work. These experiments would otherwise not have been possible in Malawi. I am also very thankful to STRATAA, and TyVAC studies for the remarkable support accorded to me throughout my study period.

I am indebted to several institutions that hosted and allowed me to carry out my research work. University of Malawi College of Medicine sponsored my research projects and hosted me as I carried out microbiology work. The University of Liverpool, Institute of Infection and Global Health, Miren Iturriza-Gomara and Jay Hinton research groups for hosting me during my one year stay in the UK, giving me laboratory space and some reagent for bifidobacteria culture work. Wellcome Trust sanger Institute through Julian Parkhill research group for sequencing my microbiome and whole-genome sequences and introducing me to microbiome data analysis. The Quadram Institute, through the Halls lab, who introduced me to *Bifidobacterium,* trained and allowed me to carry out most of the work on *Bifidobacterium*.

I would not have done this without the support of family. Chawanangwa Mahebere Chirambo, my lovely sons; Benjamin, Luthando and Rutendo, who always believed in me – this encouraged and helped me to keep moving forward. Last but not least, I would like to give Glory to the Almighty God through Jesus Christ, my Lord, for starting and finishing this good work through me, for nothing is possible without Him.

Thesis dedicated to Malawian girls, specifically teen mother

To the Malawian girls and teen mothers that I know

To the Malawian girls and teen mothers that I will know

To the Malawian girls and teen mothers that will know me

To the Malawian girls and teen mothers that I have met

To the Malawian girls and teen mothers that I will meet

To the Malawian girls and teen mothers that will meet my story

To the Malawian girls and teen mothers that face challenges everyday

To the Malawian girls and teen mothers that face inequalities everyday

To the Malawian girls and teen mothers that are not safe in the hands of their teachers

To the Malawian girls and teen mothers that are not safe in the hands of their religious leaders

To the Malawian girls and teen mothers that are not safe in the hands of their family

To the Malawian girls and teen mothers that have been failed and have failed themselves

Many times, we will fail and fall but many times plus one we will rise-again

Table of Contents

DECLARATION OF WORK DONE i

ABSTRACT iv

ACKNOWLEDGEMENTS v

LIST OF FIGURES xi

LIST OF TABLES xiv

ABBREVIATIONS xvi

CHAPTER 1: INTRODUCTION 18

1.1 The gut microbiota *18*

1.1.1 General description of the gut microbiota composition *18*
1.1.2 Factors contributing to advancements in gut microbiota research 22
1.1.3 Challenges in gut microbiota research 27

1.2 Main components of the gut microbiota *29*
1.2.1 Seeding of the gut microbiota and *Bifidobacterium* dominance 32
1.2.2 History and biology of *Bifidobacterium* 34
1.2.3 *Bifidobacterium* association with health and disease 34

1.3 Functions of the gut microbiota *37*
1.3.1 Physiological 37
1.3.2 Metabolic 38
1.3.3 Immunological 40
1.3.4 Colonisation resistance 41

1.4 Factors that influence gut microbiota composition and function *48*
1.4.1 Mothers health status before and during pregnancy 49
1.4.2 Host genetics 51
1.4.3 Gestation age; term vs preterm births 52
1.4.4 Delivery mode; vaginal vs caesarian birth 53
1.4.5 Early feeding habits; breastfeeding vs formula feeding 55
1.4.7 Nutritional status 57
1.4.8 Environmental conditions 58
1.4.9 Infection and disease 60
1.4.10 Medical interventions 61

1.5 The impact of gut microbiota - studies in human health *65*

1.6 Enteric infections *66*
1.6.1 The importance of enteric pathogens 66
1.6.2 Diarrhoeal infections 68
1.6.3 Bloodstream infection 69
1.6.4 Asymptomatic infection 71
1.6.5 Important Bacterial pathogens - Salmonella 72
1.6.6 Important viral pathogens 78
1.6.7 Important parasitic enteric infections 79

1.7 Aims and Objectives of the thesis *80*
1.7.1 Main aims 80
1.7.2 Specific objectives 80

Chapter 2: Materials and Methods 81

2.1 Study type and site ... *81*

2.2 Ethical approval ... *81*

2.3 Clinical procedures for SalExpo study ... *84*

2.3.1 Participant recruitment, follow-up, sample collection and processing for Salmonella asymptomatic children from 6 to 18 months ... 84

2.4 Clinical procedure for BIFBAC study ... *85*

2.4.1 Consenting and recruitment of healthy exclusively breastfed children ... 85

2.4.2 Stool sample collection ... 85

2.5 Laboratory procedures for SalExpo study ... *86*

2.5.1 Stool Total Nucleic Acid extraction for Taqman Array Card Assay ... 87

2.5.2 Enteric Taqman Array Card assay ... 89

2.5.3 Salmonella detection using stool culture and Real-Time PCR ... 91

2.5.4 Sample processing for 16S rRNA sequencing ... 92

2.6 Laboratory procedures for BIFBAC study ... *92*

2.6.1 Isolation of bifidobacteria ... 93

2.6.2 Bifidobacteria biochemical confirmatory test: Fructose-6-phosphate Phosphoketolase assay 94

2.6.3 Bifidobacteria confirmatory test: 16S rRNA Sanger sequencing ... 95

2.6.4 Bacteria DNA extraction for whole genome sequencing ... 96

2.6.5 Stool DNA extraction for microbiome work ... 98

2.6.6 Library preparation for 16S rRNA sequencing ... 98

2.7 Salmonella and Bifidobacterium Competition assay ... *104*

2.7.1 Screening the growth of *Bifidobacterium* isolates and *S.*Typhimurium D23580 ... 105

2.7.2 Co-culturing of *S.* Typhimurium and *Bifidobacterium* ... 106

2.7.3 Growth of *S.* Typhimurium in *Bifidobacterium* supernatant ... 108

2.8 Microbiome data analysis ... *109*

2.8.1 Data importing and demultiplexing ... 109

2.8.2 Quality checking, filtering and sequence variant calling ... 110

2.8.3 Taxonomy assignment ... 110

2.8.4 Diversity measurements ... 111

2.9 Whole-genome sequence data analysis ... *111*

2.9.1 Assembly and annotation ... 111

2.9.2 Genomic identification of Malawi reads ... 112

2.9.3 Identification of human milk oligosaccharide digesting genes ... 112

Chapter 3: Longitudinal Enteric Pathogen Exposure in Healthy 6 to 18 months old Malawian children ... 117

3.1 Introduction ... *117*

3.1.1 Hypothesis ... 119

3.1.2 Study objectives ... 119

3.2 Methods ... *120*

3.2.1 Participant recruitment, follow-up, sample collection and processing ... 120

3.2.2 Detection of multiple enteric pathogens using Taqman Array Card Assay ... 120

3.2.3 Data management and statistical analysis ... 120

3.2.4 Ethical approval ... 121

3.3 Results ... *122*

3.3.1 Participant clinical, demographic and social economic characteristics ... 122

3.3.2 Malawian children aged between 6 to 18 months were exposed to a wide range of enteric pathogens ... 124

3.3.3 Bacteria pathogen exposure ... 126

3.3.4 Viral pathogen exposure ... 132

3.3.4.1 Association between viral pathogen exposure, clinical symptoms and social economic characteristics 133
3.3.5 Protozoa and Helminths exposure 134
3.3.5.1 Association between parasite exposure, clinical symptoms and social economic characteristics 136
3.3.6 : Co-detection of multiple pathogens in symptomatic and asymptomatic children 138

3.4 Discussion 141

3.5 Conclusion 146

Chapter 4: Gastrointestinal tract microbiota changes Associated with Enteric Pathogen exposure in Malawian children 147

4.1 Introduction 147
4.1.1 Hypothesis 148
4.1.2 Objectives 148

4.2 Methods 149
4.2.1 Study type and study site 149
4.2.2 Ethical approval 149
4.2.3 Enteric pathogen detection 149
4.2.4 DNA extraction, library preparation and sequencing for microbiota profiling 150
4.2.5 Microbiota sequenced data analysis 150
4.2.6 Data management and statistical analysis 150

4.3 Results 152
4.3.1 General description of 16S rRNA sequence reads 152
4.3.2 Quality control assessment 154
4.3.3 General description of the gut microbiota of Malawian children from 6 to 18 months of age 154
4.3.4 Gut microbiota composition with clinical characteristic 162
4.3.5 Gut microbiota composition with water source and boiled drinking water 163
4.3.6 Gut microbiota composition with Bacterial pathogen exposure 164
4.3.7 Differences in gut microbiota composition with viral pathogen exposure 167
4.3.8 Differences in gut microbiota composition with parasitic pathogen exposure 168

4.4 Discussion 171
4.4.1 Strengths and weaknesses 176

4.5 Conclusion 179

Chapter 5: Gut Microbiota Profiling of Healthy Exclusively Breastfed Malawian Children and the Genomic Characterisation of Bifidobacterium Isolated from their Stool 180

5.1 Introduction 180
5.1.1 Hypothesis 181
5.1.2 Study objectives 182

5.2 Methods 183
5.2.1 Study type and site 183
5.2.2 Consenting and recruitment of healthy exclusively breastfed children 183
5.2.3 Stool sample collection 183
5.2.4 Isolation and confirmation of Bifidobacterium 183
5.2.5 Bacteria DNA extraction and sequencing 184
5.2.6 Microbiome data analysis – 16S rRNA sequence data 184
5.2.7 Whole genome sequenced data 185
5.2.8 Statistical analysis 185
5.2.9 Ethical approval 186

5.3 Results 187
5.3.1 Characteristics of study participants 187

5.3.2 Quantitative analysis of the gut microbial community in exclusively breastfed Malawian children 188
5.3.3 Identification and general features of the Malawi Bifidobacterium genomes 193
5.3.4 General features of *Bifidobacterium* longum genomes 196
5.3.5 Characterising the human milk oligosaccharide digesting gene repertoire in genomes from exclusively breastfed Malawian infants 197

5.4 Discussion 202
5.4.1 Strengths and weaknesses 206
5.4.2 Future direction 207

5.5 Conclusion 208

Chapter 6: Exploiting Anti-infection Properties of the Early Life Microbiota Member Bifidobacterium Against Invasive Salmonella Typhimurium – An Exploratory Study 209

6.1 Introduction 209
6.1.2 Overall objective 211
6.1.3 Specific objectives 211

6.2 Methods 212
6.2.1 Bacterial strains used 212
6.2.2 Culture media for *Bifidobacterium* and *S*. Typhimurium growth 212
6.2.3 Co-culturing of Bifidobacterium and S. Typhimurium 213
6.2.4 Growth of S. Typhimurium in Bifidobacterium cell-free supernatant 213

6.3 Results 214
6.3.1 Growth patterns of *Bifidobacterium* and *Salmonella* in a nutrient-rich media and chemically defined media 214
6.3.2 The growth of *Salmonella* Typhimurium when co-cultured with *Bifidobacterium* 217
6.3.3 *S.* Typhimurium growth in *Bifidobacterium* supernatant 219

6.4 Discussion 221
6.4.1 Strengths and limitations 225
6.4.2 Experiments that will been conducted in Malawi to validate and extend these exploratory findings 226

6.5 Summary 232

CHAPTER 7: GENERAL DISCUSSION 233

7.1 Introduction 233

7.2 Discussion of key findings 234
7.2.1 Malawian children are asymptomatically exposed to multiple enteric pathogens 234
7.2.2 Enteric pathogen exposure was not significantly associated with gut microbiota changes 235
7.2.3 Exclusively breastfeeding Malawian infants have a high abundance of *Bifidobacterium* possessing a wide range of HMO digesting genes 236
7.2.4 *Bifidobacterium* found in breastfeeding Malawian infants seem to possess antibacterial properties against the invasive S. Typhimurium, D23580 236

7.3 Study limitations 237

7.4 Recommendations for further research 238
7.4.1 Recommendations for further research on enteric pathogen exposure and changes in gut microbiota profiles 238

7.5 Concluding remarks 239

LIST OF FIGURES

FIGURE	TITLE	PG
1.1	Schematic representation of bacterial composition along the gastrointestinal tract	21
1.2	Timeline of the advancements in microbiome research	24
1.3	16S rRNA gene	26
1.4	Gut microbiota composition throughout the human life	31
1.5	Hypothetical figure of the proportion of Bifidobacterium in the gut over time and proportions and maturation of immune cells.	36
1.6	Functions of the gut microbiota.	38
1.7	Direct colonisation resistance mechanisms for a healthy and impaired gut	42
1.8	Factors that affect gut microbiota composition.	49
1.9	Differences in gut microbiota composition with mode of delivery	54
1.10	Differences in gut microbiota composition by geography.	60
1.11	All causes of death in SSA for under five-year-old children.	69
1.12	Estimated incidence of pathogens causing bloodstream infections in Blantyre, Malawi	70
1.13	Potential benefits and harms of asymptomatic pathogen exposure	72
1.14	S.Typhimurium specific immune development with age.	74
2.1	Recruitment and sample collection for SALEXPO study	84
2.2	Flow chart of activities for the longitudinal cohort study.	86
2.3	Layout of the enteric Taqman Array Card used for the detection of different enteric pathogens.	89
2.4	Flow chart of activities for the BIFBAC project	93
2.5	Quality score of the forward and reverse reads for SalExpo study	110
3.1	Proportions of detected pathogen groups	124
3.2	Proportion of the detected group enteric pathogens.	125
3.3	Proportion of pathogens detected in stool samples from Malawian children from 6 to 18 months of age.	125
3.4 A and B	Proportions of detected Bacterial pathogens and Proportions of detected *E.coli*	126
3.5	Proportion of bacterial pathogen exposure from 6 months to 18 months.	127

3.6	Proportions of viral pathogens detected from 6 to 18 months old children.	132
3.7	Viral pathogen exposure patterns over time	133
3.8	Proportions of protozoa detection from children aged 6 to 18 months old	135
3.9	Longitudinal protozoa exposure patterns form the 2 commonly detected protozoa.	136
3.10	Number of samples with co-detected pathogens	139
4.1	Number of sequence reads.	153
4.2	Distribution of sequencing depth.	153
4.3	Relative abundance of the gut microbiota at the genus level.	155
4.4	Relative abundance of the gut microbiota at the genus level.	156
4.5	Distribution of alpha diversity metrics.	157
4.6	Alpha diversity by age group.	159
4.7	Beta diversity by age in month.	160
4.8	Alpha diversity with gender.	161
4.9	Beta diversity by age in months.	162
4.10	Alpha and beta diversity with current diarrhoea	163
4.11	Alpha and beta diversity with current vomiting	163
4.12	Alpha diversity with water source.	164
4.13	Alpha diversity with EAEC exposure.	166
4.14	Alpha diversity with B. fragilis exposure.	167
4.15	Alpha diversity with sapovirus exposure. No significant differences were detected.	168
4.16	Alpha diversity with Giardia exposure.	169
4.17	Beta diversity with Giardia exposure.	170
5.1	Gut microbiota composition at phylum level	189
5.2a	Alpha diversity by sex	190
5.2b	Alpha diversity by Bifidobacterium culture.	191
5.3a	Beta diversity by Sex - weighted Bray Curtis	192
5.3.b	Beta diversity – weighted by Bifidobacterium Culture Bray Curtis.	193
5.4	A dendrogram showing the phylogenetic distribution of Bifidobacterium isolated from exclusively breastfed Malawian children.	195
5.5	Phylogenetic Distribution of global and Malawi Bifidobacterium longum genomes.	197
5.6	A dendrogram with the number of HMO genes present in Malawi Bifidobacterium genomes.	199
5.7	HMO digesting genes present among Bifidobacterium species isolated from exclusively breastfeeding Malawian infants.	201
5.8	HMO digesting genes present in the Malawi B. longum clusters.	201

6.1	Anaerobic growth of Bifidobacterium in BHI and CDRIM culture media at different time points.	215
6.2	Growth patterns of Bifidobacterium in BHI and CDRIM culture media	216
6.3	Growth of S. Typhimurium in CDRIM and BHI using both aerobic and anaerobic conditions.	217
6.4	S. Typhimurium growth when co-cultured with Bifidobacterium	218
6.5	Impact of pH on Salmonella growth when S. Typhimurium and Bifidobacterium are co-cultured.	219
6.6	S. Typhimurium, D23580 growth in 1 in 2 serially diluted Bifidobacterium supernatant.	210

LIST OF TABLES

TABLE	TITLE	PG
1	List of Abbreviations	xvi
1.1	Definition of terms commonly used in gut microbiota studies	20
1.2	Comparison between 16S rRNA and Shotgun metagenomic sequencing	25
1.3	Direct colonisation resistance mechanisms: healthy and impaired gut	43
2.1	Basic study information for SalExpo and BIFBAC studies	83
2.2	TAC PCR conditions	91
2.3	Primers and probes used for molecular detection of Salmonella in stool specimens	92
2.4	Forward and reverse primers for V1V2 16S rRNA gene	99
2.5	Preparation of PCR master mix	100
2.6	Media preparation for co-culture experiments	108
2.7	Details of laboratory materials and reagent used for SalExpo study	113
2.8	Details of laboratory materials and reagents used for BIFBAC study	114
3.1	Differences in anthropometric measurements for male and female participants	123
3.2	Table 3.2: Number of clinical episodes and demographic characteristics at 6, 12 and 18 months of age	123
3.3	Association between EPEC and clinical symptoms and social economic status	129
3.4	Association between other bacteria; Campylobacter and B. fragilis, and clinical features and social economic status	131
3.5	Association between sapovirus and clinical symptoms and social economic status	134
3.6	Association between parasite exposure and clinical symptoms and social economic status	137
3.7	Correlation between number of co-detection and clinical and demographic characteristics	140
4.1	The number of sequence reads for controls used in the study	154

4.2	Distribution of Shannon diversities showing the median and standard deviation of the diversity values	158
5.1	General participant clinical characteristics	187
5.2	Identity and genomic characteristics of the Malawi Bifidobacterium	194
5.3	Number of HMO metabolising genes present in Malawi Bifidobacterium genomes	200
6.1	Bifidobacterium isolates used for different exploratory experiment	212

ABBREVIATIONS

All abbreviations are defined the first time they are being mentioned. The list of all abbreviations used in this thesis are presented in Table 1.

Table 1: List of Abbreviations

Abbreviation	Full name
BIFBAC	Short study name for "Exploiting anti-infection activities of the early life microbiota member *Bifidobacterium* against bacterial enteric infections" study.
SalExpo	Short study name for "*Salmonella* exposure and development of specific immunity in Malawian children" study.
CDC	Centre for disease control
CFU	Colony-forming units
CO_2	Carbon dioxide
CoM	College of Medicine
COMREC	College of Medicine Research and Ethics Committee
CR	Colonisation resistance
CDRIM	Chemically defined rich media
DNA	Dioxyribonucleotide
HMO	Human Milk Oligosaccharides
HMP	Human microbiome project
LB	Luria-Bertani
LPS	Lipopolysaccharide
MDR	Multidrug-resistant
mg	microgram
ml	microlitre
ml	millilitres
MLST	Multi-locus sequence typing
MLW	Malawi Liverpool Wellcome Trust Clinical Research Programme
NTS	Nontyphoidal *Salmonella*

O_2	Oxygen
^oC	Degrees Celsius
P-value	Probability value
PBS	Phosphate buffered saline
QECH	Queen Elizabeth Central Hospital
QIB	Quadrum Institute Biosciences
RNA	Ribonucleic acid
Rpm	Revolutions per minute
rRNA	Ribosomal RNA
S. Enteritidis	*Salmonella* enterica serovar Enteritidis
S. Typhi	*Salmonella* enterica serovar Typhi
S. Typhimurium	*Salmonella* enterica serovar Typhimurium
SCV	*Salmonella* containing vacuole
SPI	*Salmonella* pathogenicity island
SPI-1	*Salmonella* pathogenicity island 1
SPI-2	*Salmonella* pathogenicity island 2
UNIMA	University of Malawi
UoL	University of Liverpool
WGS	Whole-genome Sequence
WHO	World Health Organisation
GIT	Gastrointestinal tract
ASV	Amplicon Sequence Variants
OTU	Operational Taxonomic Unit
MUAC	Mid-upper arm circumference
E. coli	*Escherichia coli*
EAEC	Enteroaggregative *E. coli*
EPEC	Enteropathogenic *E. coli*
ETEC	Enterotoxigenic *E. coli*
EIEC	Enteroinvasive *E. coli*
B. fragilis	*Bacteroides fragilis*
GDP	Gross domestic product

CHAPTER 1: INTRODUCTION

1.1 The gut microbiota

1.1.1 General description of the gut microbiota composition

Different human anatomical environments such as the skin, oral cavity, gastrointestinal tract (GIT/ gut), and vagina are each colonized by an array of microorganisms consisting of bacteria, archaea, fungi, virus and phages (Rinninella et al., 2019; Ursell, Metcalf, Parfrey, & Knight, 2012). These microorganisms interact with each other as well as the human host. Collectively, this microbial community is called the microbiota, and their collective genetic material is called the microbiome (Costello et al., 2009; Dong et al., 2018; Y. Fan & Pedersen, 2021; Rogers & Zhang, 2016; The Human Microbiome Project, 2012; L. K. Ursell et al., 2012; Y.E., Fischbach, & Belkaid, 2018). **Table 1.1** has defined commonly used terms in gut microbiota studies.

The microbiota composition at each specific site is directly dependent on the environmental conditions associated with that particular site (Ding & Schloss, 2014; Jun, Si, Pugatch, & Scott, 2018; Koziolek et al., 2015). Conditions such as temperature, pH, moisture content, and nutrient availability empower specific microorganisms to thrive and establish themselves in specific body sites. Of all the different body sites, the GIT comprises the largest microbial population of the human microbiota (Sai Manasa Jandhyala 2015, (Cho & Blaser, 2012; Costello et al., 2009). Of all the different microorganisms that make up the gut microbiota, bacteria dominate in numbers and functions. It is estimated that the gut contains between 10^{12} and 10^{14} bacteria with more than 1000 species. Some have estimated that the

gut microbiome contains ten times more bacterial cells than human cells. It consists of a hundred times more genes than those found in the human genome (Cho & Blaser, 2012; Gill et al., 2006; Milani et al., 2017; Sender, Fuchs, & Milo, 2016). However, recent estimates indicate that the ratio of bacterial cells to human cells is closer to 1:1 (Sender et al., 2016).

Table 1.1: Definition of terms commonly used in gut microbiota studies

Term	Definition	Reference
Gut microbiota	A collection of microorganisms consisting of bacteria, archaea, fungi, virus and phages resident in the human gut.	(Rinninella et al., 2019)
Gut microbiome	Is the aggregate of the gut microbiota and their genomic elements	(Rinninella et al., 2019)
Dysbiosis	Changes in gut microbiota composition that results in a reduction in the optimal function of the microbiota	(Rinninella et al., 2019)
Alpha diversity	Measure of microbiota diversity within a sample	(Laudadio et al., 2018)
Beta diversity	Measure of similarity or dissimilarity of the microbiota between populations	(Laudadio et al., 2018)
Richness	The number of microbiota components within a sample	(Laudadio et al., 2018)
Eveness	The relative representation of gut microbiota species within a sample	(Laudadio et al., 2018)
Probiotics	Dietary supplements containing live microbial species that can transiently colonise the gut and confer beneficial influence on the host	(C. Hill et al., 2014)
Prebiotics	A substrate that is selectively utilized by host microorganisms conferring health benefits.	(Gibson et al., 2017)

Bacteria in the gut do not just make up most of the microbial biomass, but they are also responsible for most of the metabolic activities in the human gut (Costello et al., 2009; Rowland et al., 2018). Most of the gut microbiota are anaerobic bacteria and mainly belong to *Firmicutes*, *Bacteroidetes*, *Actinobacteria,* and *Proteobacteria* phyla. *Verrucomicrobia*, *Cyanobacteria*, *Synergistetes,* and *Fusobacteria* are also

found in the gut but at minimal levels (Milani et al., 2017; Thursby & Juge, 2017). Different microbial populations colonize different parts of the GIT. The density and diversity of the gut microbiome also differ along and across the GIT. As shown in **Figure 1.1**, along the GIT, most of the bacteria reside in the large intestines (Mailhe et al., 2018). Across the GIT, the gut lumen and outer mucus layer have the most complex and diverse microbiota (Desai et al., 2016; Donaldson, Lee, & Mazmanian, 2016).

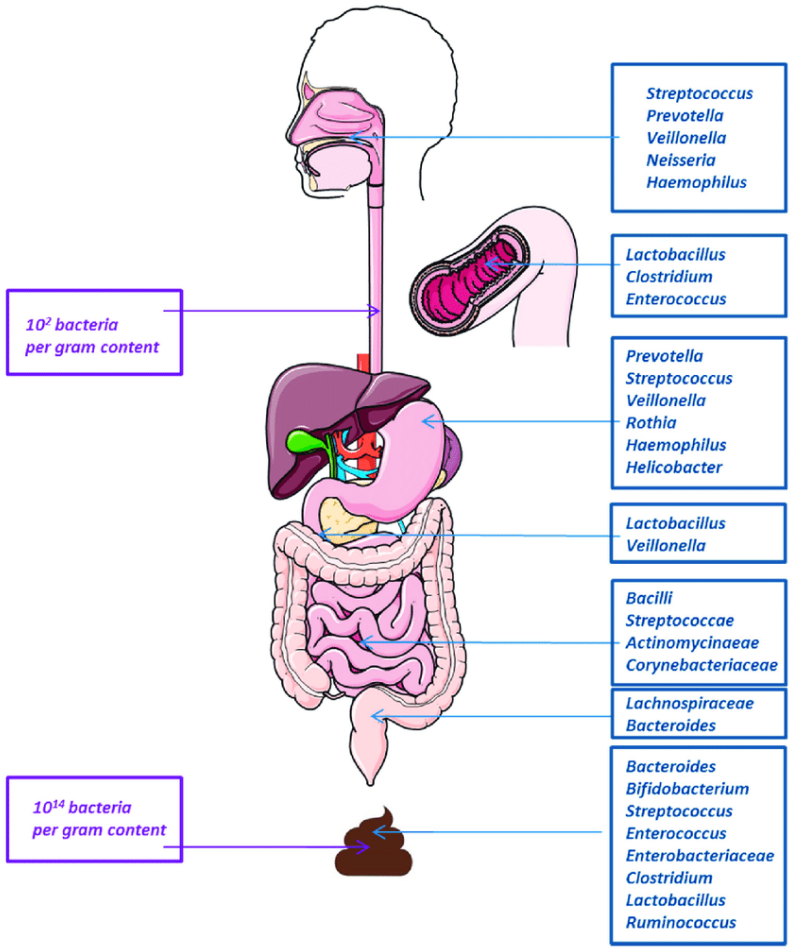

Figure 1.1: Schematic representation of bacterial composition along the gastrointestinal tract

Schematic representation of bacterial composition along the gastrointestinal tract with diversity and richness increasing along the tract. Adapted from (Pichon & Burucoa, 2019).

1.1.2 Factors contributing to advancements in gut microbiota research

In recent years, there has been a significant increase in microbiome research, with a significant proportion of studies focusing on the intestinal microbiome. Advancements in the field of microbiology, molecular biology, immunology, genomics and computational biology are also main contributing factors to the increase in studies of the gut microbiota (Sanschagrin & Yergeau, 2014; Schriefer et al., 2018). Advances in all these fields of study have also contributed to an increase and broadening of ways to understand how microbiota affects human health. Studies on microbial communities inhabiting the gut started when researchers were only using culture-based techniques. Then, the microbiota was only known to consist of about 400 different obligate anaerobic bacteria (A. Lee, Gordon, & Dubos, 1968; W. E. Moore & Holdeman, 1974). More than 90% of organisms in different environments including the gut are however unculturable (Streit & Schmitz, 2004). From the 1990s, technological advancement in the field of microbiology and molecular biology introduced culture-independent techniques that have significantly spearheaded our understanding of the gut microbiota and its functions (Fraher, O'Toole, & Quigley, 2012; 2010; Schriefer et al., 2018; Watts et al., 2017).

Advances in microbiology, complemented by innovations in sequencing technologies have made genomic sequencing to be fast, high throughput and affordable. Sequencing technologies have transitioned from first generation (Sanger) sequencing followed by second/ next-generation sequencing (Roche/454 pyrosequencing, genetic analyser2, Illumina/Solexa sequencing) and now the third generation sequencing (Pacific Biosciences (PacBio) Single-Molecule Real-Time sequencing (SMRT), Oxford Nanopore sequencing platforms and Illumina Tru-seq

Synthetic Long Read sequencing technology (van Dijk, Jaszczyszyn, Naquin, & Thermes, 2018). **Figure 1.2** shows timeline of advancements made in sequencing technologies that has promoted microbiome studies and some of the important pioneering microbiome project

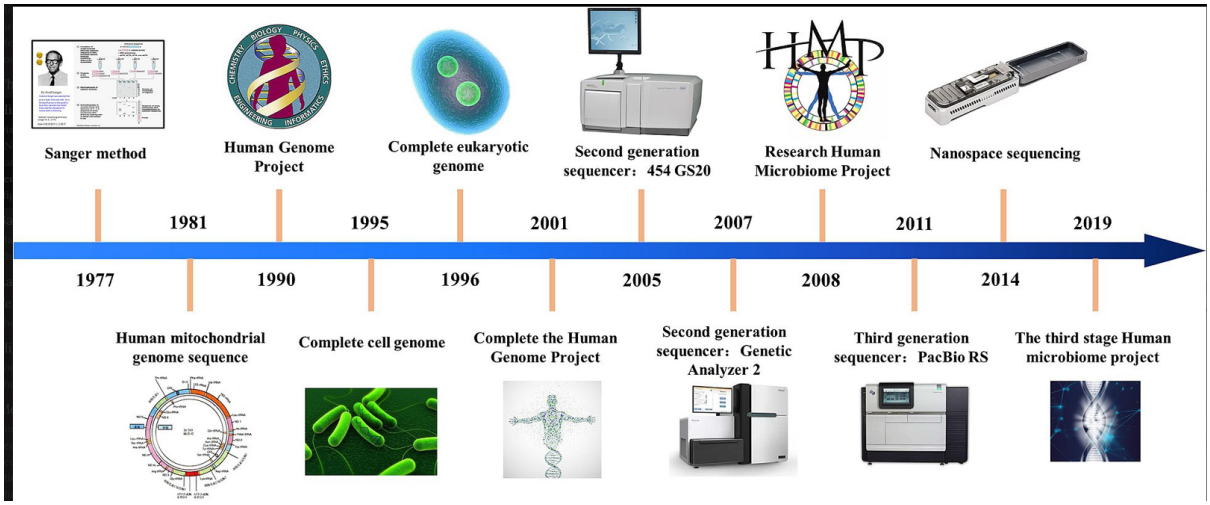

Figure 1.2: Timeline of the advancements in microbiome research

Timeline of the introduction of different sequencing technologies, major sequencing projects and important achievement in genomic sequencing (Yang et al., 2020).

Advances made in sequencing technologies have allowed researchers to conduct metagenomic studies either through amplicon-based sequencing such as 16S rRNA sequencing for bacterial organisms, or whole-genome shotgun metagenomic sequencing. Each of these two metagenomic sequencing methods have strengths and weaknesses that should be taken into account when planning a microbiome study (**Table 1.2**) (Ranjan, Rani, Metwally, McGee, & Perkins, 2016).

Table 1.2: Comparison between 16S rRNA and Shotgun metagenomic sequencing

	16S rRNA Sequencing	**Shotgun metagenomic sequencing**
Sequence region	Specific regions of the 16S rRNA gene	Whole genome sequencing of all organisms
Target organism	Bacteria	Bacteria, virus, fungi, archaea
Commonly used sequencing platform	Illumina, oxford nanopore	PacBio
Sequencing depth	Lower	High
Advantage	• Cheaper • Data relatively easy to analyse • Low biomass samples can be used	• Taxonomic analysis to species or strain level is possible • Taxonomic profiling and functional annotation of the metagenomes • Deeper characterization of organism • Uncultured microbial genomes can also be annotated
Disadvantage	• Biases can arise from primers and during PCR • Analysis limited to taxonomic profiling • Taxonomic assessment sensitive only up to genus level • Characterisation is not very deep • Amplification of different regions of the 16S rRNA gene produces different classification results • False positives are common in low biomass samples	• Expensive • Data analysis is complicated and time consuming • Contamination by host DNA

16S rRNA is a constituent of the prokaryotic ribosome that is 1500 base pairs (bp) long (**Figure 1.3**). It consists of conserved regions across different bacteria species and nine variable regions that allow for the identification of bacteria at genus. Data

generated using 16S rRNA sequencing helps in understanding the abundance and diversity of the microbiota (Han et al., 2020; Laudadio et al., 2018; Schriefer et al., 2018). Shotgun metagenomic sequencing is a whole genome-based sequencing method that helps in understanding the taxonomic composition of bacteria, archae and viruses. In addition to taxonomic identification, shotgun metagenomic sequencing helps in understanding the function of the genes present in the microbiota (Han et al., 2020; Laudadio et al., 2018; Panek et al., 2018; Ranjan et al., 2016).

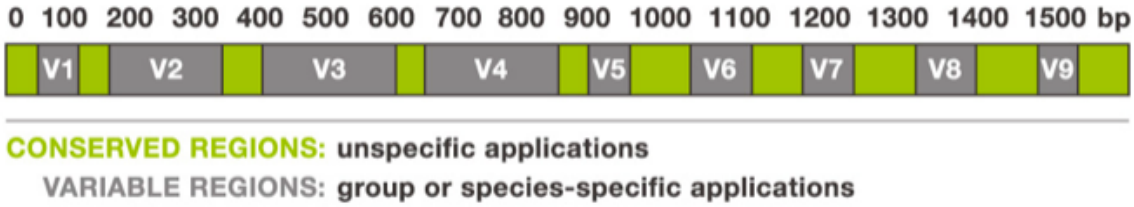

Figure 1.3: 16S rRNA gene

16S rRNA gene showing conserved regions in green and variable regions in grey. Adapted from Stack Exchange (Adapted from https://biology.stackexchange.com/questions/54823/what-causes-the-variable-conserved-structure-in-the-16s-rrna-gene

The increase in metagenomic data generated presented a data analysis challenge. Powerful and reliable computers are now available that allow computationally intensive work, like the analysis of metagenomic data do be done in a fast and efficient manner. The use of computer servers has allowed researcher using low local computing power computers to remotely connect to computers with high computing power in any part of the world. As long as a job has been sent to the supercomputer or server, it will be done and completed regardless of internet connection. Internet connection is very expensive and unreliable in developing countries like Malawi.

Besides the introduction of supercomputers, advancements in bioinformatics have been made to allow for a quick and efficient analysis of the massive data typically generated from metagenomic studies. User friendly analysis programs and pipelines have been developed that allow biologist who are not computational to pre-process, analyse, evaluate and visualize the data. Commonly used data analysis pipelines for 16S rRNA sequence data are QIIME, MOTHUR, DADA2 and UPARSE. For species level shotgun metagenomics data analysis, MetaPhlAn2, Kraken and MG-RAST are commonly used (Niu et al., 2018).

1.1.3 Challenges in gut microbiota research

Gut microbiota research is, however, not without challenges. Human faecal material is used to generate information on gut microbiota composition. This may not be the best representation of the gut content but is used as a proxy. Using mouse models, scientist have shown that microbiota composition differs in the different segments of the GIT (Klymiuk et al., 2021; D. Liu et al., 2021; Lkhagva et al., 2021). The bacteria community in faeces may be affected by sample collection, sample handling, storage and DNA extraction methods (Dave, Higgins, Middha, & Rioux, 2012; Panek et al., 2018). However, several studies have been done that have helped identify the right procedures to minimize the effect of these factors on gut microbiota composition. Protocol standardisation within and between studies is one way of ensuring reproducible and comparable results. The microbial community in faeces might also not represent very well the interaction between the organisms in the gut microbiota and human cells in the microenvironment close to or adjacent to the gut epithelium. The gut microbiota being dynamic and at the centre of a complex ecosystem

influenced by host and external environmental factors poses another challenge. This complexity poses a considerable challenge in clearly understanding the mechanism behind the different roles that the gut microbiota plays. Multi-Omic approaches to understanding the gut microbiome are helping in addressing this challenge. Genomic, proteomic, transcriptomic or epigenomic data is combined with microbiome data during analysis to better understand microbiome functions. There are also high interindividual and intraindividual variations in the gut microbiota composition at any particular stage of life (L. K. Ursell et al., 2012). Because of this variation, a core healthy microbiome, which could have been used as a standard, has not been correctly described. A microbiota mainly composed of beneficial bacterial species such as those belonging to the Bacteroidetes and Actinobacteria phylum and less of potentially disease-causing organisms called pathobionts is considered a healthy gut microbiota (Gill et al., 2006; Rinninella et al., 2019). The description of a healthy microbiome is, however, likely to differ with geographic region. This difference would to a greater extent be attributed to environmental factors which subsequently have an effect on diet and pathogen exposure. These factors have been shown to significantly influence the gut microbiota composition. The pursuit of identifying a healthy gut microbiota is therefore supposed to be specific to a particular region or population.

In addition to the mentioned challenges, conducting microbiome studies in a LMIC like Malawi has its own challenges. It is costly to conduct studies that involve sequencing facilities because either the sequencing capacity is not readily available, or because of limitations in data handling. Samples, mostly DNA sample, are therefore shipped to collaborators in developed countries. Not being able to

sequence samples locally makes sequencing/microbiome study very time-consuming (Chaguza, Nyaga, Mwenda, Esona, & Jere, 2020). Proper resourcing, support and planning is necessary to overcome this challenge. For a start, new and simpler sequencing platforms such as the oxford nanopore could be used. Using the Oxford nanopore technology, scientist at MLW in Malawi are now able to do long read sequencing for Covid-19 and Salmonella. The challenge remains in cases where science-based informed decisions are urgently needed, such as during epidemics. The Covid-19 pandemic has demonstrated how important sequencing is in informing decisions and directing policy. In addition to sequencing samples to understand pathogens, microbiome studies are critical in understanding host-pathogen interaction.

1.2 Main components of the gut microbiota

Age is one of the main factors that influence the composition of the gut microbiota. Other factors that influence gut microbiota composition are described in **section 1.4.** The bacterial composition of the gut microbiota between adults and children is similar in that the main components (making up to 90% of the total bacteria) at phyla level are *Actinobacteria*, *Firmicutes*, *Proteobacteria* and *Bacteroides* (Y. Chen, Zhou, & Wang, 2021; Lloyd-Price, Abu-Ali, & Huttenhower, 2016; Proctor, 2016). The main age related differences are however related to the proportions of taxonomic composition belonging to each of these phyla, taxonomic diversity, and functions.

The gut microbiota changes with age in terms of its richness and diversity. Studies have indicated differences in the richness and diversity of gut microbiota among infants, toddlers, adults and the elderly (O'Toole & Claesson, 2010; Toshitaka et al.,

2016). It has been believed for a long time that colonisation of the gut happens at birth when vaginal and environmental microorganisms first colonise a sterile fetus coming from a sterile environment. A study conducted by Aagaard and colleagues demonstrated that the placenta has a microbiome that is mainly comprised of a low abundance of species that are mostly non-pathogenic. At the phyla level, the placental microbiome was composed of *Firmicutes, Tenericutes, Proteobacteria, Bacteroidetes, Fusobacteria* phyla, and *E. coli* was dominant at species level (Aagaard et al., 2014). The meconium has also been shown to have a low-level species diversity (Gosalbes et al., 2013; Hansen et al., 2015; Koleva, Kim, Scott, & Kozyrskyj, 2015) and enriched with *Proteobacteria* (Hu et al., 2013). However, the implications of the gut microbiota composition at this stage of life on the host's health status are currently not clear.

In terms of taxonomic composition, the gut microbiota of infants and children is mainly composed of Actinobacteria (Selma-Royo et al., 2019). This contrasts with adults who are dominated by Firmicutes and Bacteroides (Rinninella et al., 2019). **Figure 1.4** shows the gut microbiota composition at different stages of life with different factors that may also affect gut microbiota composition at that particular life stage. In terms of diversity, adults have a more diverse gut microbiota than infants and children. Several studies have shown that this age-dependent change occurs until about three years of age when the child’s gut is mature and more adult-like (Beller et al., 2021; Rinninella et al., 2019).

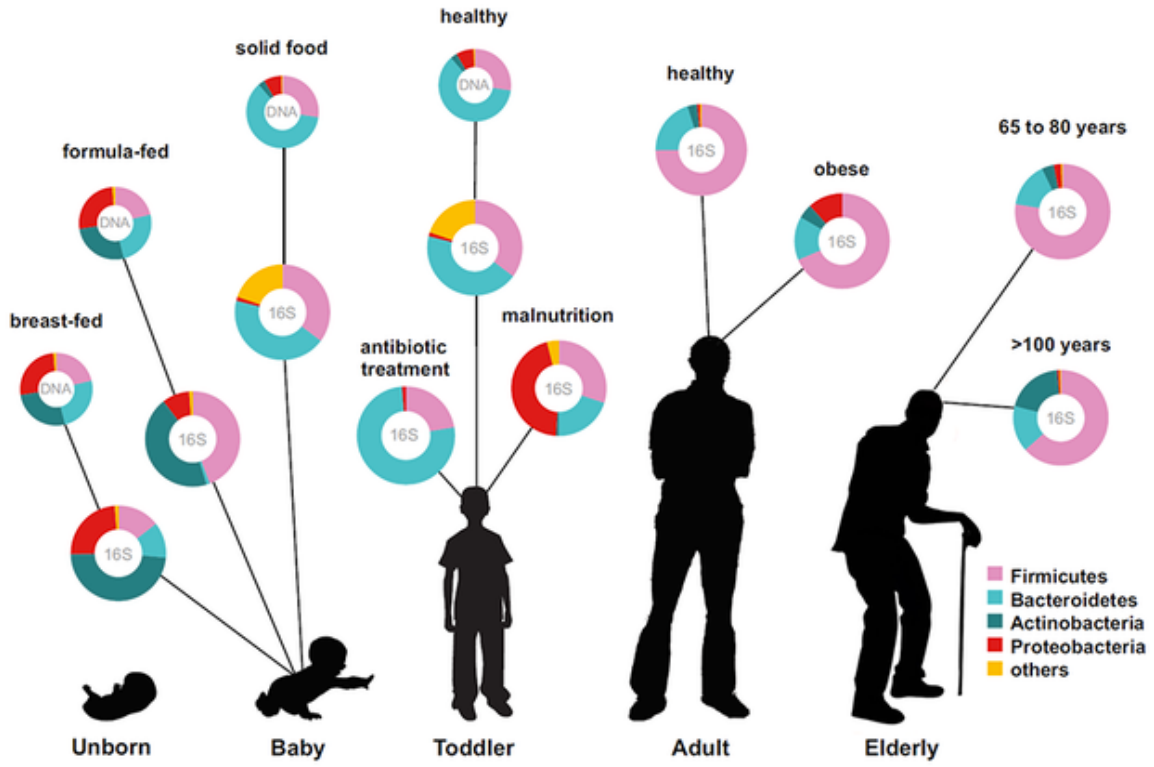

Figure 1.4: Gut microbiota composition throughout the human life

Gut microbiota composition throughout the human life – adapted from (Ottman, Smidt, de Vos, & Belzer, 2012). Gut microbiota composition at the genus level is shown for the unborn, baby, toddler, adult and elderly of life. 16S – data based on 16S rRNA sequencing, DNA – data based on metagenomic sequencing.

After birth, microbial colonisation of the gut proceeds steadily with an increase in gut microbiota diversity during and after weaning, until about three years, when it is very rich and diverse (Rinninella et al., 2019; Selma-Royo et al., 2019). In the very young, the protective effect of the immature gut microbiota is minimal. At this point, functions such as those on the immune system are not fully primed. There are not enough bacteria to optimally cover the gut mucosa and protect the underlying cells against invading pathogenic organisms. Transplacentally transferred antibodies, breastmilk and beneficial gut microbiota members present promotes infant health.

The high abundance of the beneficial bacteria, *Bifidobacterium*, helps offset the impact of an overall low gut microbiota composition and diversity. This study focuses on understanding the gut microbiota of young Malawian children. Bifidobacterium is the main gut microbiota component in young children. The following section therefore focused on understanding Bifidobacterium as the main gut microbiota component in the studied population.

1.2.1 Seeding of the gut microbiota and *Bifidobacterium* dominance

Significant colonisation of the gut by microorganisms mainly starts during child delivery. The gut is colonized by microorganisms from the surrounding environment, mainly from the mother in vaginally born children and the mothers' skin and surrounding environment in caesarian born children (Reyman et al., 2019). The abundance and diversity of the gut microbiota during early life is however very low compared to adults. During early life, the gut microbiota is dominated by the phyla Actinobacteria, especially in term children who are vaginally born and breastfeeding (F. Turroni et al., 2018). The remaining small proportion of the gut microbiota comprises the other phyla: *Proteobacteria*, *Firmicutes* and *Bacteroides*. With time, primarily in about a year, the relative abundance of these other phyla increases as the child's gut matures and becomes more adult-like (Rinninella et al., 2019). At birth, the child is exposed to the mother's gut microbiota or microorganisms in the surrounding environment (Reyman et al., 2019). Because the adult gut is dominated by *Firmicutes*, *Bacteroides* and *Proteobacteria*, the infant gut microbiota is dominated by bacteria from these phyla in the very first few days of life. *Firmicutes* and *Proteobacteria* are in very low abundance, but they are the dominant microbiota members at this time. *Actinobacteria* and *Bacteroides* abundance increase in about

seven days after birth, with *Actinobacteria* being the most dominant phyla (Selma-Royo et al., 2019). In infants born by C-section, bacteria belonging to the *Firmicutes* and *Proteobacteria* phyla remains in high abundance even after the first few weeks of life. *Clostridium sensu stricto 1* and *C. difficile* belonging to the *Firmicutes* phyla have been shown to be frequent colonizers of the gut of C-section born children in early life (Reyman et al., 2019; Werlang et al., 2018).

Among the first colonizers of a child's gut are anaerobic bacteria belonging to the *Actinobacteria* family and genus *Bifidobacterium* (Milani et al., 2017; Thomson, Medina, & Garrido, 2018). The source of *Bifidobacterium* in an infant's gut has been shown to be transferred from the mother through contact with faecal matter during vaginal birth and from breastmilk (Grönlund et al., 2007; Makino, 2018). The presence of bifidobacteria in an infant's gut is promoted by the presence of Human Milk Oligosaccharides (HMO) that cannot be digested by the body's digestive enzymes but are digested by *Bifidobacterium* (Murphy et al., 2017; Sirilun et al., 2015). In the presence of HMO, bifidobacteria that have been transferred to the infant gut rapidly multiply and become the main constituent of a child's gut microbiota (Katayama, 2016; Thomson et al., 2018). *Bifidobacterium* composition in children is thus influenced by the mode of delivery and feeding habits. Caesarian born and formula-fed children tend to have less *Bifidobacterium* than vaginally born and exclusively breastfed children, respectively (Grönlund et al., 2007; Werlang et al., 2018). *Bifidobacterium* is the most predominant member in exclusively breastfed children, accounting for up to 90% of the gut microbiota composition.

1.2.2 History and biology of *Bifidobacterium*

Bifidobacteria are Gram-positive, mostly branched nonmotile anaerobic bacteria that are found in humans, other mammals and social insects. In humans, bifidobacteria are not only found in the gut but also in the mouth and vagina. It was first isolated in the year 1899 by Henri Tissier of the Pasteur Institute in Paris. He isolated *Bifidobacterium* from stool samples collected from a breastfed infant and called it bifidus because of bifid (Y) shape (Tissier, 1900). From the time of Tissier, about 60 species and sub-species have been isolated, characterized and included in the genus *Bifidobacterium*. *Bifidobacterium* are anaerobic and grow well on media having different types of carbohydrates. They utilize a unique Fructose-6-Phospate phosphoketolase pathway to metabolise different types of carbohydrates. Although most *Bifidobacterium* are hypersensitive to Oxygen and are thus mostly strictly anaerobic, some are classified as Oxygen sensitive, tolerant and microaerophilic. The genome size of bifidobacteria ranges from 2.0 to 2.8 Megabase pairs (Mbp). The *Bifidobacterium* genome comprises a significant proportion of genes specializing in carbohydrate metabolism. In breastfeeding infants, most of these carbohydrate digesting genes are for the digestion of HMO while in young children and adults, these genes are responsible for the digestion of plant-based carbohydrates (Kujawska et al., 2020).

1.2.3 *Bifidobacterium* association with health and disease

From the time they were discovered, many studies have been conducted to investigate the role of bifidobacteria in metabolism, nutrition, immunity and control of pathogens. *Bifidobacterium* has been shown to be beneficial to the human host, especially infants and children. The presence of *Bifidobacterium* in the infant's gut is

associated with a reduction in the incidence and severity of diarrhoea infections and allergies (Enomoto et al., 2014), inhibition of pathogenic viruses and bacteria (Aw & Fukuda, 2019), promotion of barrier function of the intestinal epithelial cells (Ishizuka et al., 2009), development and maturation of the immune system and provide nutrients to the host by breaking down the non-digestible carbohydrates including the HMO (Alessandri, Ossiprandi, MacSharry, van Sinderen, & Ventura, 2019; Fanning, Hall, & van Sinderen, 2012; Groeger et al., 2013). A decrease in *Bifidobacterium* abundance in the gut seems to correlate with the maturity of immune cells and an increase in the amount of memory T cells **(Figure 1.5).** The high abundance of bifidobacteria when the immune system is not mature seems to be very important in controlling enteric infections in the very young (O'Neill et al., 2017). *Bifidobacterium* abundance seems to wane once the immune system is primed. These health-promoting properties of *Bifidobacterium* have been widely explored in the developed world and have resulted in using some of the species as probiotics that are administered to at-risk populations such as preterm infants. Probiotics are live microorganisms that, when administered in adequate amounts, confer health benefits to the host (Colin Hill et al., 2014; C. Hill et al., 2014). Probiotics aim to promote a healthy balance of gut microbiota composition, which subsequently help to ameliorate some clinical anomalies. There is a huge body of evidence that supports the use of probiotics in controlling acute infectious diarrhoea, antibiotic associated diarrhoeal and gastrointestinal diseases such as *C. difficile*, ulcerative colitis, and necrotizing enterocolitis (Robertson et al., 2020). Probiotics are however not effective in controlling or preventing acute pancreatitis and crohn's disease. However, B*ifidobacterium* and its health-promoting properties have not been explored in Malawi and most of the SSA countries where the burden of enteric

infections causing either mild gastroenteritis or invasive disease and at-risk populations such as preterm births are very high.

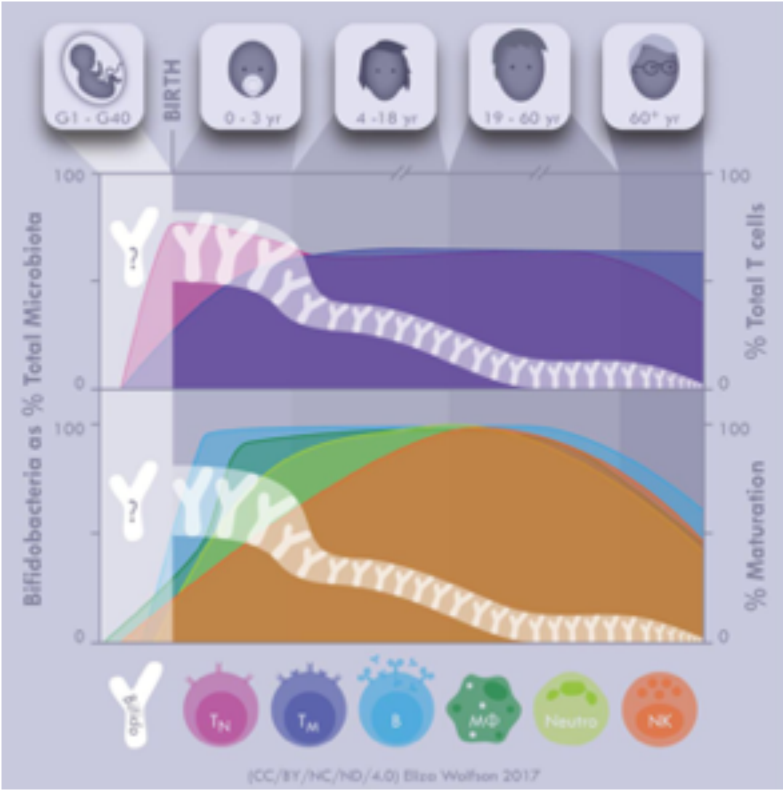

Figure 1.5: Hypothetical figure of the proportion of Bifidobacterium in the gut over time and proportions and maturation of immune cells.

Hypothetical figure of the proportion of *Bifidobacterium* in the gut over time and proportions and maturation of immune cells – Adapted from (O'Neill, Schofield, & Hall, 2017). The proportion of total main immune cells and their maturity period is shown in relation to the relative abundance of *Bifidobacterium* in the gut from in-utero to old age. Y shape is for *Bifidobacterium.* TN – naïve T cells, TM – mature T cells, B – B cells, MF - macrophages, Neutro - neutrophils, NK – Natural Killer

1.3 Functions of the gut microbiota

The roles that the gut microbiota plays in homeostasis and health are continuously developing as more insights are gained through research. A healthy gut microbiota that is continuously interacting with the host plays important physiological, metabolic, immunological and colonisation resistance functions (**Figure 1.6**).

1.3.1 Physiological

The gut microbiota has been implicated in the development, normal function, and maintenance of different human body systems. For example, physiological functions played by the gut microbiota include protecting the epithelium against physical injury (Rakoff-Nahoum, Paglino, Eslami-Varzaneh, Edberg, & Medzhitov, 2004), fat storage (Bäckhed et al., 2004), modulation of bone biomass (Bäckhed et al., 2004; Yan & Charles, 2017), induction of mucosal tolerance, angiogenesis (Andriessen et al., 2016; Kåhrström, 2012), development of lymphatic tissues (Kamada & Nunez, 2013) and modification of the nervous system (Dinan & Cryan, 2017; Graf et al., 2015; H. X. Wang & Wang, 2016).

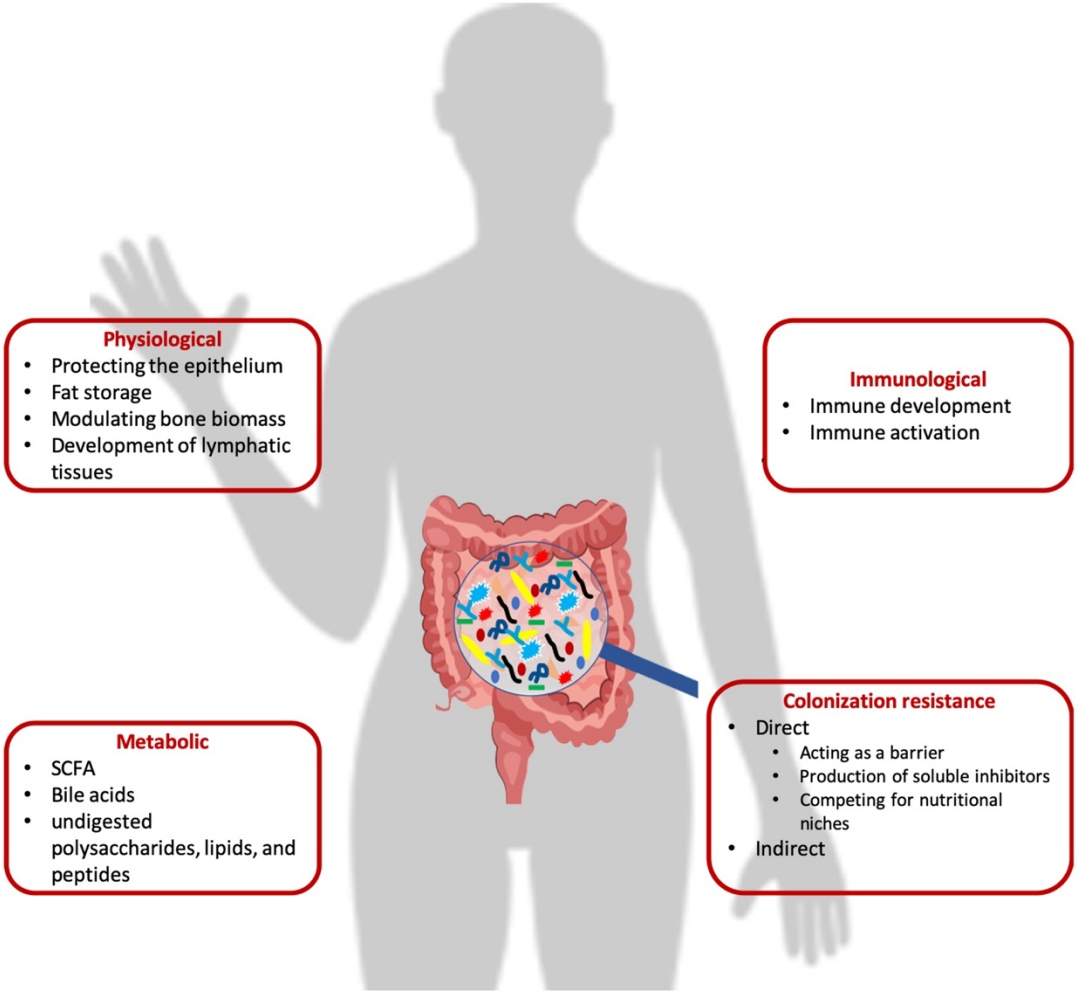

Figure 1.6: Functions of the gut microbiota.

Functions of the gut microbiota. They are categorized into four themes: physiological, immunological, metabolic and colonisation resistance. macrophages, Neutro - neutrophils, NK – Natural Killer cells.

1.3.2 Metabolic

Digestion and metabolism of food involve the use of enzymes secreted by the gut mucosa and the liver. Gut microbiota has a distinctive but extensive metabolic repertoire which complements the liver and gut mucosa in food digestion and metabolism. Microorganisms in the gut produce enzymes used in digestion and metabolism processes that are not encoded by the mammalian genome. These enzymes are used to synthesise vitamins, carbohydrates (polysaccharides) and

polyphenols (Kaoutari, Armougom, Gordon, Raoult, & Henrissat, 2013; Rowland et al., 2018). In the large intestine, bacteria ferment undigested food material from the upper digestive system to produce short-chain fatty acids (SCFAs) and gases. The three main SCFAs produced are acetate, propionate and butyrate. Short Chain Fatty Acids are the main energy source for human colonocytes and helps maintain gut barrier integrity (Antonini, Lo Conte, Sorini, & Falcone, 2019). Besides being an energy source, some SCFAs such as Butyrate, have anti-cancer properties that induce apoptosis among colon cancer cells (Rowland et al., 2018). In addition to benefiting from macronutrients such as SCFAs, the human host cannot synthesise all types of micronutrient vitamins from the ingested food material. Fortunately, gut bacteria can carry out this function in a symbiotic relationship (Hill, 1997; LeBlanc et al., 2013). Vitamins are involved in the synthesis and metabolism of glucose, protein and fatty acids, energy production, biochemical reactions in the body, nervous system function and DNA synthesis, and also in growth and development (Hercberg et al., 1998).

Therefore, the gut microbiota plays a critical role in determining the macronutrient and micronutrient components that end up being used by the host cells. This role impacts the nutritional status of the host and, subsequently, their health and disease state. The different processes and pathways used by the gut microbiota or specific components of the gut microbiota in metabolism and how they affect host health are now an important subject of research. Studies have provided associations between gut microbiota mediated digestion and metabolism and the aetiology and development of some important gastrointestinal diseases and conditions such as

obesity and hypertension (Baothman, Zamzami, Taher, Abubaker, & Abu-Farha, 2016; Tidjani Alou, Lagier, & Raoult, 2016; Wolf & Lorenz, 2012).

1.3.3 Immunological

A symbiotic relationship shaped by evolution has been described between the gut microbiota and the immune system. The immune system plays important roles in the maturation and regulation of gut microbiota composition, while the gut microbiota plays a role in the development, maturation and function of the immune system. Studies using germ-free mice demonstrated that without the microbiota, intestinal mucosal immunity was underdeveloped (Hapfelmeier et al., 2010; Pope, Tomkovich, & Jobin, 2020), the sizes of mesenteric lymph nodes and Peyer's patches was reduced, and the number of immune cells such as CD4+ LP T-cells, IgA-producing plasma cells and intraepithelial alpha-beta T-cell receptor CD8+ cells was reduced (Östman, Rask, Wold, Hultkrantz, & Telemo, 2006). This resulted in a weakened immune system. A study done by Francois Legoux and colleagues demonstrated that the thymic development of MAIT cells is controlled by metabolites such as 5-OP-RU, which are produced by the gut microbiota (Legoux et al., 2019). These MAIT cells were absent in germ-free mice.

Gut microbiota has also been shown to stimulate gut humoral immunity (IgA), enhanced macrophage activity and phagocytosis and augmentation of antigen-specific immune responses (Belkaid & Hand, 2014; Hooper, Stappenbeck, Hong, & Gordon, 2003). In addition, studies conducted in germ-free mice demonstrate that the gut microbiota can induce local intestinal immunity through their effect on antigen-presenting cells (APC), toll-like receptors (TLR) and by promoting splenic

CD4+T cells and systemic antibody responses (Belkaid & Hand, 2014; Kamada & Nunez, 2014).

1.3.4 Colonisation resistance

One well-known function of a healthy gut microbiota is that it provides conditions that disfavour pathogen colonisation of the gut lumen and controls the over-growth of potentially pathogenic gut microbiota members also known as pathobionts. This phenomenon is called colonisation resistance (CR). A vigilant immune system and CR play important roles in tolerating commensal organisms but prohibiting pathogenic organisms from invasion through the epithelium into the gut wall, promote tolerance of an already established pathogen or clearance of pathogens in the gut lumen, and also control the transition of the pathogens to the bloodstream or tissues (Kamada, Chen, Inohara, & Nunez, 2013; Lawley & Walker, 2013; Palleja et al., 2018; Sassone-Corsi & Raffatellu, 2015). Colonisation resistance mechanisms employed by the gut microbiota can either be direct or indirect (Sassone-Corsi & Raffatellu, 2015). Immature or disrupted gut microbiota, due to factors such as antibiotic intake, may not efficiently exhibit this protective effect. Directly, the gut microbiota resist pathogen colonisation by inhibiting the attachment of the pathogenic organism to the gut mucosal by acting as a barrier against invading pathogens, production of soluble inhibitors, competing for nutritional niches. Indirectly, the microbiota resists colonisation by stimulating immune defence mechanisms specifically targeted to the pathogen (Kamada, Chen, et al., 2013; Lawley & Walker, 2013). Despite the well-programmed CR mechanisms, pathogens are still able to colonize the gut and cause infections. Infections happen when the multidimensional interaction involving the microbiota, the host and the pathogen is

altered in favour of the pathogen. A schematic diagram of a healthy and impaired gut is given in **Figure 1.7** with descriptions in **Table 1.3.**

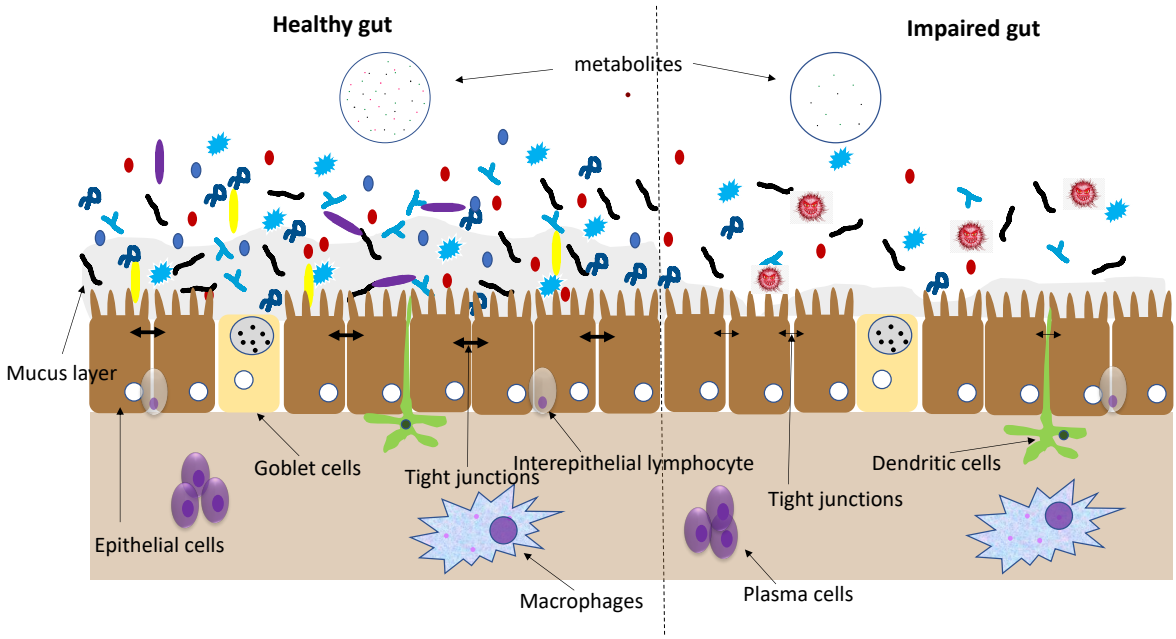

Figure 1.7: Direct colonisation resistance mechanisms for a healthy and impaired gut.

Direct colonisation resistance mechanisms for a healthy and impaired gut. Differences in the thickness of the mucus layer, number of organisms making up the gut microbiota, number of beneficial bacteria and metabolites released from the gut microbiota are shown as protective agents for the underlying epithelial cells. In addition, immune cells found in the submucosa are also shown.

Table 1.3: Direct colonisation resistance mechanisms: healthy and impaired gut

Healthy gut	**Impaired gut**
Rich and diverse microbiota	Reduced microbiota richness and diversity
Rich and diverse repertoire of metabolites	Reduced richness and diversity of metabolites
More healthy promoting microbiota members	An increase in pathobiont/ pathogens
Thick mucus layer	Reduced thickness of the mucus layer
Strong, tight junctions	Weak, tight junctions
Uncompromised epithelial layer	Compromised epithelial layer

1.3.4.1 Direct colonisation resistance: Acting as a physical barrier

The dense population of microorganisms inhabiting the gut provides a protective cover for the gut lumen against invading enteric pathogens. A healthy gut microbiota will have a dense microbial community overlying the mucus and outer membrane of the gut lumen. Together with the mucus, they provide an uncompromised barrier against pathogenic organisms. The microbiota also blocks the binding of pathogens to attachment sites present in the mucous layer and also plays a role in mucus secretion through SCFAs that it produces (Kamada, Chen, et al., 2013). Antibiotic treatment has been shown to reduce the thickness of the mucus layer, which results in increased contact between commensal or pathogenic organisms and the epithelial cell (Zou et al., 2018).

1.3.4.2 Direct colonisation resistance: Production of soluble inhibitors

Besides providing a barrier for enteric pathogen and receptor interaction, the diverse beneficial bacterial species in the gut also secrete antimicrobial substances such as bacteriocins that lyse or inhibit pathogenic organisms. *Bacillus*, a component of the

gut microbiota, has been intensively studied for its ability to inhibit enteric pathogens through the production of antimicrobial substances. In 2007, Rea et al demonstrated that *Bacillus* controlled the growth of *C. difficile* but not *Escherichia coli* and *Salmonella* Typhimurium (Rea et al., 2007). Specifically, *Bacillus thuringiensis* produces thuricin CD, a bacteriocin with antimicrobial activity against *C. difficile* and Listeria monocytogenes (L. monocytogenes). A contact-dependent killing has also been reported in Gram-negative bacteria having a type VI secretion system (T6SS). These produce toxic effector proteins, which are directed to other bacteria or eukaryotic host cell. T6SS mediated killing is utilized by Vibrio cholera against E. coli and Pseudomonas aeruginosa (C. Chen, Yang, & Shen, 2019). Salmonella, however, uses T6SS to establish itself in the gut (Sana et al., 2016).

Production of secondary bile products and short-chain fatty acids by beneficial bacteria is also important for pathogen growth or virulence inhibition*. Bacteroides thetaiotaomicron* is a vital colonizer of the mucus layer. It digests mucin peptides and O-linked glycans as an energy source, resulting in the production of short-chain fatty acids (SCFA) such as butyrate (Koh, De Vadder, Kovatcheva-Datchary, & Bäckhed, 2016). Furthermore, many microbiota species such as bifidobacteria and *Lactobacillus* can produce organic acids and SCFAs, some of which are detrimental to the growth of bacterial pathogens such as *S.* Typhimurium (Ducarmon et al., 2019; Gantois et al., 2006). Butyrate is an “anti-virulence molecule” that acts as a diffusible signal to downregulate the expression of the Salmonella Pathogenicity Island-1 Type 3 secretary system (SPI-1 T3SS) invasion genes in *Salmonella*. In contrast, the SCFA formate induces the expression of the SPI-T3SS and invasion. The SCFAs; formate and acetate are primarily located within the small intestine, whereas butyrate and propionate predominate in the colon.

1.3.4.3 Direct colonisation resistance: Competition for nutritional niches

One other important CR mechanism employed by commensal bacteria in the gut is competing for nutrients. The microbiota can create a nutritional environment not favourable for the growth of bacterial pathogens (Kamada, Chen, et al., 2013). For example, the resident microbiota can modify the intestinal composition of carbohydrates and sugars present, which are essential for the growth of pathogens. *Bacteroidetes thetaiotaomicron* excludes pathogenic *Citrobacter rodentium* by consuming carbohydrates that the pathogen would otherwise use (Kamada et al., 2012). The rich metabolome produced by the gut microbiota has also been shown to confer CR against fungal infection by competing for nutrients through the Target of Rapamycin (TOR) pathway. Carlos Garcia et al. demonstrated that the gut metabolome has antifungal activities against the dominant fungal pathogen, Candida albicans (C. albicans) and other fungal species resident in the gut. This was shown to be mediated by secretary molecules and not cells. Out of the 2,360 mutants that they tested, the kog1 mutant, a conserved subunit of TOR, was responsible for promoting antifungal activity (García et al., 2017). Kog1 is central to eukaryotic cell growth with regards to nutritional requirements and stress. The study also demonstrated that the TOR was mainly secreted from the gut microbiota components, *B. ovatus*, *R. intestinalis* and *R. faecalis* (García et al., 2017). Other studies have also shown the importance of *B. ovatus* in controlling recurrent *C. difficile* and *S.* Typhimurium.

1.3.4.4 Indirect colonisation resistance: Development and maturation of the immune system

Gut microbiota have been shown to impact stimulation of both the innate and adaptive immune system (Mazmanian, Liu, Tzianabos, & Kasper, 2005). Using germ-free mice, a study demonstrated that a gut that was pre-transplanted with faecal matter could recruit dendritic cells while germ-free mice failed to recruit dendritic cells. The number and function of intestinal macrophages and Tregs have also been shown to be compromised in germ-free animal models as compared to their pre-colonized counterparts (Mitsuyama et al., 1986; Mørland & Midtvedt, 1984; Östman et al., 2006).

Bacterial colonisation of the gut also impacts the differentiation of naïve T cells into Forkhead box P3 (Foxp3)+ Treg cells or various Th1, Th2, and Th17 (Hori, Nomura, & Sakaguchi, 2003; Legoux et al., 2019). Treg cells suppress the differentiation of naïve T cells into Th cells and have various anti-inflammatory effects, including suppressing the inflammatory activities of mast cells, basophils, eosinophils suppressing IgE, and induction of IgG4(Schülke, 2018). Several intestinal bacteria, including *Lactobacillus*, *Bifidobacterium*, *Bacteroides*, *Clostridium*, and *Streptococcus*, as well as bacterial metabolites such as butyric acid and propionic acid, have been shown to induce Treg cells in various mouse models or cell culture (Rivière, Selak, Lantin, Leroy, & De Vuyst, 2016). Treg cells are generated in the thymus (tTreg) and periphery (pTreg). pTregs, but not tTregs, control mucosal Th2 inflammation in the gut and lungs (Holohan, Van Gool, & Bluestone, 2019; Östman et al., 2006). Recent studies have shown that pTregs are promoted by intestinal bacteria and a symbiosis factor (polysaccharide A) produced by Bacteroides fragilis

(Jiang et al., 2017). Each type of Th cell plays a distinct and critical role in shaping and amplifying the immune response by producing cytokines that can suppress other types of Th cells. Th17 cells, which are abundant in the mucosa, secrete cytokines, including IL-17, IL-17F, and IL-22 (Liang et al., 2006; Qu et al., 2013). These cytokines improve the barrier function of the GI tract and protect against bacteria and fungi. In addition, th17 cells are promoted by bacterial flagella, unmethylated DNA, and adenosine triphosphate (ATP) (Y. Wang et al., 2019). Mutual regulation of Th1 and Th2 cells is crucial for immune homeostasis, but extreme activation results in chronic inflammation and autoimmune and allergic disease (Spellberg & Edwards, 2001).

A bidirectional relationship have been proposed, in which intestinal bacteria affect not only the development of the immune system but also the host immune system influences the development of the gut microbiota. Immune factors in infancy, such as the presence of maternal antibodies and intestinal gut microbiota-primed B cell trafficking, play important roles in the selection of intestinal bacteria (Belkaid & Hand, 2014). Through the placenta and breast milk, maternal antibodies such as IgG and IgA are introduced into infants. The presence of IgAs specific for intestinal bacteria are associated with differences in bacterial habitats in GI tracts (Y. Wang et al., 2019). Bunker et al. showed that intestinal bacteria in the small intestine are bound to bacteria-specific IgA, whereas, in the large intestine, most bacteria are not bound to IgA (Bunker et al., 2015). This suggests that bacteria bound by IgA effectively colonize the small intestine. Furthermore, segmented filamentous bacteria (SFB) overcolonize the GI tracts of IgA-knockout mice and strongly induce immunity (Lecuyer et al., 2014). Thus, SFB plays a role in shaping immune functions, such as

promoting IgA secretion and promoting intestinal cell lymphocytes and Th17 cells (Lecuyer et al., 2014). Based on these functions of IgA molecules, maternal IgA in breast milk may function similarly; thus, IgA may be involved in selecting bacteria that colonize the GI tract in infancy.

1.4 Factors that influence gut microbiota composition and function

The composition of the gut microbiota varies not only between individuals but also within individuals. There are several factors that are associated with these inter and intra individual variations. These variations can be influenced by factors occurring prenatally, at birth and postnatally. Some of the changes in gut microbiota composition may promote good health, while others may create conditions that encourage the growth of pathogens. Age already described in **section 1.2** is one of the main factors that influence gut microbiota composition. **Figure 1.8** shows different factors that affect gut microbiota composition.

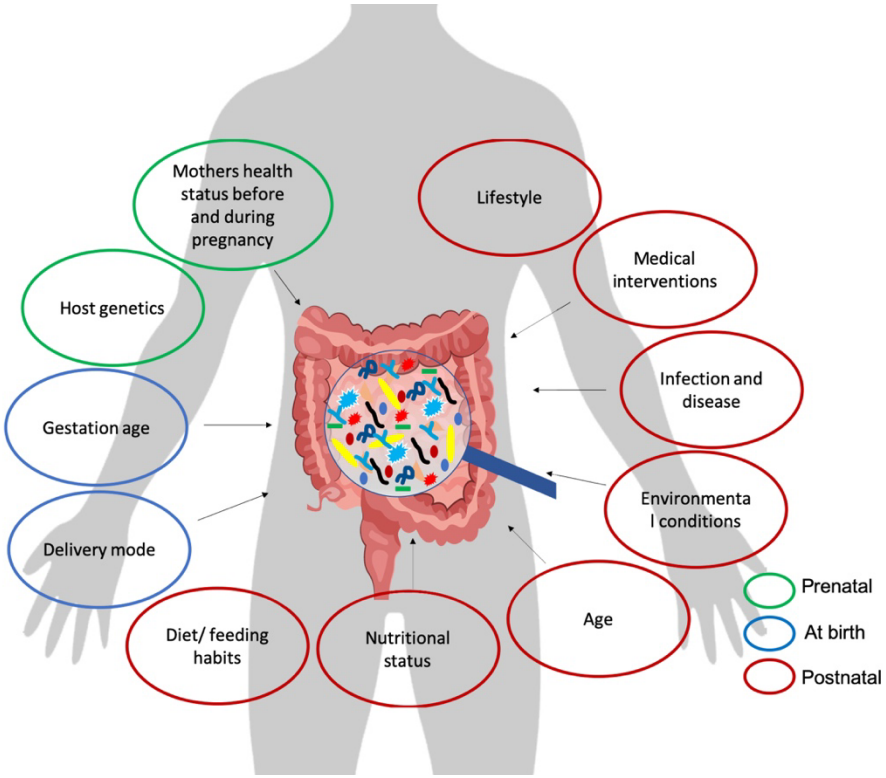

Figure 1.8: Factors that affect gut microbiota composition.

Factors that affect gut microbiota composition. Factors are grouped based on time of occurrence, such as prenatal, at birth and postnatal.

1.4.1 Mothers health status before and during pregnancy

A mother's health status is very important for fetal development and the health status of the child after birth. Factors such as HIV status, malaria infection, urinary tract infection, sexually transmitted infections, hepatitis and parasitic infections in a mother impact pregnancy outcomes or the child's health status (van den Broek, Jean-Baptiste, & Neilson, 2014; Walker et al., 2007). Screening for some of these factors is done in Malawi during antenatal clinics to avert fatal outcomes. A healthy mother has been associated with optimal fetal brain development, healthy birth weight and a reduction in the occurrence of congenital disabilities (Kamng'ona et al., 2020; Stanislawski et al., 2017). A mother's health status has also been associated

with a child's gut microbiota composition and function. Mothers nutritional status, pregnancy weight gain, disease/infection and antibiotic usage are important aspects that have been considered to impact a child's gut microbiota composition and function (Milliken, Allen, & Lamont, 2019; Sherman et al., 2018; Zou et al., 2018). A study done by Ponzo et al. (2019) looking at gestational diabetes mellitus (GDM) and the gut microbiota of the offspring demonstrated that nutritional habits in the first-trimester had more impact on the infant gut microbiota than the third-trimester maternal nutritional habits. The study demonstrated that infants from mothers with GDM had a relatively higher abundance of pro-inflammatory taxa than offspring's from healthy mothers (Ponzo et al., 2019). Uninfected children born from HIV-infected mothers have been shown to have an impaired gut microbiome which may explain the high mortality rate in this group compared to children born from HIV negative mothers (Bender et al., 2016). A mother's nutritional status, specifically obesity, is a risk factor in a child developing non-communicable diseases later in life (Godfrey et al., 2017). This effect of maternal obesity on offspring outcome may be mediated by alterations in the gut microbiome and DNA methylation (Godfrey et al., 2017).

Antibiotic usage is widespread during pregnancy. Their use has immediate beneficial effects on the mother and child, but their long-term impact on the child has been detrimental. Their usage has also been shown to affect gut microbiota composition in children. In a study by Zou et al., prenatal antibiotic exposure influenced the gut microbiota by significantly reducing levels of *Bifidobacterium* and Bacteroides (Zou et al., 2018). Antibiotic usage during pregnancy has also been associated with obesity,

changes in immune regulation and behaviour in mice (Champagne-Jorgensen et al., 2020).

1.4.2 Host genetics

It is well known that host genetics determines the physiology and function of the body and its systems. The question that microbiome research scientists have not yet been able to answer, however, has been, “To what extent does host genetics influence gut microbiota composition and diversity?”. It is not easy to determine a clear relationship because of the other important factors that influence gut microbiome composition. The impact of host genetics in gut microbiota composition is modest, and may become non-significant when other factors such as environmental conditions and demographics are considered. Genome-wide association studies are helping to reveal the association between the microbiome and host genes related to the immune system, diet, metabolism and vitamin D receptors. Several studies have previously identified “heritable bacteria taxa” meaning a pattern of single nucleotide polymorphisms (SNPs) associated with specific bacterial taxa or pathways (Bonder et al., 2016). For example, in a study conducted by Wang, J et al., 10% of the difference in microbiome Beta diversity was explained by 42 SNPs identified in the study (J. Wang et al., 2016). In a very recent study done by Daphna Rothschild and colleagues, 1,046 healthy Israel participants were recruited into their study, which aimed to investigate the extent to which human host genetics determines microbiota composition compared to other factors (Rothschild et al., 2018). This study demonstrates that environmental factors rather than host genetics significantly shape the gut microbiome (Rothschild et al., 2018). A comprehensive assessment of the association of 110 factors and gut microbiota

diversity in 1000 healthy individuals supports Rothschild's findings. The 110 factors included demographic, environmental and host genetic factors. Demographic and environmental factors were the main contributors to differences in alpha and beta diversities, and not host genetics (Scepanovic et al., 2019).

1.4.3 Gestation age; term vs preterm births

Gut microbiota composition can also be affected by neonatal factors such as gestation age and mode of delivery. Based on the gestation period, children can either be classified as preterm, term or post-term. Children born before 37 weeks of gestation are referred to as preterm babies; those born after 37 weeks of gestation are term babies, while those born after 42 weeks of gestation are classified as post-term (WHO, 2018a). According to the WHO, it is estimated that there are 15 million preterm births every year (WHO, 2018a). Malawi has the highest preterm birth rate globally, 18.1 per 100 live birth (WHO, 2018a). Persistent malaria, being under 20 years, and anaemia are some of the factors associated with preterm birth in Malawi (van den Broek et al., 2014). Complications due to preterm birth include death of babies within the first year of life and lifetime disabilities. In addition, spending more time in the hospital under intensive care, receiving more antibiotics and mostly feeding on formula milk makes them more susceptible to infections and other complications, and preterm babies have an immature immune system, underdeveloped gut barrier function and compromised gut microbiota composition. These factors contribute to the reported low gut microbiota diversity in preterm babies compared to term babies reported to have a rich gut microbiota. Using a Norwegian cohort of 519 children (160 preterms), Dahl et al. demonstrated that preterm children had a distinct gut microbiota composition and function characterised

by a significantly low bacterial diversity compared with term children at ten days postnatally (Dahl et al., 2018). This was not explained by more antibiotic, shorter breastfeeding and caesarean delivery. Birth weight, hospital environment/ microflora, length of stay in the hospital, mothers faecal, vaginal and skin microbiota may also affect a child's gut microbiota composition.

1.4.4 Delivery mode; vaginal vs caesarian birth

Following low-level intrauterine microbiota colonisation, significant colonisation takes place during delivery. The microbiota composition at this stage is mainly influenced by the mother's gut and vaginal microbiota and the surrounding environment (Bäckhed et al., 2015b; Dominguez-Bello et al., 2010). Delivery mode, whether vaginal or caesarean, has been associated with specific microbiota composition (Dominguez-Bello et al., 2010). In vaginal delivery, neonates are exposed to microorganisms that inhabit the birth canal and the vagina, which have a rich microbiota, but low diversity compared to the gut microbiota. The major vaginal microbiota genera include *Prevotella*, *Lactobacillus* and *Sneathis* (Hyman et al., 2005; Zhou et al., 2007). Moore et.al. describes difference in gut microbiota composition with delivery more (R. E. Moore & Townsend, 2019). **Figure 1.9** shows the gut microbiota compositional differences with mode of delivery.

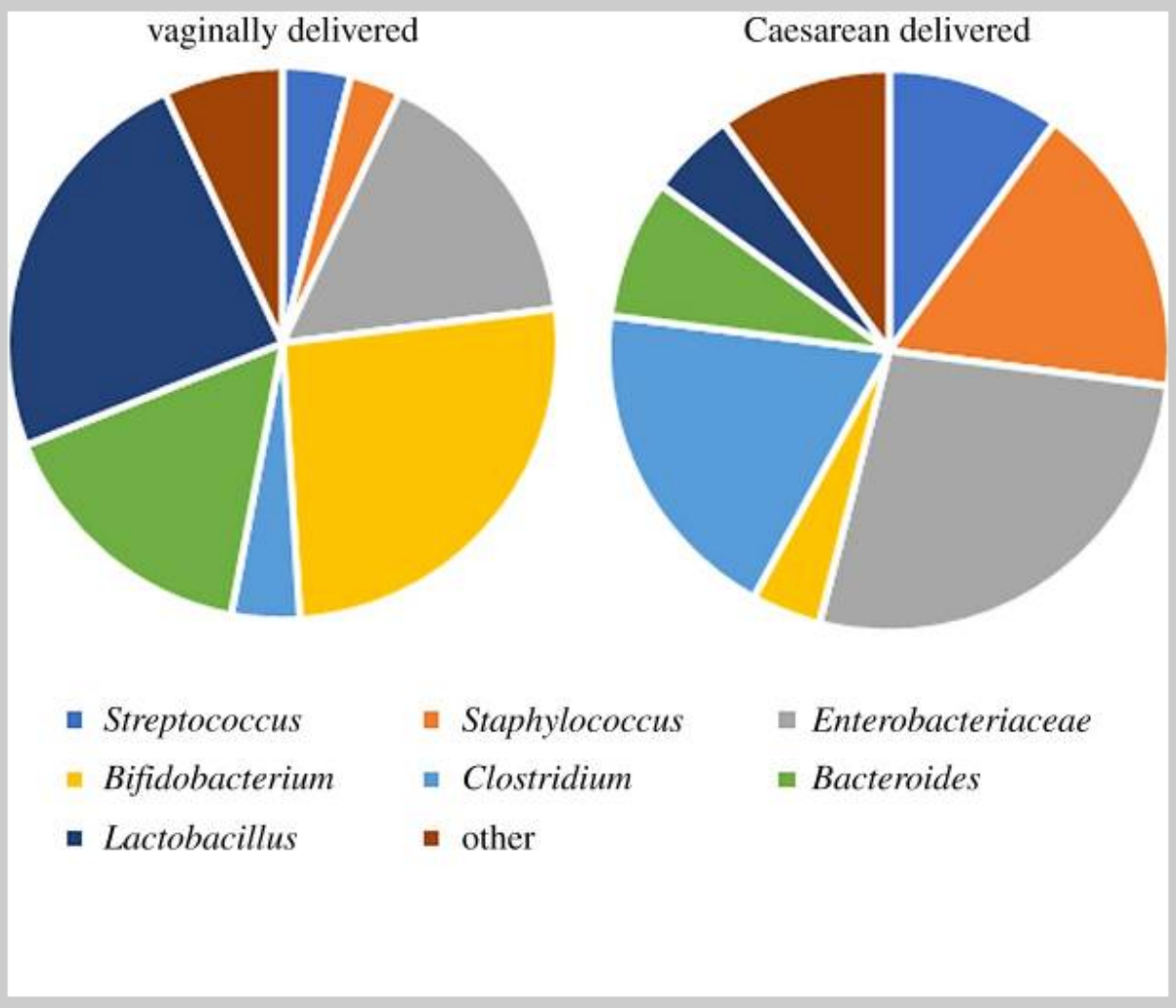

Figure 1.9: Differences in gut microbiota composition with mode of delivery

Differences in gut microbiota composition in vaginally and caesarean delivered infants. Reproduced from (R. E. Moore & Townsend, 2019)

On the other hand, caesarean born babies microbiota composition (major genera include *Acinetobacter*, *Staphylococcus* and *Micrococcineae)* has been shown to resemble the mothers' skin microbiota and the surrounding environment. (Dominguez-Bello et al., 2010). Results from a systematic review done by Erigene Rutayisire and colleagues supports these facts. In this review, the microbiota of vaginally delivered neonate was mostly composed of *Actinobacteria* and

Bacteroides, while caesarean born neonates had gut microbiota dominated by *Firmicutes* (Rutayisire, Huang, Liu, & Tao, 2016). Several studies have, however, observed that the gut microbiota composition of infants at six months of life is comparable regardless of the mode of delivery.

In Malawi, most infants are born vaginally. A study that collected mode of delivery information from 62,033 HIV positive and HIV negative women from central and southern regions of Malawi demonstrated that 80% of the women had spontaneous vaginal delivery while 14% had a caesarean section (Tenthani et al., 2018).

1.4.5 Early feeding habits; breastfeeding vs formula feeding

Whether a baby is breastfed, or formula-fed is another major factor determining the composition of the gut microbiota. Facultative anaerobic bacteria, *Staphylococcus spp., Streptococcus spp*. and *Enterobacteriaceae,* are dominant during the first days after birth (Turroni et al., 2012). Strict anaerobic bacteria outcompete the facultative anaerobic bacteria by depleting oxygen in the gut. The strict anaerobes subsequently replace the facultative anaerobes. This change in the dominant bacteria during the first days of life has been attributed to the type of food the baby is introduced to, whether breast milk or formula milk. In Malawi, most infants are exclusively breastfed in the first three months, and they continue breastfeeding until they are two years of age (Apanga et al., 2021). Breastmilk contains Human Milk Oligosaccharides (HMO), glycoproteins and a microbiota rich in the genera *Streptococcus*, *Staphylococcus*, *Weisella, Leucosticte* and *Lactococcus (Bode, 2012; Moossavi et al., 2019)*. Bifidobacteria digest the polysaccharides in human

milk(Thomson et al., 2018). The polysaccharides and anaerobic conditions in the baby's gut promote the growth of *Actinobacteria* and *Firmicutes*.
Actinobacteria mainly *Bifidobacterium* species; *Bifidobacterium breve (B. breve), Bifidobacterium longum (B. longum), Bifidobacterium infantis (B. infantis*) and *Bifidobacterium pseudocatenulatum (B. speudocatenulatum)* and. Lactic acid-producing bacteria, *Lactobacillus*, *Enterococcus* and *Clostridium* species (Firmicutes phylum) are the most predominant (Bergstrom et al., 2014; Voreades et al., 2014). The gut microbiota of babies that are fed formula milk are, on the other hand, colonised mainly by *Escherichia coli (E. coli)* and *C. difficile* in comparison with breastfed babies (S. A. Lee et al., 2015; Murphy et al., 2017; Timmerman et al., 2017).

1.4.6 Other feeding habits

A significant shift in the composition of the gut microbe community occurs when the infant is introduced to other foods and when the baby is being weaned. A study that describes the microbiome of 98 mothers: infant pairs over the first year of the infants' life demonstrated that the introduction of formula and other foods at about four months of age was associated with an abrupt increase in the phylum Bacteroides (Bäckhed et al., 2015a). Complete cessation of breastfeeding and complete dependence on adult diet was associated with the establishment of a more stable gut microbiota similar to an adult gut microbiota. Cessation of breastfeeding is also associated with a microbiota that has an increased capacity to degrade plant and animal derived polysaccharides (Bäckhed et al., 2015a; Selma-Royo et al., 2019).

Later in life, the type of diet has also been shown to impact the gut microbiota composition. Studies have demonstrated that some dietary elements are able to

induce microbial change within 24 hours. Dietary impact is being used to alter the gut microbiota therapeutically. Communities whose diet is rich in proteins are associated with a different microbiota composition than those with a diet rich in carbohydrates (Conlon & Bird, 2015). Protein-rich diets are associated with an overall increase in microbial diversity. *Bifidobacterium* and *Lactobacillus* have been reported to have increased in a randomised controlled trial where subjects were given whey and pea protein extracts (B Romond et al., 1998).

1.4.7 Nutritional status

Linked to diet is nutritional status. Nutritional status, whether malnourished or obese, has also been associated with modification of gut microbiota composition. In a study that was conducted in Dhaka, Bangladesh, to look at the relationship between nutritional interventions, gut microbiota and therapeutic responses during the first two years of life, Subramanian et al. demonstrated that severe acute malnutrition (SAM) was significantly associated with immaturity of the gut microbiota as compared to healthy controls (Subramanian et al., 2014). Nutritional interventions only resulted in partial reconstitution (Kamng'ona et al., 2020; Subramanian et al., 2014). Partial reconstitution was also observed in a study of Malawian twin pairs where children with kwashiorkor only had a transient maturation of metabolic function after being given ready-to-use therapeutic food (Smith et al., 2013). In another study conducted in rural Malawi, prenatal and postnatal nutritional intervention promoted infant gut microbiota diversity at 18 months but not gut microbiota maturation (Kamng'ona et al., 2020). A meta-analysis of microbiome sequenced data from 184 children from five different countries in Africa and Asia specifically shows that malnutrition was associated with depletion of species

belonging to *Bacteroidaceae, Eubacteriaceae, Lachnospiraeae,* and *Ruminocoeae* families and enrichment in *Enterococcus faecalis*, *Escherichia coli* and *Staphylococcus aureus* (Million et al., 2016). Gut microbiota influence the amount of energy extracted from food and used by the host and affects energy usage and storage. In a study where stool from malnourished Malawian children was transplanted into germ free piglets demonstrated that the gut microbiota does influence nutritional status and that it can transmit malnutrition. Studies have therefore demonstrated that malnutrition is associated with gut microbiota dysbiosis and that nutritional intervention promotes gut microbiota diversity. Obesity, on the other hand, is associated with increased gut microbiota composition, diversity, and SCFAs, which suggests increased gut microbiota mediated metabolic function (Hou et al., 2017; Riva et al., 2017). It is, however, not clear whether increased gut microbiota diversity impacts gut microbiota mediated biological processes. This necessitates the need for studies looking at nutrition and gut microbiota function.

1.4.8 Environmental conditions

The surrounding environment of the host determines the type of microorganisms to which the host is exposed. This subsequently impacts what type of pathogens with which the host can be infected, and indirectly affects host immunity. Geography and seasonality affect the surrounding environment and also diet. In caesarean born babies, the surrounding environment is the primary determinant of the early colonizers of the gut (Dominguez-Bello et al., 2010; Tun et al., 2017). Studies comparing gut microbiota composition among different ethnicities or between different settings have shown marked differences in the associated gut microbiota composition. A study by Yatsunenko et al. (2012) compared the gut microbiota

composition between populations with notable differences in terms of geography, social, economic status and cultural settings. They demonstrated significant differences in gut microbiota composition between participants from the three populations: Malawians, Amerindians, and Americans (**Figure 1.10**). The rural Malawian and Amerindian participants clustered together but separated from the participants from the metropolitan area in the United States (Yatsunenko et al., 2012). A study comparing gut microbiota composition of 6 months old infants living in rural Malawi and urban Finland also demonstrated differences in gut microbiota composition between these two sites (Grześkowiak et al., 2012). Another study looked at gut microbiota composition in different ethnicities within the United States of America (USA) identifies 12 microbial genera and families that reproducibly varied by ethnicity (Brooks, Priya, Blekhman, & Bordenstein, 2018). The influence of ethnicity on gut microbiota composition and function has, however, not been further explored. However, the differences observed may be linked to different factors, including environmental conditions, feeding behaviours, maternal adult microbiota, and underlying health conditions.

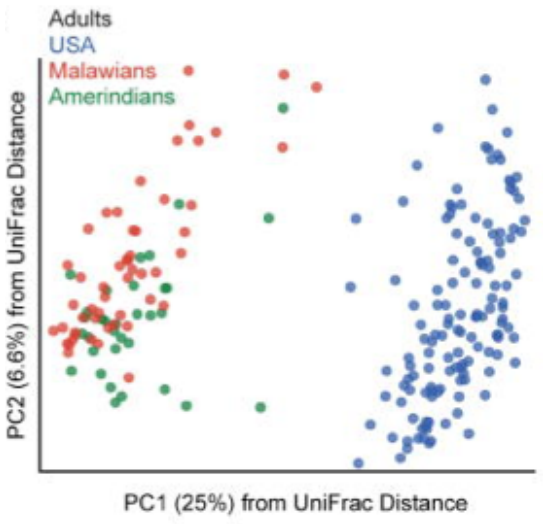

Figure 1.10: Differences in gut microbiota composition by geography.

Gut microbiota composition differences for three populations with marked differences in geography, socio-economic status and cultural setting. Reproduced from (Yatsunenko et al., 2012)

1.4.9 Infection and disease

Infections, whether gut localized or involving distal organs, are being linked to gut microbiota compositions. Diseases such as inflammatory bowel diseases (IBD), type 2 diabetes, allergies, obesity and colorectal cancer have also been associated with gut microbiota dysbiosis (Cho & Blaser, 2012). Inflammatory responses induced by pathogenic organisms alter the gut microbiota composition and promote microorganisms' proliferation and translocation (B. Wang, Yao, Lv, Ling, & Li, 2017). Microorganisms adapted to withstand immune defense, such as pathogens, are likely to survive in an inflamed environment. Bacterial products such as LPS, cell adherence, tissue invasion or increased resistance to phagocytosis may trigger and promote inflammation (Chow & Mazmanian, 2010). Organisms such as *S.* Typhimurium have acquired virulence factors that trigger an inflammatory immune response that works to its advantage (Stecher, 2015). These virulence factors are

Salmonella pathogenicity islands, Type 3 secretary systems, effector proteins and fimbriae. Using these virulence factors, Salmonella induced intestinal inflammation alters the gut microbiota composition and promotes its growth (Stelter et al., 2011).

There is also emerging evidence of the link that is there between the gut microbiota and the lung. This relationship has a bearing on common respiratory diseases such as chronic obstructive pulmonary disease (COPD), cystic fibrosis, lung cancer, asthma. The fact that gastrointestinal symptoms accompany chronic respiratory diseases and that patients with inflammatory bowel disease and gastroesophageal reflux are likely to develop pulmonary disease prompted scientist to study the role of gut microbiota in respiratory infections (Jové Blanco et al., 2020). In asthma, high gut microbiota diversity is protective. Breastfeeding, which also influences gut microbiota composition, has been shown to protect against asthma and allergic diseases in children (Chunxi, Haiyue, Yanxia, Jianbing, & Jin, 2020). In addition, specific early gut microbiota members such as *Bifidobacterium* have been shown to be protective against asthma. A low abundance of *Bifidobacterium*, *Akkermansia* and *Faecalibacterium* and a high abundance of *Candida* and *Rhodotorula* have been shown to increase a child's risk of developing allergies and asthma (Fujimura et al., 2016).

1.4.10 Medical interventions

Medical interventions such as antibiotics, vaccines, prokinetics, laxatives and pre/probiotics are also thought to impact the gut microbial community. Many studies have been done on antibiotics and pre/probiotics with a large body of evidence showing that they affect the gut microbiota composition. Antimicrobials are globally

used, including Malawi, where both prescribed and over-the-counter use is common. Antimicrobials have been shown to have a profound impact on the gut microbiota. Over the counter, antibiotic usage is also becoming common in Malawi.

The use of antibiotics in early life has profound effects on the development of the gut microbiota, with some results showing long term effects. The use of antibiotics in infants shifts the composition of the gut microbiota toward a high abundance of Proteobacteria and low abundance of Actinobacteria populations (Palleja et al., 2018), decreases the overall diversity of the infant's microbiota, and selects for drug-resistant bacteria (Palleja et al., 2018). According to some epidemiological surveys, antibiotics in early life increase the risk of developing allergic diseases such as asthma, atopic disease, eczema, and type 1 diabetes (Metsälä et al., 2015; Mueller et al., 2015).

Tanaka et al. carefully examined the effects of antibiotic exposure in the early postnatal period on the development of the intestinal microbiota (Tanaka et al., 2009). The faecal microbiotas of newborns orally administered a broad-spectrum antibiotic for the first four days of life were analyzed for two months. In the first week of the infants' lives, the antibiotic-administered subjects showed less diversity in their faecal bacterial communities, with the attenuation of some bacterial groups, especially *Bifidobacterium*, as well as unusual colonisation with Enterococcus (Tanaka et al., 2009). At one month of age, overgrowth of Enterobacteriaceae and Enterococcus was observed in infants in the antibiotic-treated group. It is possible that these microbiota changes were related either to the antibiotic treatment, or to the underlying illness requiring the treatment. Interestingly, however, caesarean-

delivered subjects whose mothers were intravenously administered a closely related antibiotic in the same period showed similar, although weaker, associations with microbiota development (Tanaka et al., 2009), attributed to ingestion by the infant through breastmilk. This suggests that the changes are related to the drug rather than illness in the infant. These results indicate that antibiotic exposure at the beginning of life significantly affects the development of neonatal intestinal microbiotas.

Most studies done on pre/probiotics have been done in developed countries. Some and not all pre/probiotics have been associated with increased gut microbiota diversity and richness (Fijan, Sulc, & Steyer, 2018; Korpela et al., 2018). Studies on the gut microbiota and pre/probiotics are not available in Malawi because pre/probiotics are not commonly used.

Although scientists are keen to understand the interaction between the gut microbiome and oral vaccine response, more studies need to be done to understand how prokinetics and laxatives impact the gut microbiota. Several studies have now been done to understand the gut microbiome and oral rotavirus vaccine efficacy (Carvalho & Gill, 2019; A. H. Kim, Hogarty, Harris, & Baldridge, 2020; Parker et al., 2018). These studies have been driven by the low rotavirus vaccine efficacy observed in Africa and Asia. For example, a study done in Ghana and Pakistan demonstrated that the gut microbiota significantly correlates with vaccine immunogenicity (V. Harris et al., 2018). Such findings have led to the hypothesis that the gut microbiota composition in children in Africa and Asia is associated with low vaccine efficacy (V. C. Harris et al., 2017; Parker et al., 2018).

1.4.11 Lifestyle

The type of lifestyle that people lead, such as exercising, smoking and excessive drinking of alcohol, is associated with their health status. For example, living a more sedentary lifestyle is associated with the occurrence of cardiovascular diseases, type 2 diabetes, metabolic syndrome and cancer (Owen, Sparling, Healy, Dunstan, & Matthews, 2010; Tremblay, Colley, Saunders, Healy, & Owen, 2010), smoking is associated with lung disease. Recent studies are starting to show that lifestyle impacts host health by, among others modulating the microbial community in the gut. Studies in mice and humans have both demonstrated that exercise increases gut microbiota diversity and modifies microbiota composition. In a study by Clarke et al. comparing athletes and healthy controls, athletes had a more microbiota diversity, with 22 phyla and 113 genera than healthy controls with 11 phyla and 65 genera, respectively (Clarke et al., 2014). Another study on athletes demonstrated that exercise affects the gut microbiota composition and that a much more significant impact is observed in the gut microbiota function when they studied the gut microbiota at a metagenomic and metabolomic level. Smoking and alcohol have both been shown to modify the gut microbiota. Alcohol is, however, known to deplete bacteria with known anti-inflammatory properties, which results in a leaky gut (Boltin & Niv, 2014) and lower alcohol intake has been shown to be protective against conditions such as H. pylori disorders (Brenner, Rothenbacher, Bode, & Adler, 1997; Murray et al., 2002).

1.5 The impact of gut microbiota - studies in human health

Results from gut microbiota studies have provided evidence on the important roles that it plays in health and disease states. Most early gut microbiota studies were observational studies that focused on determining the association between particular health or disease states and gut microbiota components. Based on these studies, it was shown that some gut microbiota components were associated with particular disease states.

This led to interventional studies. In interventional studies, whole stool samples (as in the case of faecal microbiota transplant) or specific components or communities (as in artificial microbial communities) of the microbiota were used to prevent, control or treat particular disease states.

Lately, animal models are commonly used as a follow-up to observation studies with the aim of understanding the mechanisms involved. The gut microbiota plays important roles in digestion and metabolism, immune development and function, controlling pathogenic infection, maintaining intestinal homeostasis, development of the lymphatic system and brain maturity (Jandhyala et al., 2015). Changes occur in the composition of the gut microbiota, and these are influenced by different factors such as age, diet, host genetics, environment and infections(Conlon & Bird, 2015; Eggesbø et al., 2011; W. Fan, Huo, Li, Yang, & Duan, 2014; Kamada, Seo, Chen, & Nunez, 2013; Noverr & Huffnagle, 2004). Disruption of the gut microbiota may have detrimental consequences to human health. Changes in the gut microbial composition can result in loss of diversity and an overgrowth of opportunistic and pathogenic microorganisms. Studies have shown associations between changes in

the gut microbiota and pathological intestinal conditions like malnutrition, obesity, diabetes, and inflammatory bowel diseases; including ulcerative colitis and Crohn's disease (Gillis et al., 2018). Our understanding of the gut microbiota in relation to different disease states has allowed scientist to use the gut microbiota in controlling or treating infections, restoring health after an infection and promoting healthy states. Faecal Microbiota Transplant (FMT) is one gut microbiota based therapy that has significantly been successful in treating recurrent *Clostridium difficile (C. difficile)* infections (Hasegawa et al., 2012; Lau & Chamberlain, 2016). More studies need to be done to harness the powerful role of the gut microbiota in controlling other infections in populations where they cause more challenges, including sub-Saharan Africa (SSA).

1.6 Enteric infections

1.6.1 The importance of enteric pathogens

The intestinal tract is continually exposed to a broad range of biological stimuli that can significantly disrupt gut microbiota composition and normal function. Food and environmental materials, commensal and pathogenic microorganisms are all capable of inducing this biological stimulus. Of importance among these are the pathogenic microorganisms. Being open to the environment, the human gut is exposed to an extensive range of pathogens. Exposure to enteric pathogens and burden of enteric diseases is higher in regions where the quality of drinking water and level of sanitation are low (GBDa, 2018). This is the case with developing countries of Asia and SSA. In SSA, enteric pathogen infection, either causing diarrhoea or BSI, present a considerable threat to the existing poor public health system and the country's economy. In this region, enteric pathogen colonisation of the gut is thought

to increase following the introduction of supplementary feeding, which happens between 3 and 6 months after birth. The source of infection is mostly from contaminated water and food (Colston et al., 2020). A large scale study, the Global Enteric Multicenter Study (GEMS), which was done to identify the aetiology and population-based burden of pediatric diarrhoeal disease in SSA and South Asia, have shown that most attributable cases of moderate-to-severe diarrhoea are caused by rotavirus (Kotloff et al., 2013).

The study also demonstrated surprisingly high rates of multiple pathogens being detected from one sample, even in the absence of diarrhoea (Kotloff et al., 2013). This shows that the gut in children from this setting is frequently exposed to multiple enteric pathogens, and that this exposure is not always associated with clinical symptoms. Although pathogen exposure may be beneficial in that it could induce an immune response, it also poses a risk for clinical disease and pathogen transmission. In SSA, including Malawi, bloodstream infections still remain a big challenge that is now complicated with the emergence of multidrug-resistant pathogens. A study aimed at describing longitudinal trends in bacteremia and antimicrobial resistance at the largest referral hospital in Blantyre, the southern part of Malawi, demonstrated a reduction in BSI but an expansion in ESBL and fluoroquinolone resistance in the most commonly isolated Gram-negative bacteria (Musicha et al., 2017). Of the 26 174 blood culture isolates, 13 366 (52.1%) were resistant to first-line antibiotics used in Malawi at that time (amoxicillin or penicillin, chloramphenicol and co-trimoxazole). For the 19 year census period, NTS serovars *S.* Typhimurium and *S.* Enteritidis were the leading cause of BSI in under-five-year-old children (Figure 7).

1.6.2 Diarrhoeal infections

Although there is a notable reduction in the incidence of diarrhoea, globally, diarrhoeal illnesses remain an important cause of morbidity and mortality. The impact is however more pronounced in SSA where levels of sanitation are low. According to **Figure 1.11**, the Global burden of diseases report on causes of deaths in SSA children under the age of five ranked diarrhoea as the third with neonatal and lower respiratory illnesses being the first and second respectively (GBDa, 2018). In Malawi, a reduction in the incidence of diarrhoeal illnesses has been reported (Tizifa et al., 2021). This reduction has been associated with improvement in the management of malaria and the other public health interventions. Acute diarrhoea is one of the common reasons for seeking hospital care although it does not require antibiotics to manage it. Acute diarrhoea causes dehydrating and can thus easily become life threatening. Acute diarrhoea is managed through the administration of oral rehydration solution (Rissman, Deavenport-Saman, Corden, Zipkin, & Espinoza, 2021). Depending on pathogen and host factors, simple gastroenteritis can become chronic diarrhoea that may require antibiotic treatment. Besides causing clinical symptoms that can be life threatening, diarrhoea illnesses facilitate the spread and transmission of pathogen into the environment. This may lead to more infections and sometime result into disease outbreaks.

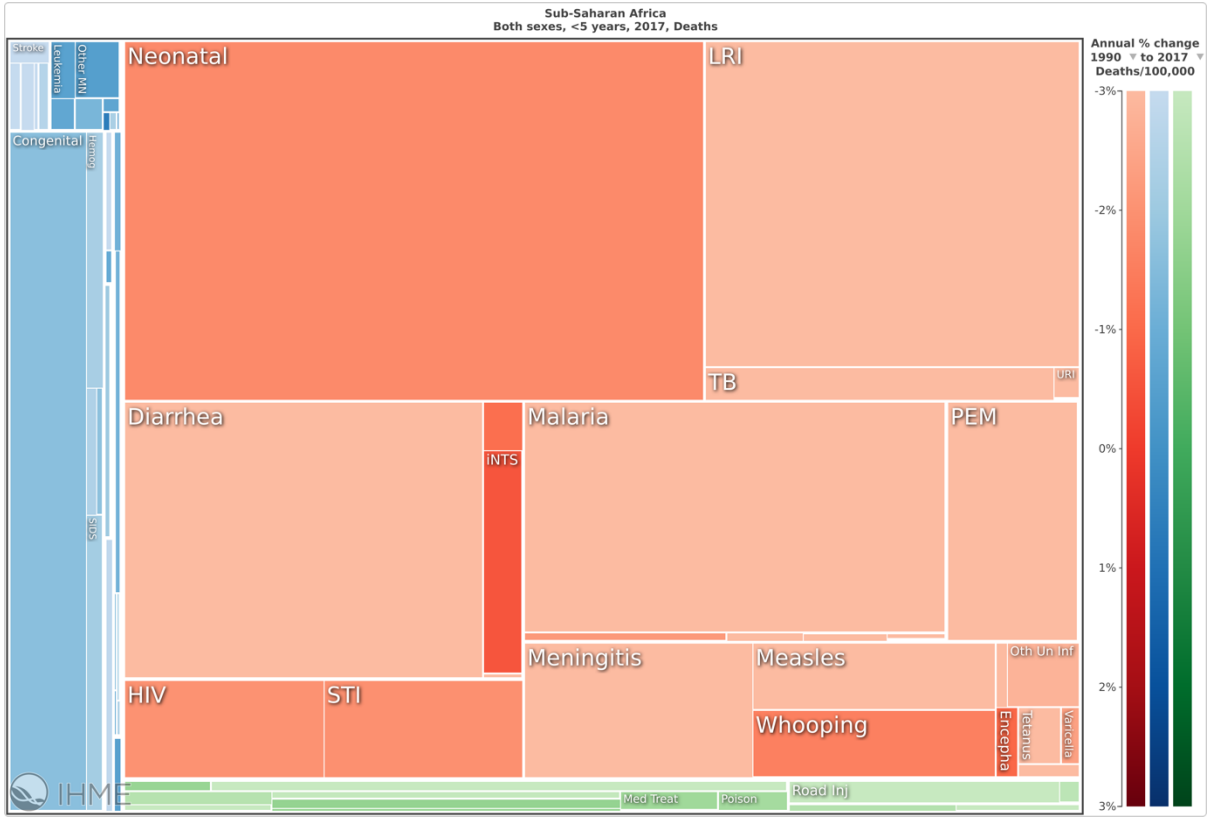

Figure 1.11: All causes of death in SSA for under five-year-old children.

All causes of death in SSA for children under five years old. Diarrhoea is among the leading causes of death in SSA children. Neonatal, lower respiratory infections and malaria are other important causes of mortality in the region (GBDb, 2018)

1.6.3 Bloodstream infection

Bloodstream infections are an important cause of morbidity and mortality worldwide (McNamara et al., 2018). Some important enteric pathogens causing BSI are *S.* Typhi, *S.* Paratyphi and some Non Typhoidal Salmonella (NTS) serovars (GBD, 2019b; Gilchrist & MacLennan, 2019; Msefula et al., 2019), E coli (Pascual et al., 2016; Riley, 2014; Temkin et al., 2018), *Bacteroides fragilis (B. fragilis)* and *C. difficile* (Brook, 2010). In Malawi iNTS, specifically *S*. Typhimurium, is the leading cause of BSI in young children and immunocompromised adults. **Figure 1.12** shows

BSI causing pathogens isolated at MLW (Musicha et al., 2017). Bloodstream infections are usually not self-limiting and require antibiotic treatment. The development of resistance to most standard antibiotics is a big challenge in the fight against bacteremia. Antibiotic resistance in pathogens causing bacteremia has been associated with a high mortality rate (Seboxa et al., 2015).

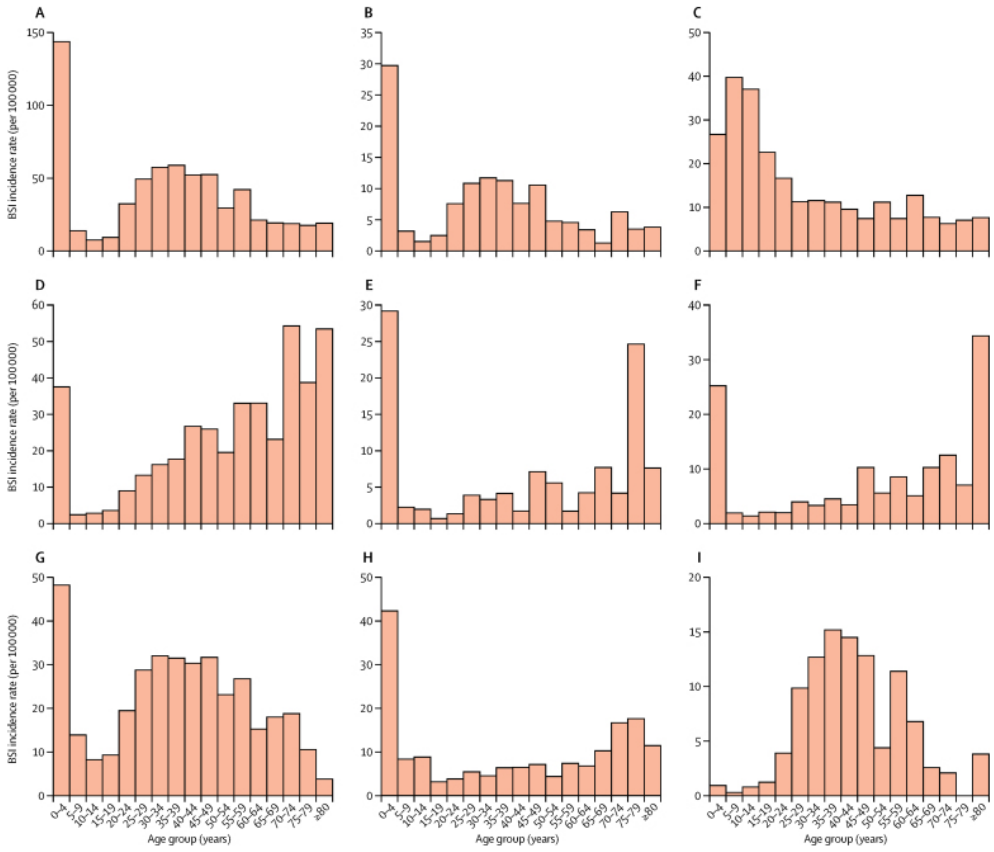

Figure 1.12: Estimated incidence of pathogens causing bloodstream infections in Blantyre, Malawi

Estimated incidence of pathogens causing blood stream infections in Blantyre, Malawi. *(A)* Salmonella *Typhimurium. (B)* S *Enteritidis. (C)* S *Typhi. (D) Escherichia coli. (E)* Klebsiella *spp. (F) Other Enterobacteriaceae. (G)* Streptococcus pneumoniae. *(H)* Staphylococcus aureus. *(I) Yeast. Adapted from (Musicha et al., 2017)*

1.6.4 Asymptomatic infection

A high burden of enteric pathogens has however, also been reported in the absence of diarrhoea and BSI (Kotloff, 2017; Platts-Mills et al., 2018). Thus, making asymptomatic pathogen exposure of public health interest as it may have positive and negative implications to the host or on the control of infections (**Figure 1.13**). One favorable implication is that gut exposure to pathogens, even asymptomatically, has been shown to enhance gut microbiota colonisation resistance mechanism (Stacy et al., 2021). Gut microbiota exposure may also result in protective immunity (Ly & Hansen, 2019). Asymptomatic exposure can, however, lead to pathogen transmission in the environment, is a potential risk factor for diarrhoea, invasive disease, stunting and a reduction in vaccine effectiveness.

A study by Nyirenda et al. (2015) also demonstrated the huge burden of asymptomatic *Salmonella* exposure in young Malawian children. The study found that up to 47% of children recruited into the study were exposed to *Salmonella* within the gut over 12 months (Tonney Stophen Nyirenda, 2015). More than half of the *Salmonella* isolates were identified as *S.* Typhimurium. Almost all the S. Typhimurium isolated from Malawi are sequence type (ST) 313 strains that are invasive and resistant to multiple antimicrobial agents (iNTS-GBD-Collaborators, 2019).

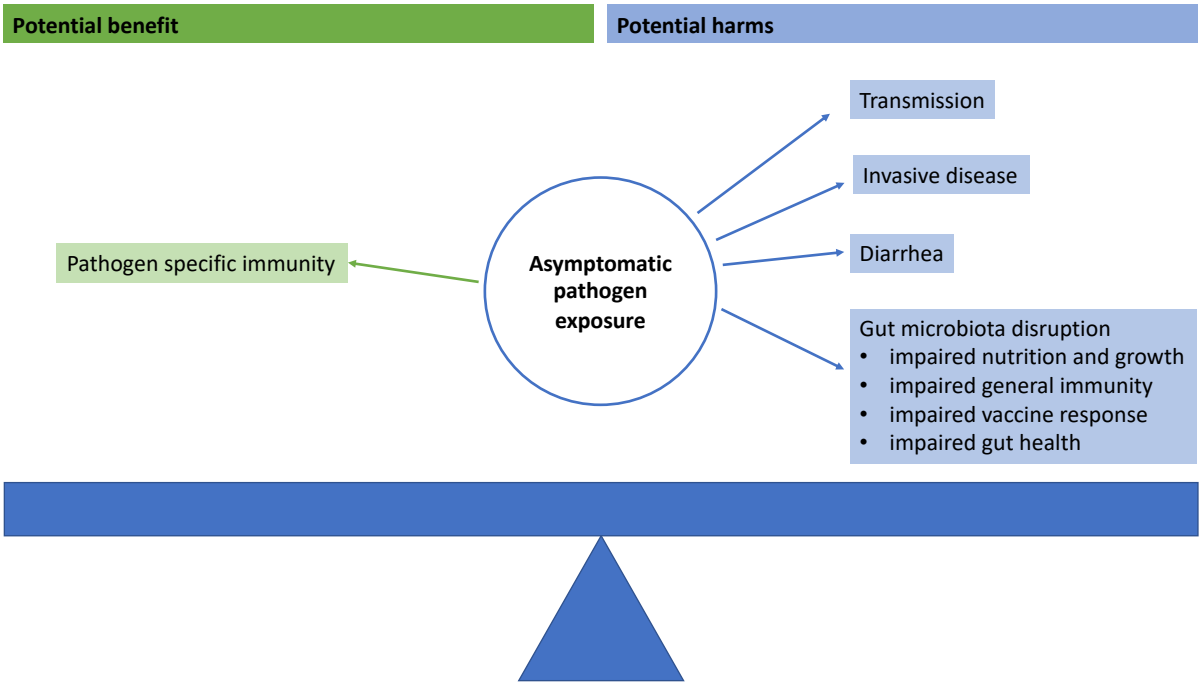

Figure 1.13: Potential benefits and harms of asymptomatic pathogen exposure.

1.6.5 Important Bacterial pathogens - Salmonella

Some of the important bacterial enteric pathogens are *E. coli*, *Shigella, Salmonella*, *Campylobactor* and *Vibrio cholerae*. A more detailed explanation has been given for *Salmonella* compared with other pathogens because it is an important enteric pathogen in SSA and specifically Malawi, as described in section 1.6.3 below.

1.6.5.1 Epidemiology of Salmonella

Salmonellosis is an important cause of morbidity and mortality worldwide, with approximately 1.3 billion cases and about 3 million deaths (GBD, 2019a, 2019b). The causative agent, *Salmonella*, is a Proteobacteria belonging to the Enterobacteriaceae family, and the genus has two species; *Salmonella bongori* and *Salmonella enterica. Salmonella bongori* has one subspecies*, bongori* while *S. enterica* has six subspecies; *S. enterica* subspecies *enterica*, *S. enterica* subspecies *salamae*, *S. enterica* subspecies *arizonae*, *S. enterica* subspecies *diarizonae, S.*

enterica subspecies *hautenae* and *S. enterica* subspecies *indica*, with about 2600 serotypes/serovars (Andino & Hanning, 2015; Lamas et al., 2018). About 99% of the *Salmonella* causing disease in humans and animals belong to the species *Salmonella enterica*. A majority of these serovars have a wide range of hosts where they usually cause gastroenteritis, with only a few serovars that are adapted to the human host (Andino & Hanning, 2015). In humans, *Salmonella* can cause Typhoid fever (TF) or Non-typhoidal Salmonella (NTS) disease. In the case of NTS, the gut can be asymptomatically exposed to NTS (eNTS) or NTS can cross the gut barrier and cause invasive NTS (iNTS). Typhoid fever is caused by Salmonella Typhi, Salmonella Paratyphi A and Salmonella Paratyphi B. All the other *Salmonella enterica* subspecies *enterica* serovars do not cause TF but NTS and are thus all referred to as NTS serovars.

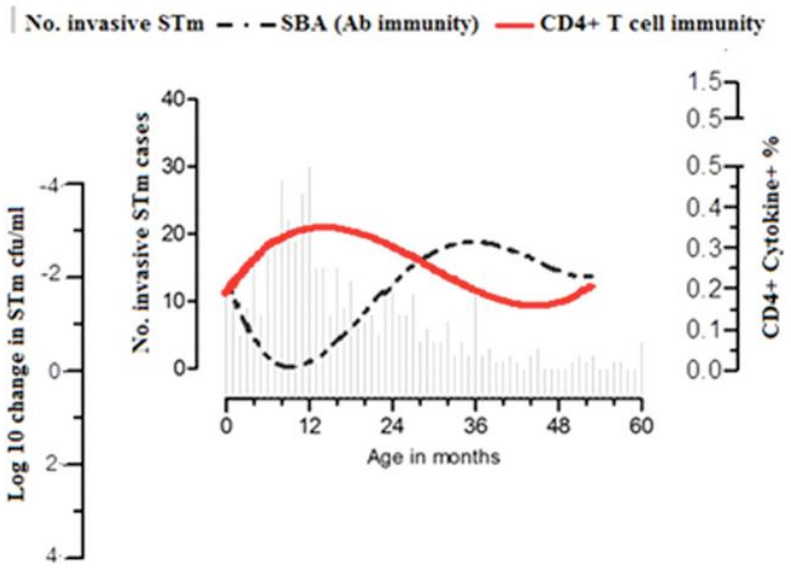

Figure 1.14: S.Typhimurium specific immune development with age.

A sequential acquisition of T cell and antibody immunity to S. Typhimurium in Malawian children. Children acquire S. Typhimurium specific CD4+ cells early in life when exposure to S. Typhimurium is highest. S. Typhimurium specific antibody response develop by 24 months of age.

1.6.5.2 Pathogenesis of Salmonella – Typhoid fever and NTS

Salmonella transmission is through the faecal-oral route. Infection is mainly through the ingestion of contaminated food and water. Infection may also occur from person to person and through direct contact with animals, although these routes are rare (Kariuki et al., 2006; WHO, 2018b). Ingested *Salmonella* colonizes the distal ileum, where they invade the intestinal epithelial cell. *Salmonella* specifically targets the specialized microfold cells (M cells) underlying the Peyer's patches (PP) (Gopinath, Carden, & Monack, 2012). Besides the M cells, *Salmonella* can also infect dendritic and $CD18^{+}$ phagocytic cells. Salmonella Pathogenicity Island 1 (SPI-1) encodes virulence genes that aid the Salmonella to invade the epithelial cells through the

Type III secretion system. Once through the PP M-cells, the bacteria infect phagocytic cells found in the underlying structure of the lymphoid tissue (Kurtz, Goggins, & McLachlan, 2017). The bacteria can migrate to the draining mesenteric lymph nodes through the afferent lymphatics and later invade the bloodstream and other systemic tissues (Johari, Besari, Wan Ghazali, & Yusof, 2019). In the systemic tissue, *Salmonella* persist and proliferate in the liver, bone marrow and spleen phagocytic cells. *Salmonella* uses virulence genes encoded by Salmonella Pathogenicity Island 2 (SPI-2) to evade degradation by macrophages.

1.6.5.3 Invasive Nontyphoidal Salmonella disease

Invasive non-typhoidal *Salmonella* is attributed chiefly to *Salmonella enterica*, subspecies *enterica* Typhimurium (*S.* Typhimurium), and *Salmonella enterica* subspecies *enterica* Enteritidis (*S.* Enteritidis), and usually cause gastroenteritis in healthy individuals but cause BSI in individuals with risk factors such as young age with malnutrition or malaria, and adults with a compromised immune system such as those with HIV/AIDS (Ao et al., 2015; Feasey, Dougan, Kingsley, Heyderman, & Gordon, 2012; Melita A. Gordon, 2008; M. A. Gordon, 2011). These are an important cause of morbidity and mortality, mainly in SSA, where there is a high population of people with risk factors associated with the disease. In 2017, 535,000 cases of invasive disease were estimated to have been caused by iNTS, with the highest incidence occurring in SSA and children under the age of 5 years with the peak of the disease being at 13 months of age in Malawi(GBD, 2019a; Tonney S. Nyirenda et al., 2014). The overall case fatality rate was 14% but it was 41.8% in HIV infected individuals (GBD, 2019b). Studies have provided evidence that supports the hypothesis that iNTS serovars have adapted to the human host (Post et al., 2019).

1.6.5.4 Antibiotic resistance in invasive Nontyphoidal Salmonella disease

Antibiotic resistance is a huge challenge facing the treatment and management of invasive *Salmonella* disease. Resistance to first-line antibiotics has been reported in most SSA countries. An increase in resistant iNTS infections predominantly caused by *S.* Typhimurium has been observed in SSA (Gilchrist & MacLennan, 2019). These isolates are sequence type 313 (ST313). *S.* Typhimurium ST313 isolates are mostly multidrug-resistant (MDR). Studies in Malawi and Kenya have also reported resistance to third-generation cephalosporin ceftriaxone aided by the acquisition of extended-spectrum B-lactamases (ESBLs) in ST313 *Salmonella* isolates (Gilchrist & MacLennan, 2019). Antibiotic resistance has also been reported in *S*. Typhi isolates (Karkey, Thwaites, & Baker, 2018).

1.6.5.5 Vaccines for invasive Nontyphoidal Salmonella disease

There is, however, no licensed vaccine for NTS serovars. Several promising vaccine targets are at an advanced stage of development (Gilchrist & MacLennan, 2019; MacLennan, Martin, & Micoli, 2014; Tennant & Levine, 2015; Tennant, MacLennan, Simon, Martin, & Khan, 2016). The live conjugate vaccines; VD 1931 and CVD 1944, and Glycoconjugate Bivalent OAg-CRM197 are undergoing preclinical trials while OMV-based Bivalent iNTS GMMA and glycoconjugate Trivalent iNTS COPS:FliC – Typbar vaccines are undergoing phase one and two trials. To better understand iNTS disease in preparation for future vaccines the Vac-iNTS consortium that is advancing GMMA-based iNTS vaccine was established. Through this consortium, Phase I trials will be done in Europe and Africa. Malawi, through iNTS research

scientist at MLW is taking part in these studies by conducting serosurvey of iNTS in Africa.

1.6.5.6 Typhoid fever

On the other hand, *Salmonella enterica* subspecies *enterica* Typhi (*S.* Typhi), *Salmonella enterica* subspecies *enterica* Paratyphi A, B and C (*S.* Paratyphi A, B and C) are all human host restricted *Salmonella* serovars causing invasive disease. They are a common cause of enteric fever (typhoid and paratyphoid fever) mainly in East Asia and typhoid fever in SSA where children and adolescents are disproportionately infected . It usually occurs in healthy non-compromised individual. HIV infection which is common is SSA, seem to be protective against clinical presentation of typhoid (Melita A. Gordon, 2008; M. M. Levine & Farag, 2011). Infection may cause serious complications such as intestinal perforations (Khanam et al., 2021).

1.6.5.7 Antibiotic resistance in Typhoid fever

Antibiotic resistance in Typhoid fever infections has been reported in SSA and Asia where the disease burden is high. Outbreaks have been reported in countries such as Malawi, Zimbabwe and Pakistan. In 2017 Pakistan typhoid fever outbreak*, S.* Typhi isolates were resistant to third generation cephalosporins (Shakya et al., 2021). Antibiotic resistance, climate change and urbanization are threats to typhoid fever management and control (Birkhold, Mwisongo, Pollard, & Neuzil, 2021). Thus, heightening the need for an effective vaccine.

1.6.5.8 Vaccines for Typhoid fever

An increase in antibiotic-resistant cases has heightened the need for vaccines against Typhoid fever. There are two licensed vaccines for *S.* Typhi: the oral, live

attenuated Ty21a vaccine and Vi polysaccharide (parenteral) vaccines. Ty21a is licensed for children under the age of 6 years while the Vi polysaccharide vaccine is licensed for children from 2 years of age. Several others are under development (Milligan, Paul, Richardson, & Neuberger, 2018; Neuzil, Pollard, & Marfin, 2019). Notable on the list of types of typhoid vaccine candidates are the Typhoid Conjugate Vaccines (TCV) that have been reported to confer long term protection. Typhoid Vaccine Acceleration Consortium was established following the high immunogenicity report on TCV to facilitate the introduction of TCVs in low and middle income countries. Malawi though researchers at MLW, is part of TyVAC and was the only African study site. Results on the efficacy and safety of the Vi-TCV administered to Malawian children aged 9 months to 12 years of age shows that the vaccine is efficacious and safe (Patel et al., 2021). Nepal TyVAC study reported that there is no evidence that protection conferred by Vi-TV wanes within two years (Shakya et al., 2021).

1.6.6 Important viral pathogens

Rotavirus remains a significant cause of diarrhoeal illnesses in infants and children across the world. In developing countries, mainly in South Asia and SSA, rotavirus is the leading cause of diarrhoeal related deaths (Kotloff, 2017). Estimates of the number of fatalities attributed to rotavirus occurring each year range from 120,000 to 215,000 (Platts-Mills et al., 2018). Despite having low vaccine effectiveness, rotavirus vaccines enrolled in an area with a high disease burden have reduced rotavirus specific and broad range diarrhoeal infections and deaths (Bar-Zeev et al., 2016; Burnett, Jonesteller, Tate, Yen, & Parashar, 2017; Carvalho & Gill, 2019).

Norovirus, enterovirus, adenovirus and sapovirus are also important viral pathogens (Kotloff, 2017).

1.6.7 Important parasitic enteric infections

Parasitic infections caused by protozoan parasites and helminth infections are a common cause of gastrointestinal infection, with an estimated 1.5 billion people inhabiting one or more intestinal parasites (Hotez et al., 2008). Helminths are multicellular organism, while protozoan parasites are unicellular. *Ascaris lumbricoides*, *Trichuuris trichiuria*, *Ancylostoma duodenale* and *Necator americanicus* are the main helminths parasites that inhabit the human gut. The main intestinal protozoan parasites are *Giardia intestinalis*, *Entamoeba histolytica*, *Cyclospora cayetanensis* and *Cryptosporidium* species. Intestinal parasitic infections are more prevalent in tropical regions of the developing world due to the lack of adequate safe water and poor sanitation. The most significant number of intestinal parasites is found in school-aged children and preschool children (Eyamo, Girma, Alemayehu, & Bedewi, 2019; Y. M. Fan et al., 2019). Intestinal parasites significantly impair the nutritional status of these children by causing diarrhoea/ dysentery, impairing the absorption of micronutrients, loss of appetite and loss of blood through bleeding of intestines (Butera et al., 2019; Mekonnen & Ekubagewargies, 2019). Complications include intestinal obstruction and rectal prolapse.

Transmission is through ingestion of not properly prepared food contaminated with eggs or larvae or through skin penetration of infective larvae. In the human intestines, helminths lay thousands of eggs excreted to the environment through the faeces (Hardwick, Werkman, Truscott, & Anderson, 2019). Parasites and helminths

infections are managed by deworming exercises usually carried out in school going children (Farrell, Truscott, & Anderson, 2017; Keiser & Utzinger, 2019).

1.7 Aims and Objectives of the thesis

1.7.1 Main aims

The main aims of this thesis are to

1. Determine the gut microbiota patterns associated with asymptomatic enteric pathogen colonisation of the gut lumen
2. Describe and characterise the early gut microbiota member, *Bifidobacterium*, among Malawian children.

1.7.2 Specific objectives

1. To describe and quantify enteric pathogen exposure events among Malawian children
2. To describe the relationships between gut microbiota composition and asymptomatic exposure to enteric pathogens among healthy Malawian children
3. To isolate and describe *Bifidobacterium* isolated from heathy exclusively breastfed Malawian infants
4. To investigate anti-*Salmonella* properties of *Bifidobacterium* strains isolated from Malawian children

Chapter 2: Materials and Methods

2.1 Study type and site

This thesis presents data from samples that were previously collected in a longitudinal cohort study investigating ***Salmonella*** **exposure and development of immunity in Malawian children (SalExpo study)** in Chapters 3 and 4, and a single centre cross-sectional pilot study **Exploiting anti-infection activities of the early life gastrointestinal tract microbiota member *Bifidobacterium* against Bacterial enteric infections (BIFBAC study) in** Chapters 5 and 6.

The longitudinal cohort study was conducted in 2013 at Zingwangwa Health Centre. The cross-sectional study was conducted in 2018 at Ndirande Health Centre. Ndirande Health Centre and ZHC are clinics in Malawi that operate under Blantyre health district. Ndirande Health Centre and ZHC are about 6 and 5 km from Queen Elizabeth Central Hospital (QECH), the country's main referral hospital. Malawi Liverpool Wellcome Trust Clinical Research Program, where initial sample processing for both studies was done, is based within QECH premises.

2.2 Ethical approval

Ethical approval for SalExpo study was granted by College of Medicine Research and Ethics Committee (P.01/13/1327). The study was compliant with Good Clinical Practice and was conducted according to the Helsinki declaration (World-Medical-Association, 2013). Written informed consent was obtained from mothers/guardians of all study participants before their children were recruited into the study.

Ethical approval for the BIFBAC study was granted by the College of Medicine Research and Ethics Committee (P.10/17/2296) and the University of Liverpool Research and Ethics committee (Reference number: 2867). The study was compliant with Good Clinical Practice and was conducted according to the Helsinki declaration (World-Medical-Association, 2013). Written informed consent was obtained from mothers/ guardians of all study participants before their children were recruited into the study.

Table 2.1: Basic study information for SalExpo and BIFBAC studies

	SalExpo study	**BIFBAC study**
Full study title	***Sa****lmonella* **expo**sure and development of specific immunity in Malawian children	Exploiting anti-infection activities of the early life microbiota member ***Bif****idobacterium* against **Bac**terial enteric infections
Year of recruitment	2013	2018
Study site	Zingwangwa Health Centre, Blantyre, Malawi	Ndirande Health Centre, Blantyre, Malawi
Ethical approval	COMREC - P.01/13/1327	COMREC - P.10/17/2296 UoL – Reference number 2867
Study design	longitudinal cohort study	single centre cross-sectional pilot study
Sample size	60 participants with 630 samples collected longitudinally	30 participants

Inclusion criteria	• A child who is 6 months old
	• healthy at the time of recruitment
		• A child who is less than or equal to 18 weeks old
		• Exclusively breastfed
		• Healthy at the time of recruitment
		  ○ S/he has no fever
		  ○ S/he has no diarrhoea

Exclusion criteria	• Children born preterm
	• Known HIV-positive children
	• HIV exposed children
	• Children presenting with acute illness
	  ○ Fever > 38^{o}C
	  ○ diarrhoea
	• Children residing outside the study's geographical area
		• Exposure to HIV
		• Preterm birth
		• History of diarrhoea within the past 1 month

2.3 Clinical procedures for SalExpo study

2.3.1 Participant recruitment, follow-up, sample collection and processing for Salmonella asymptomatic children from 6 to 18 months

For the longitudinal study, consenting, recruitment and sample collection was previously completed (Tonney Stophen Nyirenda, 2015). Briefly, healthy children were recruited into the study based on the inclusion and exclusion criteria in **Table 2.1**. Children were recruited at the age of 6 months and followed up each month for one year until 18 months **(Figure 2.1).** Demographic data, clinical data, stool and blood samples were collected at the time of recruitment. Stool samples and clinical data were also collected every month. Stool samples were separated into 3 aliquots and stored at -80^{o}C. DNA extraction from the stool samples, with bead-beating, was done within one week of sample collection so that samples were optimal for 16S microbiota assessment. For other analyses, stool samples were stored at -80^{o}C until the time of the test.

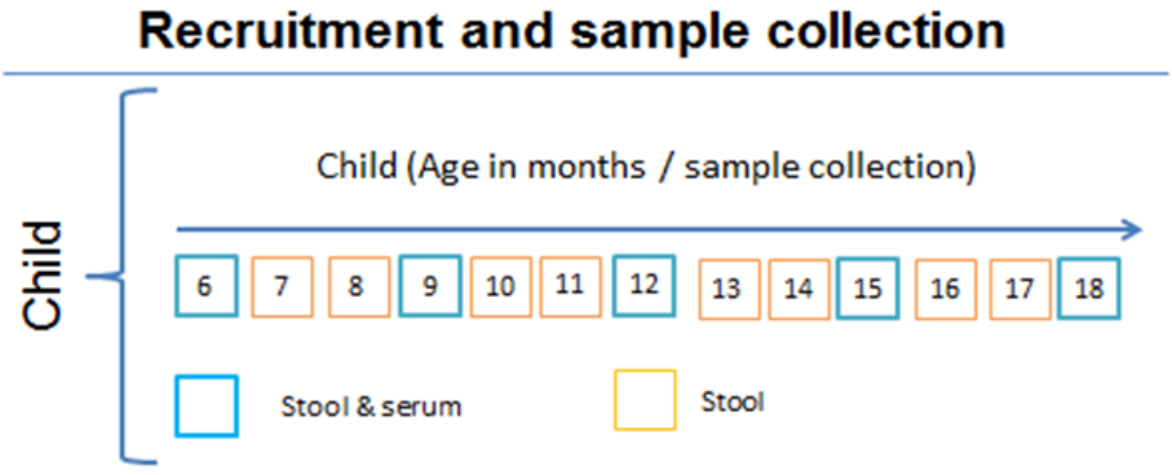

Figure 2.1: Recruitment and sample collection for SALEXPO study

Adapted from (Tonney Stophen Nyirenda, 2015)

2.4 Clinical procedure for BIFBAC study

2.4.1 Consenting and recruitment of healthy exclusively breastfed children

Participant engagement was conducted among parents or guardians who brought their children for routine under five antenatal by a study nurse who had undergone training on the study protocol and good clinical practice. Interested parents or guardians of the infants, who were sensitised about the study and its aims and objectives, had a private and detailed study discussion before consenting to join the study. Only children whose parents consented to have their children take part in the study were recruited. Thirty healthy exclusively breastfed infants were recruited, with informed consent, according to the inclusion and exclusion criteria in Table 2.1.

2.4.2 Stool sample collection

The same procedure for stool sample collection was used in the 2 studies reported here. The study nurses collected faecal samples from the infant's nappies using a sample bottle and collection scoop. Parents/guardians were trained by the study nurses on how to collect the samples whenever a sample could not be provided at the time of recruitment. Mothers were requested to collect the stool samples during the day. Samples collected from home were brought to the clinic within 4 hours of sample collection. Samples were labelled with a sample identification number (sample ID), date of collection, and collection time. The bottle was tightly closed, then placed in a biohazard bag in a cooler box with icepacks, then transferred to MLW laboratory for processing the same day.

2.5 Laboratory procedures for SalExpo study

The laboratory procedures carried out under the longitudinal study involved 16S rRNA sequencing for gut microbiota profiling, Taqman Array Card (TAC) assay, using total nucleic acid (TNA), for detection of multiple enteric pathogens and *Salmonella* detection by stool culture and qPCR. **Figure 2.2** is a flow diagram of all the procedures.

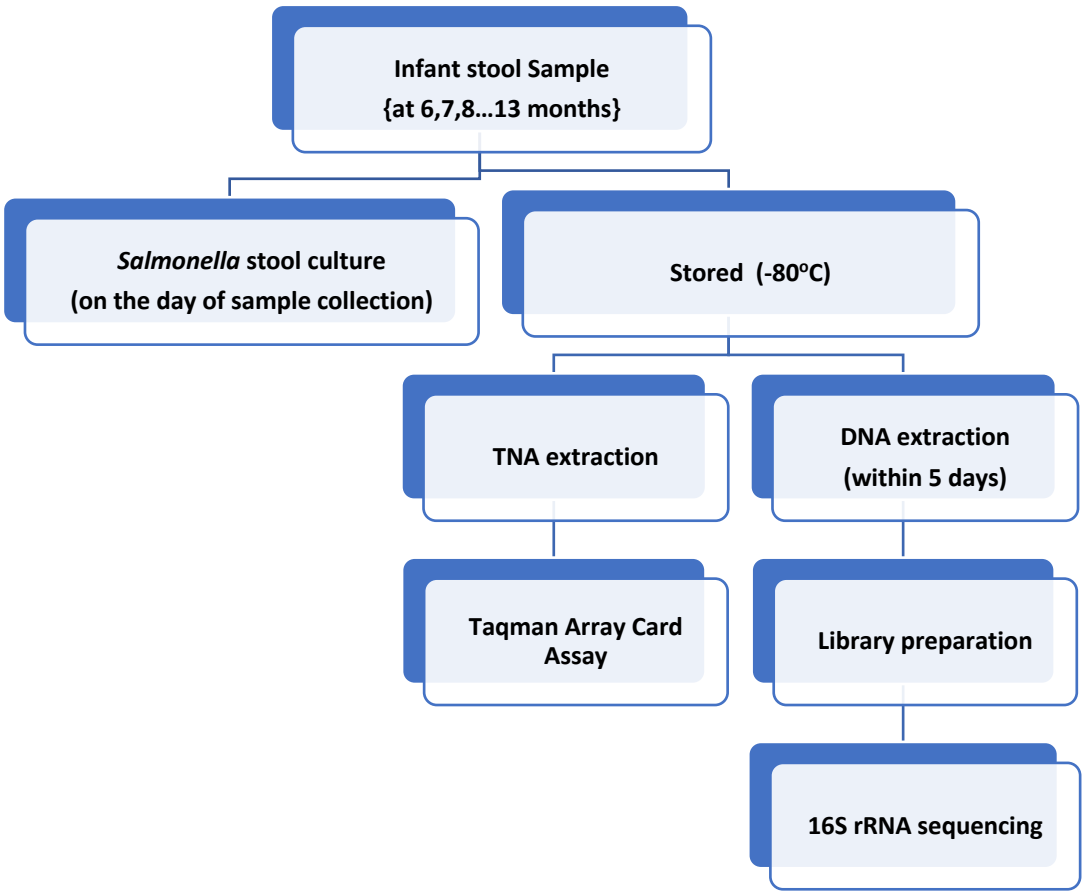

Figure 2.2: Flow chart of activities for the longitudinal cohort study.

Total Nuclei Acid (TNA), Deoxyribonucleic acid (DNA) ribosomal ribonucleic acid (rRNA)

2.5.1 Stool Total Nucleic Acid extraction for Taqman Array Card Assay

According to manufacturer instructions, total nucleic acid (TNA) was extracted from stool samples using QIAamp Fast DNA Stool Mini Kit, with some modifications. Two main changes to the manufacturers' procedure were; the addition of a bead beading step for mechanical disruption of the faecal sample; and the addition of Phocine Herpesvirus (PhHV) and MS2 extrinsic controls, which were used to evaluate the efficiency of extraction and amplification procedures.

About 200 mg of stool stored at -70^{o}C was aliquoted into 2 ml screw cup tubes compatible with the Qiagen Tissue Lyser machine. 200 µl of distilled water was used as a negative control for each extraction batch. To the stool, 390 mg of sigma beads were added. This was followed by the addition of 1 ml inhibitEX buffer with PhHV and MS2.

InhibitEX buffer and extrinsic control mixture were prepared as follows: InhibitEX was first checked for precipitate. In case of a precipitate, the buffer was incubated at 37^{o}C for 5 to 15 minutes, depending on how quickly the precipitate disappeared. Buffer was then mixed thoroughly. For each sample, 1 ml of inhibitEX buffer was mixed with 1 µl of PhHV and 1 µl of MS2. Reagents for N +1 samples were prepared to account for pipetting error. For 10 samples, 11 ml inhibitEX buffer, 11 µl of PhHV and 11 µl of MS2 were mixed.

The tube containing stool sample, beads and inhibitEX buffer was then vortexed for 1 minute, followed by bead beating using a Tissue Lyser machine (QIAGEN, Netherlands). Bead beating was done at 20,000 g for 1 minute followed by 5 minutes resting time. For each sample, the bead beating process was done 3 times. After the three bead-beating steps, samples were incubated at 95^{o}C for 5 minutes using a heat

block. This was followed by vortexing for 15 seconds and centrifuged at 20,000 g for 1 minute to pellet the faecal material. 25 µl of proteinase K was aliquoted into a clean 2 ml microcentrifuge tube. 600 µl of the stool supernatant was added to the 25 µl proteinase K. To this mixture, 600 µl of Buffer AL was added, followed by vortexing for 15 seconds to form a homogeneous solution. The mixture was then incubated at 70^{o}C for 10 minutes using a heat block. Samples were then centrifuged for 30 seconds to remove drops on the lids from the incubation step. 600 µl of 96 - 100 % ethanol was then added to the supernatant and buffer AL mixture, followed by vortexing and brief centrifuging. Spin columns with labelled lids were placed in 2 ml collection tubes, 600 ml of the lysate was added to the spin column. The tubes were then centrifuged at 20,000 g for 1 minute. The spin column was retained and placed in a new collection tube while the used collection tube with filtrate was discarded. This step was repeated until all the lysate from one sample (1800 µl) had gone through its labelled spin column. The spin-column was then placed in a new collection tube. To denature proteins, 500 µl of reconstituted first wash buffer - Buffer AW1, was added to the column, followed by centrifuging at 20,000 g for 1 minute. The spin column was then placed in a new collection tube whilst the collection tube with filtrate was discarded. This was followed by adding 500 µl of reconstituted second wash buffer – Buffer AW2, followed by 3-minute centrifuging at 20,000 g. Buffer AW2 contains 70% ethanol and is used to wash out salts.

The spin column was retained while the used column was discarded. The spin column was placed in a new collection tube and centrifuged without adding anything to it for 3 minutes at full speed. This was to avoid Buffer AW2 being carried over to subsequent steps. To finally elute the TNA, the spin column was transferred into a new microcentrifuge tube labelled with sample ID, sample type, and extraction date. 200

µl of ATE buffer was then directly added to the membrane on the QIAamp spin column, followed by 3 minutes incubation at room temperature. After incubation, the TNA was eluted by centrifuging at 20,000 g for one minute. Total nucleic acid was stored at 80ºC.

2.5.2 Enteric Taqman Array Card assay

The extracted TNA was used to detect enteric pathogens using a customised enteric TAC assay (J. Liu et al., 2013). A TAC assay is a semi-quantitative 384 well real-time PCR for detection of multiple enteric pathogens (J. Liu, Kabir, F, Manneh, J et al. , 2014). The TAC cards used in this study was semi-custom made and could test up to 8 samples, detecting 29 different enteric pathogens using 48 targets. Each of the 384 wells contains Primers and probes for different targets. Targets on the enteric TAC include bacteria, viruses, fungi, protozoa, and helminths, as shown in **Figure 2.3.** In addition to the mentioned targets, the TAC card also has 16S, MS2, and phocine herpes Virus (PhHV) targets as nucleic acid extraction and amplification controls.

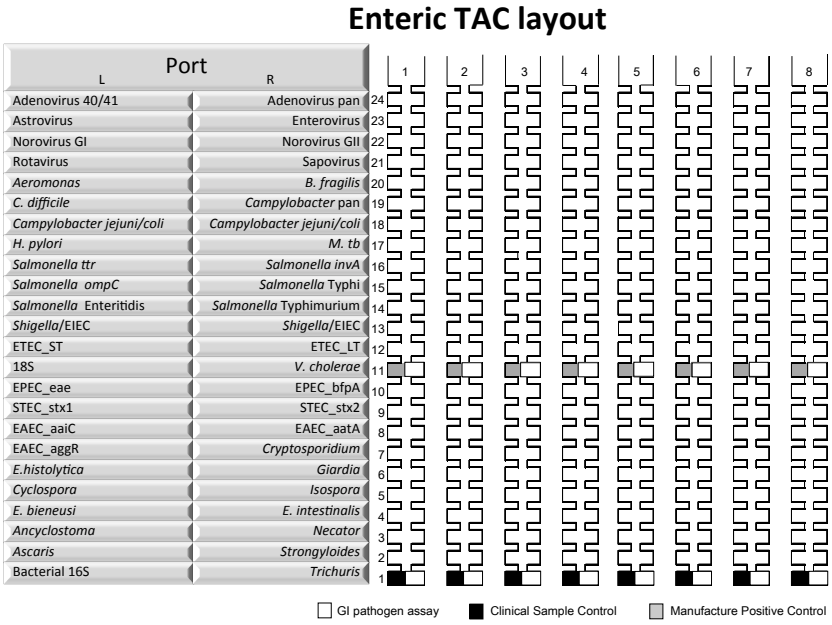

Figure 2.3: Layout of the enteric Taqman Array Card used for the detection of different enteric pathogens.

2.5.2.1 Procedure for TAC

Pathogen detection was conducted according to the manufacturers' instructions, with slight modifications. Master mix was prepared for eight samples by adding Ag-Path-ID 2X RT-PCR buffer (425 µL), Ag-Path-ID Enzyme mix (34 µL) and nuclease-free water (221 µL). Eighty microliters of the master mix was aliquoted into eight 2 ml microcentrifuge tubes. To each tube, 20 µl of TNA or nuclease-free water for negative technical control was added. The tubes were vortexed and centrifuged to remove air bubbles. One hundred microliters of the PCR reaction mix for each sample was loaded into appropriate wells on a TAC card that had reached room temperature. After loading, the card was centrifuged for 1 minute at 331 g. Centrifuging was done two times to allow the samples to fill all the respective wells containing primers and probes for different targets. The card was visually examined to ensure that the wells had been appropriately filled; a small uniform volume of sample was expected to remain in the loading port if the samples were filled correctly. After loading the plate with the reaction mix, a sealer was used to seal the card. Then the assay was run on Vii7A PCR machine (Life technologies) under the conditions indicated in **Table 2.2** once the plate was sealed correctly:

Table 2.2: TAC PCR conditions

Step	**Temperature**	**Time**
RT (1 cycle)	45°C	20 minutes
Denaturation (1 cycle)	95°C	10 minutes
PCR (40 cycles)	95°C	15 seconds
	60°C	1 min (data collection)

A run was deemed successful when:

1. The negative test control reactions did not exhibit fluorescence curves (amplification plots) that crossed the threshold line.
2. Positive control targets; MS2 and PhHV resulted in reaction fluorescence curves that crossed the threshold for each target assay.

A sample was deemed positive when an appropriately shaped amplification curve crossed the threshold, and both positive and negative controls worked properly as above.

2.5.3 Salmonella detection using stool culture and Real-Time PCR

In addition to the TAC assay, *Salmonella* detection was also done using stool culture on the day of sample collection using routine MLW laboratory stool culturing procedure and an in-house qPCR using DNA extracted from frozen stool samples. For the qPCR, TTR and InVA, primers and probes were used (**Table 2.3**) (Chirambo AC, 2020).

Table 2.3: Primers and probes used for molecular detection of Salmonella in stool specimens

	Primer name	Primer direction	Primer code/Probe description
1	*INVA*	Forward	5'-AGCGTACTGGAAAGGGAAAG-3'
2	*INVA*	Reverse	5'-CACCGAAATACCGCCAATAAAG-3'
3	*INVA*	Probe	FAM-TTACGGTTCCTTTGACGGTGCGAT-BHQ1
4	*ttr*	Forward	5'-CTCACCAGGAGATTACAACATGG-3'
5	*ttr*	Reverse	5'-AGCTCAGACCAAAAGTGACCATC-3'
6	*ttr*	Probe	FAM-CACCGACGGCGAGACCGACTTT-BHQ1

2.5.4 Sample processing for 16S rRNA sequencing

For microbiome work, DNA extraction was previously conducted on fresh stool samples. QIAamp Fast DNA Stool Mini Kit (QIAGEN, Netherlands) was used according to the manufacturers' instruction with an added bead beating step. DNA was stored at -80^{o}C until it was shipped to Wellcome Trust Sanger Institute (WTSI), where library preparation was done. The V1V2 region of the 16S rRNA genome was targeted with a 300 base pair (bp) read length. Mock community A was included in the library preparation step as a sequencing control (bei genomic DNA resources - catalogue No. HM-276D). The prepared libraries were used for sequencing using the Illumina Miseq™ platform.

2.6 Laboratory procedures for BIFBAC study

As in **Figure 2.4** below, laboratory procedures for the BIFBAC study involved

1) Isolation of bifidobacteria by culture and testing of growth characteristics
2) Identification of *Bifidobacterium* by using WGS
3) Determination of the gut microbiota composition by 16S rRNA sequencing

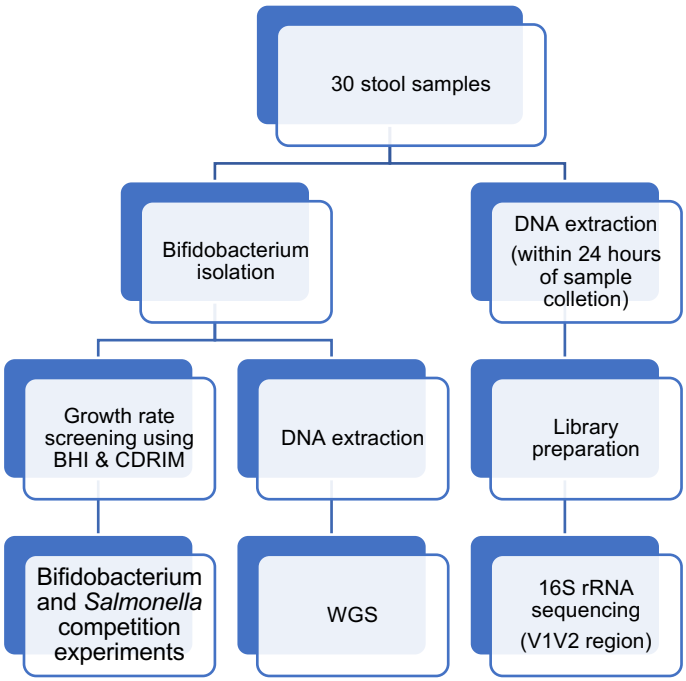

Figure 2.4: Flow chart of activities for the BIFBAC project

Stool samples collected from exclusively breastfed children were used to isolate *Bifidobacterium*. Isolated bifidobacteria was first confirmed using Fructose-6-phosphate Phosphoketolase biochemical assay and 16S Sanger sequencing before conducting whole genome sequencing (WGS).

2.6.1 Isolation of bifidobacteria

Bifidobacteria isolation was done using de Man, Rogosa and Sharpe (MRS) selective media containing 50 mg/litre Cysteine and Mupirocin. Each batch of plates was cultured with sterile PBS as a negative control. Two different known *Bifidobacterium* strains donated by the Lindsay Hall Laboratory at the QIB were used as positive controls. Samples were brought to the laboratory using cooler boxes within 4 hours of being received in the field. All work was done in a level II safety cabinet. Briefly, 100 mg of faecal sample was diluted in 900 µl of sterile PBS. From this dilution, four 10-fold serial dilutions were made. A total of five dilutions (10^{0}, 10^{-1}, 10^{-2}, 10^{-3} and

10^{-4}) were made from each single sample. 100 µl from each dilution was aliquoted onto a media plate and inoculated by a spreader across the whole plate. Inoculated plates were incubated anaerobically at 37°C for 36 hours. Plates were then assessed for bifidobacteria growth. Three single colonies were picked from each culture plate with a presumptive bifidobacteria growth. These colonies were sub-cultured on 3 MRS agar plates and incubated anaerobically at 37°C for 36 hours. Presumptive bifidobacteria colonies were repeatedly sub-cultured until pure colony growth was achieved. On average, this was achieved after five sub-cultures. After achieving pure growth of presumptive bifidobacteria, one pure colony from each culture plate was sub-cultured in MRS broth and incubated for 36 hours. After 36 hours of incubation, the MRS broth containing bifidobacteria was centrifuged at 4000 rpm for 10 minutes at level 9 and 3 of acceleration and brakes, respectively. The supernatant was discarded. The pellet which contained the bacteria was resuspended in Reinforced Clostridial Medium with 30% glycerol and kept as stock. The prepared stock was aliquoted into screw-top tubes and stored at -80°C.

2.6.2 Bifidobacteria biochemical confirmatory test: Fructose-6-phosphate Phosphoketolase assay

To confirm at the bench that the isolated organisms were bifidobacteria, a Fructose-6-phosphate Phosphoketolase assay (F6FPK) was conducted. This bench assay was used to identify isolates for subsequent DNA extraction and genomic analysis The F6FPK assay is the definitive assay used to identify *Bifidobacterium* (Orban & Patterson, 2000). The test was done using bacteria grown anaerobically from frozen stocks in 20 ml MRS broth at 37°C for 36 hours. *Lactobacillus* culture and distilled water were used as negative controls, while a known bifidobacteria strain obtained

from the Hall laboratory was used as a positive control. The liquid culture was centrifuged for 10 minutes at 4000 rcf, room temperature, acceleration and brakes level 9 and 3, respectively. The supernatant was discarded. The pellet was washed twice with 10 ml Phosphate buffer at each time and centrifuged at the same conditions but for 15 minutes. After centrifuging two times, the sample was incubated for 5 minutes in 4 ml of 180 ug/ml of cetrimonium bromide. After the 5 minutes incubation, 0.25 ml of a mixture of sodium fluoride solution (NaF 3 mg/ml) and sodium iodoacetate (5 mg/ml) in water was added to the sample. 0.25 ml of sodium fructose-6-phosphate (80 mg/ml in water) was also added to the sample. This was followed by vortexing and 30 minutes incubation at 37^{o}C. After incubation, 1.5 mls of hydroxylamine hydrochloride (13 g/100 ml) was added to the sample, followed by vortexing and 10 minutes incubation at room temperature. Finally, 1 ml of TCA, 1 ml OF 4N HCL and 1 ml of ferric chloride were added to the sample. The sample was vortexed, and colour change was observed. A positive result was defined as the colour changed from milky to dark red or brown, while a negative result was defined as no colour change.

2.6.3 Bifidobacteria confirmatory test: 16S rRNA Sanger sequencing

For final confirmation of *Bifidobacterium* isolates, DNA samples extracted from the isolates were sent to Eurofins for 16S rRNA Sanger sequencing. At Eurofins, sequences were generated using Big Dye terminator chemistry (version 3.1) following their standard protocol. Sequencing reaction cleanup was done on a Hamilton Starlet robotic workstation (Hamilton Robotics GmbH, Martinsried, Germany) by gel-filtration through a hydrated Sephadex matrix filled into appropriate 96 well filter plates followed by a subsequent centrifugation step. Finally, all reactions were run on ABI3730xl capillary sequencers equipped with 50 cm capillaries and

POP7 polymer (Thermo Fisher Scientific, Waltham, MA USA). To identify the isolated organisms, sequence reads received from Eurofins in fasta file format were blasted against the NCBI database using default settings.

2.6.4 Bacteria DNA extraction for whole genome sequencing

To characterise the isolated *Bifidobacterium*, DNA was manually extracted, quantified using Qubit, then sent for WGS at Wellcome Trust Sanger Institute (WTSI). DNA extraction for WGS was done only on confirmed *Bifidobacterium* isolates. The phenol Chloroform DNA extraction method described by Del Pickard and adopted at QIB – Hall lab was used.

A single colony of pure *Bifidobacterium* was inoculated in 20 ml of MRS broth and incubated anaerobically at 37^{o}C for 36 hours. *Bifidobacterium* cells were obtained by spinning down the *Bifidobacterium* culture at 4000 rpm for 10 minutes and discarding the supernatant. Cells were washed twice using 30 ml of PBS. After washing, the cells were resuspended in 2 ml of 25% sucrose in TE (25 g Sucrose in TE consisting of 10 mM Tris and 1 mM EDTA pH 8.0). 50 μl of 100 mg/ml Roche Lysozyme (in 0.25 M Tris pH 8.0) was added to the cells. Cells were incubated at 37^{o}C for 1 hour or until the bacterial suspension became viscous due to cell lysis. After incubation, 100 μl of 20 mg/ml of Proteinase K solution was added. This was followed by adding 30 μl of 10 mg/ml of RNase A, 400 μl of 0.5M EDTA and 250 μl of fresh 10% Sarkosyl NL30. This was left on ice for 1 to 2 hours or until lysis was complete (slightly turbid but viscous or almost transparent but should still be viscous). Cells were then left to incubate overnight in a water bath which was at 50^{o}C. After the

overnight incubation in a water bath, an appropriate volume of E buffer was added to fill the cell mixture to 5 ml.

From this point, subsequent DNA extraction steps involved the use of Phenol-Chloroform. These steps were done in a fume cupboard whilst putting on the appropriate personal protective equipment (PPE), which included safety glasses to ensure health and safety.

The DNA sample mixture was mixed with 5 mls Phenol:Chloroform:Isoamyl Alcohol (25:24:1) in a phase lock tube (PLG; Qiagen MaXtract High Density, Qiagen cat no 129065). The mixture was centrifuged at 8000 g for 15 minutes to separate the phases. After separating the phases, the aqueous phase was transferred into a new fresh PLG tube and 5 mls of Phenol:Chloroform: Isoamyl Alcohol was added. This was then centrifuged at the same speed but for 10 minutes to separate the phases. This step was repeated one more time with 5 minutes for centrifuging. After three extractions with 5 ml Phenol:Chloroform:Isoamyl Alcohol, the aqueous phase was transferred into a fresh PLG tube. Extraction with Chloroform: IAA (24:1) only was done twice by adding 5 ml of Chloroform: IAA and centrifuging at 4000 rpm for 5 minutes. The final aqueous phase was transferred into a 15 ml Falcon tube. Absolute ethanol was added in a ratio of 5:2 of ethanol and DNA mixture. Ethanol was used to precipitate the DNA. DNA ethanol mixture was incubated at -20^{o}C for 15 minutes. At this stage, DNA was seen having a cotton wool effect in some of the samples. DNA was collected by spinning down the DNA ethanol mixture at 8000 g for 10 minutes without breaks. The pellet was washed twice in 10 ml of 70% ethanol. After washing, the DNA pellet was dried overnight at room temperature. Dried DNA was resuspended

in 250 – 300 μl of E buffer. DNA was then quantified as previously described using a Qubit machine. Samples with DNA concentrations higher than two ng/μl were sent to WTSI, where they were subjected to library preparation followed by WGS using Illumina Hiseq 2500 platform.

2.6.5 Stool DNA extraction for microbiome work

To determine the gut microbiota profiles for stool samples collected from these exclusively breastfed children, we extracted whole stool DNA, prepared the libraries and sent the prepared libraries to WTSI for 16S rRNA sequencing. DNA extraction was done using the MPBio fast DNA kit for soil as described above and according to the manufacturer's instruction with minor modifications. 200 mg of faecal sample was used for DNA extraction and distilled water as technical negative control. DNA was stored at -20^{o}C until the time of library preparation.

2.6.6 Library preparation for 16S rRNA sequencing

Library preparation involved PCR amplification of the V1V2 region of the 16S rRNA gene, cleaning the amplified products, quantifying the PCR products, and pooling the PCR products into libraries.

2.6.6.1 PCR amplification

DNA extracts were first tested for quantity and quality using Qubit 2.0 dsDNA BR assay (Invitrogen) and NanoDrop 1000 (Thermo Scientific), respectively. The DNA concentration was adjusted up or down based on the initial concentration to achieve a concentration of approximately 2.5 ng/μl. A ratio of absorbance of 1.8 to 2.0 was required for the absorbance measurement taken at 260 nm to 280 nm. After adjusting

for quantity, the stool DNA was PCR amplified, targeting the V1V2 region using 16S rRNA forward and labelled reverse primers from Integrated DNA Technologies (IDT), Iowa, United States of America. There were eight indexed forward primers labelled 501 to 508 and 12 indexed reverse primers labelled 701 to 712 **(Table 2.4).**

Table 2.4: Forward and reverse primers for V1V2 16S rRNA gene

9 forward (FW) primers, 12 reverse (RV) primers	Total Primer sequences
V1FW_SA501	AATGATACGGCGACCACCGAGATCTACACATCGTACGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD501	AATGATACGGCGACCACCGAGATCTACACAAGCAGCATATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD502	AATGATACGGCGACCACCGAGATCTACACACGCGTGATATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD503	AATGATACGGCGACCACCGAGATCTACACCGATCTACTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD504	AATGATACGGCGACCACCGAGATCTACACTGCGTCACTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD505	AATGATACGGCGACCACCGAGATCTACACGTCTAGTGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD506	AATGATACGGCGACCACCGAGATCTACACCTAGTATGTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD507	AATGATACGGCGACCACCGAGATCTACACGATAGCGTTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V1FW_SD508	AATGATACGGCGACCACCGAGATCTACACTCTACACTTATGGTAATTGTAGMGTTYGATYMTGGCTCAG
V2RV_SD701	CAAGCAGAAGACGGCATACGAGATACCTAGTAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD702	CAAGCAGAAGACGGCATACGAGATACGTACGTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD703	CAAGCAGAAGACGGCATACGAGATATATCGCGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD704	CAAGCAGAAGACGGCATACGAGATCACGATAGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD705	CAAGCAGAAGACGGCATACGAGATCGTATCGCAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD706	CAAGCAGAAGACGGCATACGAGATCTGCGACTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD707	CAAGCAGAAGACGGCATACGAGATGCTGTAACAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD708	CAAGCAGAAGACGGCATACGAGATGGACGTTAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD709	CAAGCAGAAGACGGCATACGAGATGGTCGTAGAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD710	CAAGCAGAAGACGGCATACGAGATTAAGTCTCAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD711	CAAGCAGAAGACGGCATACGAGATTACACAGTAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT
V2RV_SD712	CAAGCAGAAGACGGCATACGAGATTTGACGCAAGTCAGTCAGCCGCTGCCTCCCGTAGGAGT

The starting stool DNA template volume was 2.5 µl. Each DNA sample was amplified in triplicates (96 samples in two PCR plates) and was pooled together during magnetic bead clean up. One PCR master mix was made which covered 192 (96 x 2) samples plus 10% master mix overhang. Master mix was made by mixing nuclease-free water, 5x Q5 PCR buffer, 10 mM dNTPs, Q5 Taq PCR enzyme, forward and reverse primer as detailed in **Table 2.5.**

Table 2.5: Preparation of PCR master mix

Reagents	Master Mix (n= 1)	Master Mix (n= 212)
Nuclease free water	14.25 µl	3021
5x Q5 PCR Buffer	5 µl	1060
10 mM dNTPs	0.5 µ	106
Q5 Taq PCR enzyme	0.25 µ	53
Total volume in Master mix	**20**	**4240**
Ten µM forward primers (500 numbered primers)	Added individual 1.25µ	
Ten µM reverse primers (700 numbered primers)	Added individual 1.25	
Template DNA	Added individual 2.5	
Total volume	**25**	

The master mix was transferred to each well in the 96 well PCR plate, then the forward indexed and reverse indexed PCR primer individually added to the wells. The following PCR conditions were used:

PCR cycle for 16S amplification

Start and Taman polymerase activation	98°C for 2 minutes
Denaturing	98°C for 30 seconds
Annealing	50°C for 30 seconds
Extension	72°C for 90 seconds
Number of cycles	20
Final extension	72°C for 5 minutes
Hold	10°C

Twenty PCR cycles were used to minimise nonspecific DNA amplification, which may lead to background contamination.

2.6.6.2 Clean up PCR products using Agencourt AMPure XP beads

After PCR amplification, the product was cleaned and concentrated using magnetic beads. Briefly, centrifugation of the PCR plate was done at 1,000 g at 20°C for 1 minute to collect condensation. The triplicate PCR products were combined into a new 96 well PCR plate. AMPure XP beads brought to room temperature were vortexed for 30 seconds to ensure that the beads were evenly dispersed. Beads were added to appropriate wells using the same volume (70 μl) of PCR product. Beads were mixed with the PCR product by gently pipetting the entire volume up and down at least ten times. The amplicon PCR plate with the samples was then incubated at room temperature for 5 minutes without shaking. This was then placed on a magnetic stand for 2 minutes or until the supernatant cleared. The supernatant was removed with the amplicon PCR plate on the magnetic stand. With the amplicon PCR plate still on the magnetic stand, the beads were washed with freshly prepared 80% ethanol by adding 200 μl of freshly prepared 80% ethanol to each sample well followed by 30 seconds incubation. Once the beads were separated from the supernatant, the supernatant was carefully removed and discarded. The wash step was done twice. After the second wash, any remaining supernatant was removed using a multichannel pipette using smaller volume and fine pipette tips. With the amplicon PCR plate still on the magnetic stand, the beads were air-dry for 15 minutes. The amplicon PCR plate was then removed from the magnetic stand. Using a multichannel pipette, 27.5 μl of 10 mM Tris pH 8.5 was added to each well of the amplicon PCR plate followed by up and down pipette mixing, changing tips after each column. Once the beads were fully

suspended, the plate was left to incubate for 2 minutes. The plate was then placed back on the magnetic stand for 2 minutes or until the supernatant had cleared. Using a multichannel pipette, 25 μl of the supernatant was carefully transferred from the amplicon PCR plate to a new 96-well PCR plate. According to the manual, the quality of the cleaned PCR product was checked by randomly selecting ten samples and testing them using nanodrop.

2.6.6.3 PCR product quantification and construction of library equivalent molar (equimolar) mix

Before generating a library equimolar mix, the cleaned PCR amplicons were quantified using Qubit. Determining the cleaned amplified DNA concentration is crucial in determining the actual volumes to be pooled together to form a library.

To determine the sample volume required to be pooled together for the library, the highest and lowest sample concentration in the library was determined; these were used to decide the amount of DNA from each sample to be added to make up the final equimolar mix. The required volume for each sample was calculated using the following formula: Volume (μl) = Amount of DNA in the most concentrated sample (ng) / Individual sample concentration (ng/μl). For example, if the highest concentration was 80 μg/ml and the lowest concentration was 2 μg/ml, then at least 0.4 μl from the highest concentration (40 ng) and 16 μl (40 ng) from the lowest concentration was used.

Highly concentrated samples were diluted with molecular grade water to a factor that enabled the inclusion of the samples with low concentration. After dilution, the samples

were quantified again using the same Qubit working reagent. About 20 to 30% extra Qubit working reagent was prepared for highly concentrated samples that required dilution and re-quantification.

Once all samples had known concentration and the required volume to be pooled together had been calculated, we prepared a PCR tube for each set of 12 samples (e.g. samples A1 - A12 tube 1, samples B1 - B12 tube two etc.). The previously calculated required volume from Samples was pooled into the appropriate tubes (required volume from samples A1 - A12 and B1 - B12 were pulled into Tube 1 and 2, etc.). This step was set as a control measure; in case of any error with one set of samples, only that set would be discarded. Samples from the respective tubes were then pooled into one final PCR tube, which formed the equimolar mix. The equimolar mix was vortexed, and the DNA concentration was quantified using Qubit. This was compared to the expected concentration, which was calculated as follows:

Concentration (ng/ml) = (DNA amount x No. of Samples) /

Total of the Sample volumes added to Equimolar Mix

The equimolar mix was split into two labelled tubes; 1 tube was stored at -20°C as a backup, the other was cleaned up before submitting for sequencing.

2.6.6.4 Library (equimolar mix) clean up using gel purification

Gel purification of the equimolar mix was conducted to remove primer dimers and contaminants. A 1% agarose gel and a 100 bp ladder was used to run out the equimolar library mix. The appropriately sized band (400 to 500 bp for the V1V2 16S region) was cut from the gel with a sterile scalpel and was purified using the Wizard SV gel and PCR clean up kit by Promega. To retrieve the cleaned DNA from the filter

column, 50 μl of nuclease-free water was used. After cleaning, the library was split into two labelled tubes; 1 was stored at -20°C as a backup, the other was submitted for Illumina MiSeq sequencing. The correct library size and concentration were determined on an Agilent Tape station.

Bacterial 16S libraries were submitted to the MiSeq sequencing team at Wellcome Trust Sanger Institute for sequencing. Sequencing was done using the MiSeq 600 cycle kit, version 3. The default MiSeq read one and read two Illumina sequencing primers, and the index sequencing primers, which form part of the kit, were used. To assess and validate the sequencing, negative control, Staphylococcus Aureus and Mock community A controls were used. The Mock community used was from Bei Resources catalogue number HM-276D. This Mock community control has genomic DNA from 21 different organisms.

2.7 Salmonella and Bifidobacterium Competition assay

To explore the biological antibiotic properties of *Bifidobacterium* against *Salmonella*, we conducted a *Bifidobacterium* and *Salmonella* competition assay by

1) culturing *Salmonella* together with *Bifidobacterium*

2) culturing *Salmonella* in *Bifidobacterium* supernatant.

Before conducting the competition assays, the growth of all *Bifidobacterium* isolates and *Salmonella* in a rich media and a chemically defined media were screened by measuring OD values every 15 minutes for 24 hours. This was followed by determining the actual growth curves of selected isolates by measuring CFU/ ml over time using the method of Miles and Misra (Hedges, Shannon, & Hobbs, 1978; Slack & Wheldon, 1978). The details of these experiments are as follows:

2.7.1 Screening the growth of *Bifidobacterium* isolates and *S*.Typhimurium D23580

S. Typhimurium and all isolates confirmed as bifidobacteria using Sanger sequencing and F6FPK biochemical test were screened for their growth rate. To determine the optimal media to use in the experiments, two types of media were examined. A nutrient-rich medium – Brain Heart Infusion (BHI) and a chemically-defined Iron-restricted medium (CDRIM) were used. To screen for growth rate, *S.* Typhimurium *and Bifidobacterium* from frozen stocks were cultured aerobically and anaerobically at 37°C in 5 ml of LB and MRS broth with tubes loosely covered. *Salmonella* was cultured for 4 hours, while *Bifidobacterium* was cultured for 48 to 72 hours. After incubation, 500 μl of the broth was collected, mixed and centrifuged at 8000 g for 5 minutes to sediment the cells. The supernatant was discarded. Cells were washed twice by resuspending the cells in 1000 μl sterile PBS, centrifuging at 8000 g for 5 minutes and discarding the supernatant. After washing, the bacterial cells were resuspended in 500 μl of PBS.

Two hundred microliters of prepared CDRIM or BHI were added, as appropriate, to multiple wells of a 96 flat well plate. Two microliters of the bacterial suspension were added to the plate in triplicate to make a 1:100 dilution of cells and medium. The plate was sealed with a breathable seal and placed in a plate reader with a shaker, which was in an anaerobic chamber. The experiment was set to run for 24 hours and to read the optical density (OD) at 595 nm every 15 minutes, with 1 minute of shaking and another 1 minute of rest. This experiment was done only once.

2.7.2 Co-culturing of *S.* Typhimurium and *Bifidobacterium*

To monitor the growth of *Salmonella* in the presence of live *Bifidobacterium*, *Salmonella* and representative *Bifidobacterium* isolates were co-cultured. The isolated bifidobacteria were confirmed and identified as described in **Chapter 5**. The isolated *Bifidobacterium* belonged to 4 species; *longum, breve, bifidum* and *pseudocatenulatum*. The majority of the isolates belonged to the *longum* species, which had three clades (clade A, B & C). Two samples; MBB11C for clade A, and MBB30D for Clade C were purposively and randomly chosen. *S.* Typhimurium ST313 D23580, the widely-used reference iNTS strain isolated from a Malawian child, was also used (Kingsley et al., 2013).

S. Typhimurium and *Bifidobacterium* strains were inoculated on LB and MRS agar from stock. These were incubated anaerobically at 37^{o}C for 24 and 48 hours, respectively. One *S*. Typhimurium colony and one *Bifidobacterium* colony was picked from the LB and MRS agar plates and was inoculated in LB broth and MRS broth respectively. Both were cultured using the same conditions. Two hundred microliters of *Bifidobacterium* was quantified using the Miles and Misra technique (Hedges et al., 1978; Slack & Wheldon, 1978).

Quantified *Bifidobacterium* and *S.* Typhimurium were cultured together anaerobically at 37°C by adding 200 µl of *Bifidobacterium* in 20 ml of CDRIM and approximately 1X10^{6} CFU/ml of *Salmonella* to the *Bifidobacterium* CDRIM mixture. The falcon tubes used for culture were loosely covered. One ml of the culture suspension was taken at 0, 3, 6, 9, 12, 15, 18, 21, and 24 hours by properly mixing by inverting. Then 100 µl of *Bifidobacterium* and 10 µl of *Salmonella* was collected and used to

determine CFU/ml using the Miles and Misra technique (Hedges et al., 1978; Slack & Wheldon, 1978). *S.* Typhimurium was incubated aerobically on LB agar for 24 while bifidobacteria were incubated anaerobically on MRS agar for 48 to 72 hours. Booth were incubated at 37°C. After plate incubation, colonies were counted, and Colony Forming Units (CFU) were calculated. Growth curves were plotted with time in hours on the X-axis and CFU on the Y-axis. This experiment was done once but in triplicate with triplicate CFU counts. Media used to culture Bifidobacterium and Salmonella were prepared as indicated in **Table 2.6**.

Table 2.6: Media preparation for co-culture experiments

Media	**Preparation**
MRS broth	55 g of MRS media was dissolved in 1 litre of water and autoclave. After autoclaving, 0.05% of 1M L-cysteine-HCL reagent previously sterilised by filtering was added to the MRS media.
MRS agar	MRS agar was prepared as in MRS broth with the addition of 15 g of agar before autoclaving
LB broth	20 g of LB median was dissolved in 1 litre of water and autoclave.
LB agar	LB agar was prepared as in LB broth with the addition of 15 g of agar before autoclaving
BHI	37 g of BHI was dissolved in 1 liter of water and autoclaved
CDRIM	CDRIM was prepared by adding 5.64 g of 5 x M9 minimal salt, 5 g of Polypeptone, 2 g yeast extract in 800 ml molecular grade water, then autoclaved. After autoclaving, 20 ml of 20% α-D-Glucose, 0.2 g (0.02%) of L-cysteine HCl, 2 ml of 1M $MgSO_4$, and 0.2 ml of 1M $CaCl_2$ was added. Water was added to fill up to 1 litre.

2.7.3 Growth of *S.* Typhimurium in *Bifidobacterium* supernatant

To monitor the growth of *Salmonella* in *Bifidobacterium* supernatant, a single colony of bifidobacteria was cultured anaerobically in 20 ml of CDRIM at 37^{o}C for 24 hours. After incubation, the culture was centrifuged for 20 minutes at 8000 g using the

Eppendorf 5810R centrifuge to collect the conditioned supernatant. The collected supernatant was filtered using a Minisart® CA 0.45-micron non-pyrogenic filter. The pH for each sample supernatant was monitored using pH-indicator paper. The supernatant was serially diluted using CDRIM to a final volume of 5 ml. To control and buffer the effect of pH on *S.* Typhimurium and bifidobacteria growth, the starting pH was measured and adjusted according to 6.5 to 7 using 0.02% $NaHCO_3$. One hundred and ninety-eight microliters of the filtered supernatant were aliquoted into appropriate wells of a 96 well plate and 2 µl of mixed pre-cultured *Salmonella* with a known concentration (approximately 1 X 10^6 CFU/ml) was added to the wells. The plate was sealed and placed in a plate reader in an anaerobic chamber. The plate reader was set to run for 24 hours at 37^{o}C and measure ODs every 15 minutes with 1-minute shaking and another 1-minute rest. Optical densities were measured at 595 nm. A graph with time on the X-axis and the measured OD in log scale on the Y-axis was plotted. *Salmonella* grown in CDRIM alone was used as a control. Samples for this experiment were tested in triplicate, and the experiment was done once.

2.8 Microbiome data analysis

The Illumina MiSeq FASTQ data for both studies were analysed using Quantitative Insights into Microbial Ecology (QIIME) (Bolyen et al., 2019; 2010). The data reported here were analysed using QIIME2, introduced in January 2018 (Bolyen et al., 2019).

2.8.1 Data importing and demultiplexing

The raw sequence data files were in FASTQ file format, which are not compatible with QIIME2. Sequence data in fastq format were converted into QIIME2 artifacts (QIIME2 compatible file format [.qza]). A data directory containing raw demultiplexed sequence

reads in compressed fastq format, metadata (.tsv) and import (.csv) files were used as input files to generate the primary QIIME2 artifact for downstream analysis. Sequence reads were already demultiplexed using the Illumina sequencing pipeline.

2.8.2 Quality checking, filtering and sequence variant calling

Quality checking was done using the Babraham FASTQC application version:0.11.8.

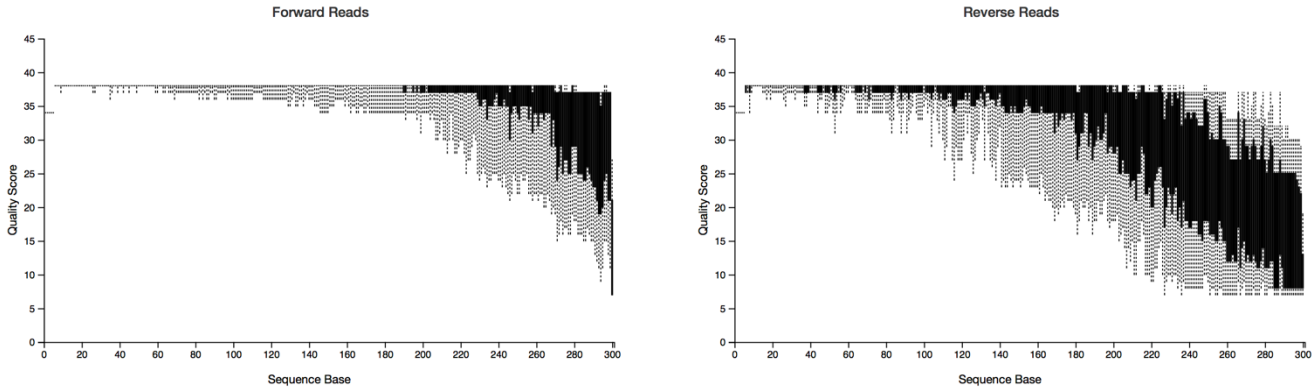

Figure 2.5: Quality score of the forward and reverse reads for SalExpo study

Dada2 pipeline within QIIME2 was used for quality filtering and sequence variant calling (Callahan et al., 2016). In both data sets, the quality of reads (**Figure 2.5)** required no trimming for the forward reads, but for the reverse reads, trimming was done at nucleotide base position 260 and 280 for the SalExpo and BIFBAC studies, respectively.

2.8.3 Taxonomy assignment

Qiime2 was also used to assign taxonomy to the sequence variants. Before assigning taxonomy, the classifier was trained using greengenes 13_8 99% reference sequence and taxonomic classification datasets according to the qiime2 tutorial “Training feature classifiers with q2-feature-classifier” (Qiime2, 2021).

2.8.4 Diversity measurements

Alpha and Beta diversity measurements were computed using Phyloseq package in R. For Alpha diversity, Shannon's diversity index, observed OTUs, evenness and Chao1 phylogenetic diversities were calculated. To estimate Beta diversity across samples, principal coordinates analysis (PCoA) plots for Bray-Curtis distance measurements were computed.

For categorical variables in both cross-sectional and longitudinal data, variations in community composition were assessed by using permutational multivariate analysis of variance (PERMANOVA) test using Pseudo-F statistical method with 4,999 permutations used for significant testing. For continuous data in the longitudinal study, pairwise difference comparisons, linear mixed effect models, volatility analysis and rate of change by assigning a static time point and tracking rate of change over time were analysed.

2.9 Whole-genome sequence data analysis

2.9.1 Assembly and annotation

FASTQ files obtained after WGS were checked for quality using Babraham FASTQC application (version:0.11.8) prior to genome assembly using Unicycler (Wick, Judd, Gorrie, & Holt, 2017). In addition to the Malawi genomes, 88 publicly available *Bifidobacterium* genomes representing all *Bifidobacterium* species and subspecies were retrieved from the NCBI genome database on 4 December 2020. All *B. longum* assemblies were also retrieved from NCBI for comparison with the Malawi *B. longum* genomes. All genomes were annotated using Prokka (Seemann, 2014). tRNA genes were identified using tRNAsan-SE v1.4 (Chan, Lin, Mak, & Lowe, 2019; Lowe &

Chan, 2016) while rRNA genes were identified using RNAmmer v1.2 (Lagesen et al., 2007). Maximum Likelihood trees displayed were generated using RaxML (Stamatakis, 2014). Trees were viewed using iTOL (Letunic & Bork, 2016) and Phandango (Hadfield et al., 2018).

2.9.2 Genomic identification of Malawi reads

Python3 Average Nucleotide Identity (pyANI v0.2.7) and Kraken2 were used to identify the Malawi genomes (Ondov et al., 2016; Pritchard, 2016; Yoon, Ha, Lim, Kwon, & Chun, 2017). pyANI with default BLAST+ settings was used to calculate the Average Nucleotide Identity of the Malawi genomes with the 88 publicly available *Bifidobacterium* genomes. The cut-off point for species delineation was set at 96% identity (Yoon et al., 2017). Kraken2 developed by John Hopkins Centre for Computational Biology, was also used for taxonomic classification of the Malawi genomes by using MiniKraken2_v1_8GB database. MiniKraken2_v1_8GB database was built from RefSeq bacteria, archaea and viral libraries.

2.9.3 Identification of human milk oligosaccharide digesting genes

The Malawi *Bifidobacterium* genomes were assessed for genes responsible for the digestion on HMOs by BLASTing the Malawi genomes against a custom database of 90 HMO genes. HMO digesting genes were identified by assessing the start and the end of the HMO gene sequence within the query genome, percent identity (percent identity of greater than 85% was used as a cutoff for assessment), length of the sequence, gap-open, e-value and bit-score. Differences in HMO digesting gene abundance for the four *Bifidobacterium* species and the three *B. longum* clusters were statistically determined using student t-test.

Table 2.7: Details of laboratory materials and reagent used for SalExpo study

Material/ reagent	Source	Reference
QIAamp Fast DNA Stool Mini Kit	QIAGEN, Netherlands	19593
Zirconia-silicate; 0.1 to 0.15 mm	Cole-Parmer	11079101Z
Ethanol (Molecular grade)	Sigma-Aldrich, Dorset, UK	E7023- 500ml
Taqman array card	University of Virginia	(J. Liu et al., 2013)
PhHv	University of Virginia	(J. Liu et al., 2013)
MS2	University of Virginia	(J. Liu et al., 2013)
lysozyme (10mg/ml)	Sigma-Aldrich, Dorset, UK	10837059001
MicroAmp® Fast Optical 96-Well Reaction Plate, 0.1 ml	Life Technologies	4346907
MicroAmp® Optical Adhesive seals	Life Technologies	4311971
Platinum® Quantitative PCR SuperMix-UDG	Life Technologies	11730-025
Molecular grade water	Sigma-Aldrich, Dorset, UK	W4502
Sartedt Safe Seal -1.5 ml	Sartedt	72.706
Sartedt Safe Seal -2 ml	Sartedt	72.695
Ag-Path-ID Enzyme	Life Technologies	4387391

Table 2.8: Details of laboratory materials and reagents used for BIFBAC study

Material/ reagent	Source	Reference
Qubit 2.0 Fluorometer	Invitrogen	Q32866
MPbio fast DNA spin kit for soil	MP Biomedicals	116560200
Ethanol	Sigma-Aldrich, Dorset, UK	E7023- 500ml
Taqman array card	University of Virginia	(J. Liu et al., 2013)
PhHv	University of Virginia	(J. Liu et al., 2013)
MS2	University of Virginia	(J. Liu et al., 2013)
Phosphate Buffered Saline (PBS)	Sigma-Aldrich, Dorset, UK	P4417-50TAB
Sartedt Safe Seal -1.5 ml	Sartedt	72.706
Sartedt Safe Seal -2 ml	Sartedt	72.695
96 well sterile round bottom plates	Thermo Fisher Scientific	10344311
Sterile spreaders	Sigma-Aldrich, Dorset, UK	HS8171B
MRS broth	BD™ Difco™	288130
Agar	VWR	20767.232
LB broth	Thermo Fisher Scientific	12795027
Sterile 10µl plastic loops	BioLab	731175
RCM broth	Oxoid	M0149
Mupirocin	PanReac AppliChem	A4718
L-cysteine hydrochloride	Sigma-Aldrich	C1276
Glycerol		10219682
pH test strips 0-14 pH, resolution: 1.0 pH unit	Sigma-Aldrich, Dorset, UK	P4786-100EA
Oxoid AnaeroGen 2.5L Sachets	Thermo Fisher Scientific	10269582
50 ml falcon tubes	Thermo Fisher Scientific	14-432-22
15 ml falcon tubes	Thermo Fisher Scientific	14-959-53A
2.0ml cryovial		

Petri dishes	Scientific Laboratory Suppliers	PET2002
Potassium phosphate	Sigma-Aldrich, Dorset, UK	P9791
HCl	Fluka	84435
Sodium Hydroxide	Fluka	35274
Cetrimonium bromide	Sigma-Aldrich, Dorset, UK	H5882
Chloroform:IAA (24:1)	Sigma-Aldrich, Dorset, UK	
Sodium fluoride	Sigma-Aldrich, Dorset, UK	S7920
Sodium iodoacetate	Sigma-Aldrich, Dorset, UK	I2512
D-Fructose 6-phosphate disodium salt hydrate	Sigma-Aldrich, Dorset, UK	F3627
Hydroxylamine hydrochloride	Sigma-Aldrich, Dorset, UK	159417
TCA	Sigma-Aldrich, Dorset, UK	522082
Hydrogen chloride	Fluka	84435
Ferric chloride	Sigma-Aldrich, Dorset, UK	451649
Sucrose	Sigma-Aldrich, Dorset, UK	
Tris EDTA	Sigma-Aldrich, Dorset, UK	93302
Lysozyme	Roche, West Sussex, UK	10837059001
Proteinase K	Roche, West Sussex, UK	03 115 828 001
RNase A	Roche, West Sussex, UK	10 109 142 001
EDTA	Sigma-Aldrich, Dorset, UK	324503
Sodium iodoacetate	Sigma-Aldrich, Dorset, UK	I2512
Sarkosyl NL30	Sigma-Aldrich, Dorset, UK	442753R
Qiagen phase lock tube	Qiagen	129065
Phenol:Chloroform:Isoamyl Alcohol (25;24:1)	Sigma-Aldrich, Dorset, UK	77618
Phenol:Chloroform-Isoamy alcohol (IAA) solution	Sigma-Aldrich, Dorset, UK	P3803
Chloroform:IAA	Sigma-Aldrich, Dorset, UK	77617
E buffer	Sigma-Aldrich, Dorset, UK	
Nuclease free water	Millipore	H20MB0106
dNTPs	Thermo Fisher Scientific	R0181

Q5 High-fidelity polymerase kit	New England Biolabs	M0491S/L
96 well PCR plates, unskirted	Eppendorf	0030 133374
PCR plate seal	Applied Biosystems	4311971
Agencourt AMPure XP beads	Beckman Coulter	A63881
Qubit assay tubes	Thermo Fisher Scientific	Q32856
Quanti-iT PicoGreen dsDNA Assay Kit	Life Technologies	P11496
Phase-lock gel tubes	Qiagen, Manchester, UK	

Chapter 3: Longitudinal Enteric Pathogen Exposure in Healthy 6 to 18 months old Malawian children

3.1 Introduction

Enteric pathogens are an important cause of diarrhoeal and BSI. Diarrhoea remains a leading cause of morbidity and mortality in under 5-year-old children, especially those from the SSA region (GBDa, 2018; GBDb, 2018). According to the 2016 Global Burden of Diseases on diarrhoea, childhood wasting, unsafe water, and unsafe sanitation were the main risk factors for diarrhoea (GBDa, 2018). A study conducted in Blantyre, Malawi, on the aetiology of diarrhoea among hospitalized children demonstrated that rotavirus, adenovirus*, Cryptosporidium,* ETEC and EPEC were associated with diarrhoea (Iturriza-Gómara et al., 2019). Although a study investigating trends in antimicrobial resistance in BSI isolate from 1998 to 2016 in Malawi, reported an overall decrease in pathogens causing BSI, BSI still remains an important cause of morbidity and mortality in SSA. (Iroh Tam et al., 2019). The study demonstrated that *S.* Typhimurium was the leading cause of BSI. The administration of antibiotics in managing BSI presents a huge global challenge of antimicrobial resistance (Iroh Tam et al., 2019; Nwafia, Ohanu, Ebede, & Ozumba, 2019; WHO, 2020). In the Malawi study, 51.1% of the total isolates were resistant to first line antibiotics; amoxicillin or penicillin, chloramphenicol, and co-trimoxazole, and most of the *Salmonella* isolates were MDR.

A high detection rate of enteric pathogens has however also been reported in the absence of clinical symptoms such as diarrhoea. As highlighted in the **Introduction Chapter sections 1.6.4**, asymptomatic enteric pathogen exposure is potentially harmful in that it can cause invasive disease, can

disrupt gut microbiota composition which can impair metabolic and immunological pathways, it has been linked to poor vaccine effectiveness, poor growth and pathogen shedding which is essential for pathogen transmission (M. M. Levine & Robins-Browne, 2012; McMurry et al., 2020; Rogawski et al., 2018) (Haak et al., 2020; Iqbal et al., 2019; Riba et al., 2020). Enteric pathogen infection and diarrhoea are risk factors to reduced cognitive development (Network Investigators, 2018; Walker et al., 2007). Exposure to enteric pathogens whether resulting in symptomatic or asymptomatic outcome, may have immediate or long-term impact on gut microbiota composition and function and subsequently on overall child health.

Children in developing countries such as Malawi are very early in life exposed to risk factors that predispose them to increased enteric infection burden. The reported risk factors of diarrhoea, childhood wasting, unsafe water, and unsafe sanitation, are very common in Malawi. Malawi is one of the low-income countries. In 2020, it was number nine on the list of countries of 20 countries with the lowest GDP (Statista, 2021). Malawi is a country with only 26.23% of the population using basic sanitation services (WHO, 2019). Despite reports of enteric pathogen exposure in asymptomatic populations, previous studies in Malawi have only looked at enteric pathogens in children seeking care for diarrhoeal illnesses. This represents a small proportion of diarrhoea episodes since many more cases are managed in the community. Understanding symptomatic and asymptomatic enteric pathogen exposure is critical in informing treatment for symptomatic exposure and general public health interventions. Work presented in this chapter was aimed at understanding the

type of pathogens that Malawian children are exposed to from 6 to 18 months of age and understanding the exposure patterns of these pathogens.

3.1.1 Hypothesis

Healthy Malawian children are exposed to a wide range of enteric pathogens from 6 - 18 months of age.

3.1.2 Study objectives

1. To determine enteric pathogen exposure patterns in healthy asymptomatic Malawian children from 6 months to 18 months of age
2. To determine the association between enteric pathogen exposure and clinical symptoms during follow-up.
3. To determine the association between enteric pathogen exposure and nutritional status
4. To determine the association between enteric pathogen exposure and water source

3.2 Methods

3.2.1 Participant recruitment, follow-up, sample collection and processing

The procedure used for participant recruitment, follow-up, sample collection and sample processing is as described in section *2.3.1*

3.2.2 Detection of multiple enteric pathogens using Taqman Array Card Assay

TNA was extracted from frozen stool samples using Qiagen Fast DNA stool minkit with a bead beating step added to the manufacturer's instruction. A custom made enteric TAC assay with 29 primer targets was used to detect 19 enteric pathogens present in the extracted TNA. The procedure used for TNA extraction and TAC assay is as described in section *2.4.1.*

3.2.3 Data management and statistical analysis

Cross-sectional analysis was done to determine if there were any associations between the detected pathogens and clinical symptoms and social economic characteristics. Categorical data were analysed using Chi-squared test or Fishers exact test in cases where the number of events were less than 5. Continuous data were analysed using Wilcoxon Mann Whitney test. Univariable and multivariable logistic regression models were used to describe the association between pathogen detection and clinical symptoms, nutritional status and water source. Clinical features tested were: diarrhoea, vomiting, cough, shortness of breath and fever. For nutritional status, we assessed stunting using height for age z-scores, and acute malnutrition using MUAC and weight for age z-scores. All statistical analyses were

performed using R studio Version 1.2.1335. Associations were considered positive at 95% significance.

3.2.4 Ethical approval

Details on ethical approval are as described in section *2.2*

3.3 Results

3.3.1 Participant clinical, demographic and social economic characteristics

Sixty healthy participants attending routine under-five clinics at ZHC were recruited at 6 months of age. Of the 60 participants, 10 were lost to follow-up. Of the remaining 50 participants, 58% were female. At the time of recruitment, 100% of the participants were healthy (as defined in chapter 2) and breastfeeding. Anthropometric measurements: weight ($p = 0.008$) and height ($p = 0.009$) indicated a significant difference between males and females but no significant difference was observed with MUAC ($p = 0.108$). Overall, the weight and height of male children was significantly higher than their female counterparts. Anthropometric measurements, weight and height, increased with time indicating child growth (**Table 3.1**). There were no differences between males and females in any recorded clinical characteristics except for current cough ($p = 0.042$) and breastfeeding ($p = 0.038$) (Table 3.2). Shortness of breath was very uncommon. Only 3 episodes were recorded (2 for history and 1 for current). Due to the small number of recorded episodes, shortness of breath was not included in all further analysis. Although all children were breastfeeding at 6 months of age, more male children were exclusively breastfeeding at this age compared to females. More than 75% of participants' households

Table 3.1: Differences in anthropometric measurements for male and female participants

	6 Months **Median (IQR)**	**12 months** **Median (IQR)**	**18 months** **Median (IQR)**
Weight (kgs)	7.40 (6.65 – 8.05)	8.80 (7.95 -9.25)	10.00 (9.00 – 10.95)
Height (cm)	67.00 (64.00 – 69.00)	73.00 (72-00 – 75.00)	80.00 (77.50 – 82.00)
MUAC (cm)	14.50 (13.88 – 15.50)	15.00 (14.40 – 16.00)	15.00 (14.50 – 16.00)

reported that they were using tap water as their source of drinking water.

Table 3.2: Number of clinical episodes and demographic characteristics at 6, 12 and 18 months of age

	Total number of episodes (n =569)	6 Months (n=56) n positive (%)	12 months (n=47) n positive (%)	18 months (n=47) n positive (%)
History of diarrhoea	99	12 (21.43)	9 (19.15)	7 (14.89)
Current diarrhoea	54	1 (1.79)	9 (19.15)	8 (17.02)
History of vomiting	30	8 (5.38)	4 (8.51)	3 (6.38)
Current vomiting	21	1 (1.79)	4 (8.51)	3 (6.38)
History of cough	127	16 (28.57)	17 (36.17)	14 (29.79)
Current cough	110	7 (12.50)	9 (19.15)	10 (21.28)
History of shortness of breath	2	0 (0)	0 (0)	1 (2.13)
current shortness of breath	1	0 (0)	0 (0)	1 (2.13)
History of antibiotics	139	17 (30.36)	15 (31.91)	8 (17.02)
History of antimalarials	8	2 (3.57)	1 (2.13)	1 (2.13)
Current fever	11	0 (0)	1 (2.13)	0 (0)
	Overall (%)	**6 Months (n=56) n positive (%)**	**12 months (n=47) n positive (%)**	**18 months (n=47) n positive (%)**
Water source: Tap water Borehole River Bottled water	 80.67 11.25 6.15 1,76	 44 (78.57) 6 (10.71) 5 (8.93) 1 (1.79)	 37 (78.72) 5 (10.64) 4 (8.51) 1 (2.13)	 39 (82.98) 4 (8.51) 3 (6.38) 1 (2.13)
Drinking boiled water	33.22	13 (23.21)	13 (27.66)	15 (31.91)
Breastfeeding	96.84	55 (98.21)	46 (97.87)	42 (89.36)

3.3.2 Malawian children aged between 6 to 18 months were exposed to a wide range of enteric pathogens

Four hundred and forty-one samples from 50 participants were available for testing, using the TAC multiplexed enteric pathogen array. Multiple enteric pathogens including viruses, bacteria, protozoa and helminths were tested. A total of 1639 TAC pathogen targets (grouped according to pathogen) were detected as positive. Ninety eight percent of the participants had at least one pathogen detected with a median of 3 pathogens being detected per sample. Bacteria were the most frequently detected pathogen group followed by viruses and parasites (**Figure 3.1**). Helminth detection was uncommon, with only one *Ancylostoma* being detected. Pathogens contributing more than 1.0% of the total pathogen load were used for further analysis (**Figure 3.2**). Overall, a wide range of pathogens were detected at each tested time point although the trend for relative abundance for some pathogens seemed to be stable while others decreased or increased (**Figure 3.3**).

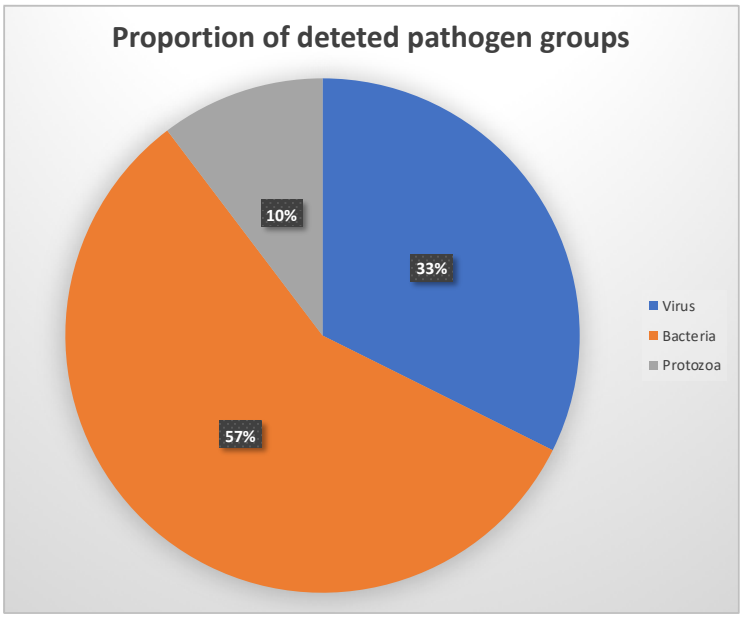

Figure 3.1: Proportions of detected pathogen groups.

Results for pathogens having several targets were grouped into 1. Number of samples = 441. Total number of detected pathogen = 1639

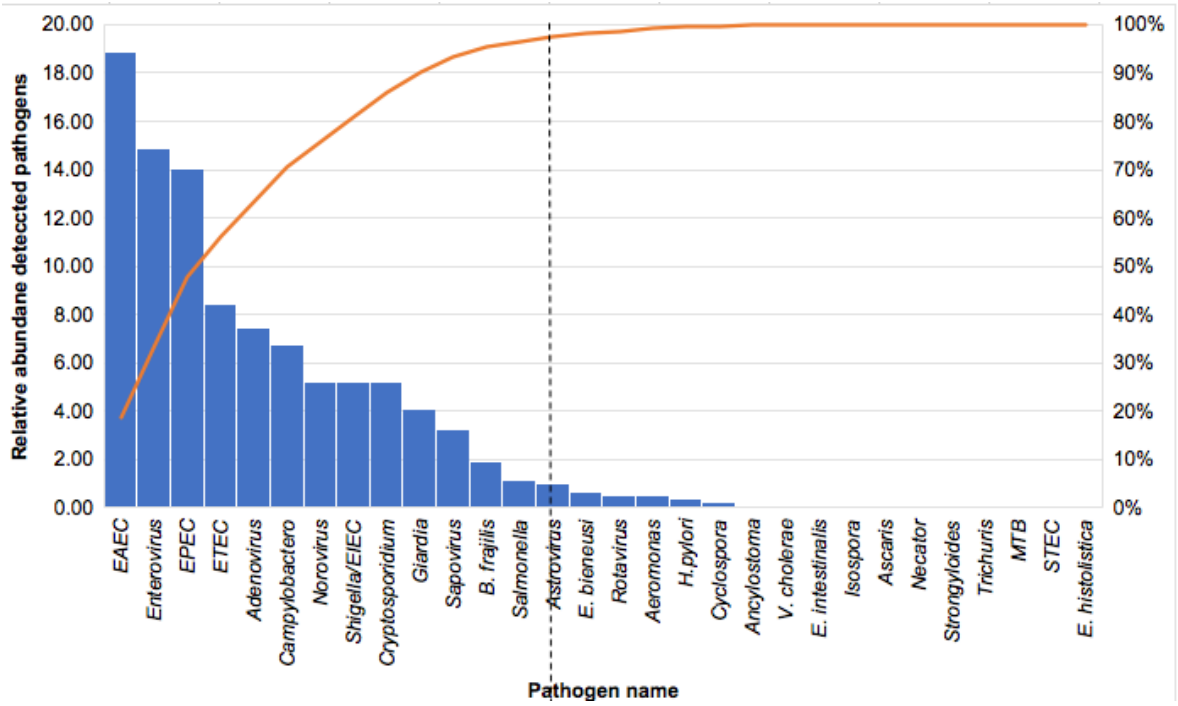

Figure 3.2: Proportion of the detected group enteric pathogens.

The black dotted line indicated the cutoff of pathogens that contributed more or less than 1% of the detected pathogen load. Pathogens to the left of the mark contributed more than 1% and were used for further analysis, while pathogens on

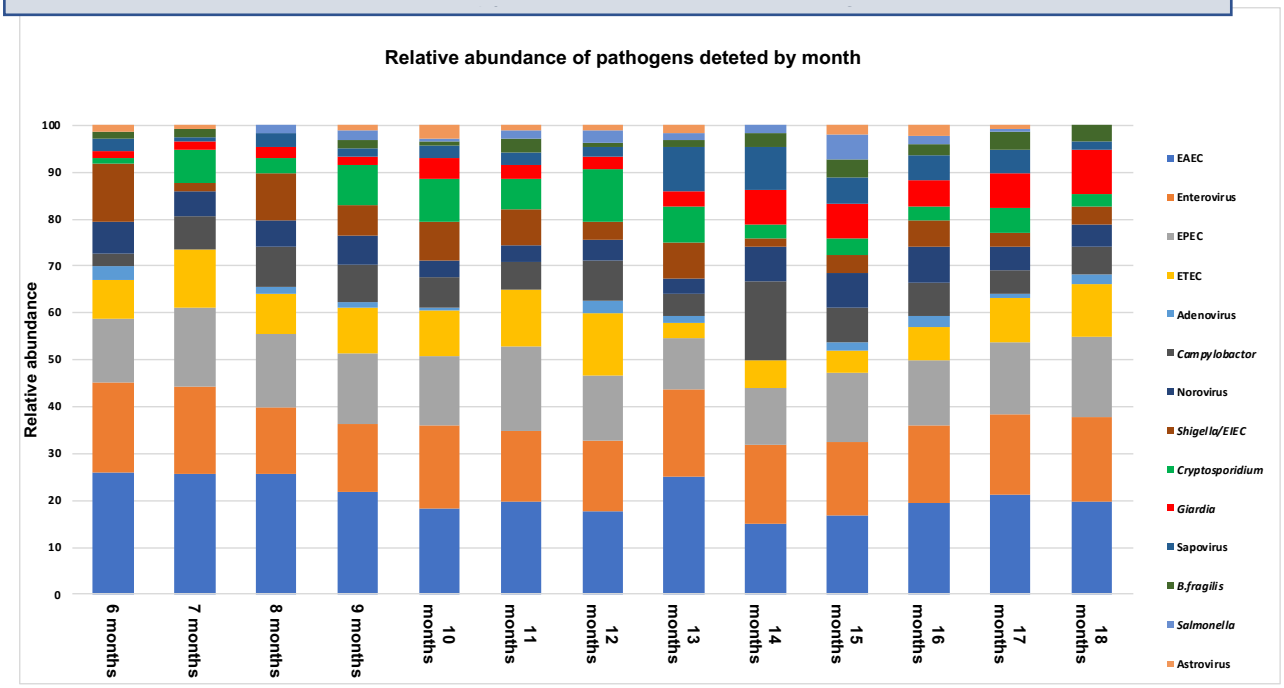

Figure 3.3: Proportion of pathogens detected in stool samples from Malawian children from 6 to 18 months of age.

3.3.3 Bacteria pathogen exposure

Bacteria were detected on 944 occasions among 50 participants (**Figure 3.1**). The most frequently detected bacterial pathogens were *E. coli* followed by *Campylobacter*, *Shigella*/EIEC, *B. fragilis* and *Salmonella* (**Figure 3.4 A**). Of the *E. coli*, EAEC was the most common (48% of the detected *E. coli*), followed by EPEC and ETEC (**Figure 3.4 B).**

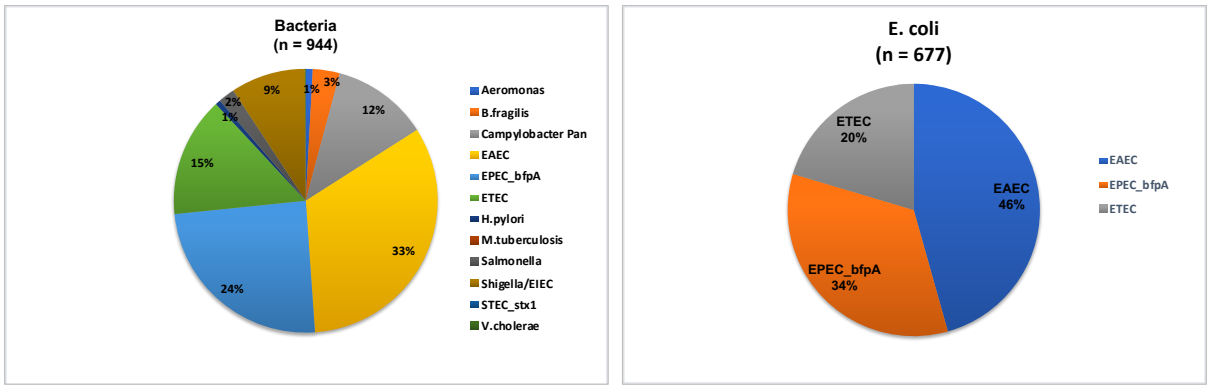

Figure 3.4B (left): Proportions of detected all Bacterial pathogens and
Figure 3.5B (right): Proportions of Detected E.coli

Proportions of all bacterial pathogens- 4A, proportions of detected E.coli pathotypes detected – 4B.

Age of detection: Apart from *Salmonella*, all the other commonly detected bacteria pathogens were detected from 6 months. The earliest *Salmonella* exposure was detected at 8 months (2 participants) and seems to gradually increase with age. The median monthly detection rates of EAEC, EPEC, ETEC, *Campylobacter*, *Shigella*/EIEC, *B. fragilis* and *Salmonella* detection for all time points were 69% (IQR 66% - 73%), 49% (IQR 40% - 62%), 24% (IQR 16% - 38%), 26% (IQR 12% - 47%), 22% (IQR 13% - 28%), 6% (IQR 3% - 9%) and 7% (IQR 2% - 11%) respectively. There was no significant difference in the rate of exposure overtime for all the other bacterial pathogens (**Figure 3.5**). There was variability in exposure patterns within and between individuals for all bacterial pathogens except for EAEC.

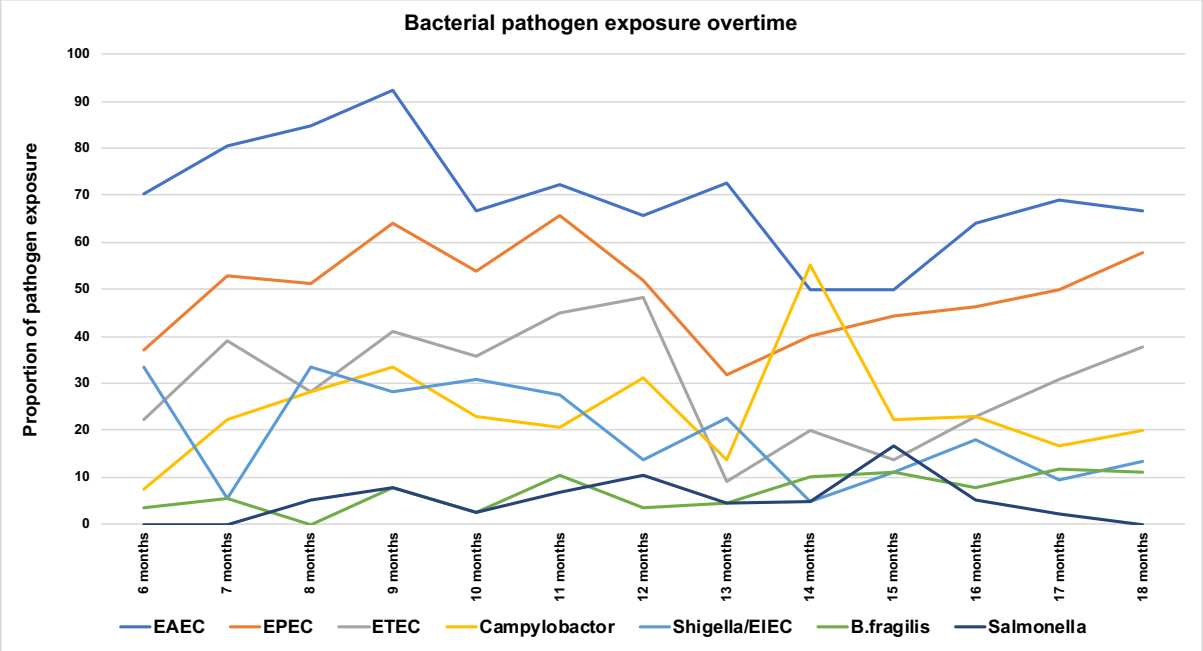

Figure 3.5: Proportion of bacterial pathogen exposure from 6 months to 18 months.

Association between bacteria pathogen exposure and clinical symptoms

Univariate analysis demonstrated a significant association between EAEC detection and history of vomiting (P value = 0.040) and history of cough (P value = 0.025). This association was lost after adjusting for gender and other clinical symptoms in multivariable analysis. No significant association was observed with any other clinical symptoms or social economic characteristics (**Table 3.3**).Detection of EPEC was associated with current diarrhoea and not any of the other tested variables. Children with EPEC exposure were 0.44 times more likely to report having current diarrhoea (P = 0.034). This negative association was present even after adjusting for gender and other clinical symptoms in multivariable analysis (**Table 3.4**).

Both univariate (P value = 0.040) and multivariable (P value = 0.017) analysis demonstrated a significant association between ETEC and history of cough (Table 3.5). Children with ETEC exposure were 0.4 times more likely to report having had a cough between the current and previous follow-up visits to the hospital. There was a

borderline association between ETEC and water source in the univariable analysis (P value = 0.053) and no significant association using multivariable analysis.

Table 3.3: Association between EPEC and clinical symptoms and social economic status

	EAEC (Pos =322)			EPEC (Pos = 227)			ETEC (Pos = 140)		
Feature (n pos)	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**
History of diarrhoea(61)	0.455	0.800	1.10 (0.52, 2.20)	0.516	0.817	0.92 (0.49, 1.76)	0.908	0.574	0.82 (0.40, 1.66)
Current diarrhoea (42)	0.109	0.137	0.57 (0.27, 1.20)	0.063	0.034*	0.44 (0.21, 0.94)	0.635	0.825	1.09 (0.51, 2.30)
History of vomiting (21)	0.040*	0.175	0.46 (0.15, 1.41)	0.291	0.560	0.73 (0.25, 2.12)	0.931	0.460	1.56 (0.48, 5.03)
Current vomiting (18)	0.577	0.816	1.15 (0.37, 3.59)	0.901	0.343	1.70 (0.57, 5.10)	0.553	0.682	1.26 (0.42, 3.77)
History of cough (86)	0.025*	0.118	0.57 (0.29, 1.15)	0.266	0.397	0.76 (0.40, 1.44)	0.040*	0.017*	0.40 (0.19, 0.85)
Current cough (86)	0.413	0.480	0.82 (0.48, 1.41	0.340	0.288	1.31 (0.80, 2.16)	0.947	0.811	0.93 (0.55, 1.59)
History of antibiotics (84)	0.184	0.678	1.17 (0.56, 2.41)	0.542	0.748	1.11 (0.58, 2.16)	0.737	0.292	1.48 (0.71, 3.08)
History of antimalarials (6)	0.751	0.812	0.81 (0.13, 4.80)	0.838	0.882	0.88 (0.17, 4.60)	0.097	0.111	4.17 (0.72, 24.17)
Current fever (10)	0.862	0.918	1.08 (0.26, 4.45)	0.791	0.729	0.79 (0.22, 2.92)	0.608	0.641	1.37 (0.37, 5.11)
Water source: Tap water (316)(ref)									
Borehole (40)	0.239	0.160	0.61 (0.31, 1.22)	0.654	0.432	0.77 (0.40, 1.49)	0.415	0.697	1.15 (0.58, 2.28)
River (22)	0.619	0.727	0.85 (0.33, 2.16)	0.388	0.218		0.053	0.084	
Bottled water (7)	0.293	0.560		0.073	0.255		0.984	0.978	

						0.58 (0.24, 1.38) 0.37 (0.07, 2.06)			3.75 (0.12, 1.14) 3.69 (0, inf)
Boiled drinking water Yes (116)	0.599	0.674	0.90 (0.53,1.50)	0.433	0.532	1.16 (0.73, 1.83)	0.183	0.249	1.33 (0.82, 2.18)

Exposure to *Shigella*/EIEC was not associated with any clinical features or water source. Exposure to *Campylobacter* was positively associated with current diarrhoea and water source using both univariable and multivariable analysis (**Figure 3.6**). Children exposed to *Campylobacter* were 2.8 times (P value = 0.007) more likely to report current diarrhoea than those that were *Campylobacter* negative.

Children exposed to *B. fragilis* were 3.93 times (P value = 0.018) more likely to be using other sources of drinking water and not tap water (**Figure 3.7**). *Salmonella* exposure was not associated with any of the clinical symptoms and social economic status indicators.

Table 3.4: Association between other bacteria; Campylobacter and B. fragilis, and clinical features and social economic status

	Campylobacter (Pos = 106)			***B. fragilis (Pos = 31)***		
Feature	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**
History of diarrhoea (61)	0.295	0.645	1.19 (0.57, 2.49)	0.438	0.552	1.53 (0.16, 2.65)
Current diarrhoea (42)	0.006**	0.007**	2.82 (1.33, 6.02)	0.594	0.876	0.90 (0.23, 3.44)
History of vomiting (21)	0.263	0.232	2.06 (0.63, 6.78)	0.63	0.803	7.38 (0.07, 8.07)
Current vomiting (18)	0.638	0.313	0.53 (0.16, 1.81)	0.143	0.226	2.65 (0.55, 12.88)
History of cough (86)	0.255	0.293	0.65 (0.29, 1.45)	0.79	0.974	1.02 (0.30, 3.42)
Current cough (86)	0.051.	0.094	1.61 (0.92, 2.81)	0.49	0.642	1.23 (0.51, 2.98)
History of antibiotics (84)	0.469	0.709	0.85 (0.38, 1.92)	0.735	0.641	1.34 (0.40, 3.42)
History of antimalarials (6)	0.557	0.586	1.63 (0.28, 9.55)	0.989	0.988	8. 06 (0, inf)
Current fever (10)	0.214	0.501	1.61 (0.40, 6.37)	0.795	0.958	1.06 (0.12, 9.27)
Water source: Tap water (316) Borehole (40) River (22) Bottled water (7)	 0.007** 0.422 0.984	 0.004** 0.370 0.987	 2.75 (1.38, 5.47) 1.54 (0.60, 3.95) 2.67 (0, inf)	 0.863 0.990 0.994	 0.574 0.991 0.994	 2.52 (0.15, 2.89) 1.42 (0, inf) 8.21 (0, inf)
Boiled drinking water Yes (116)	0.073	0.397	1.28 (0.73, 2.25)	0.003**	0.018*	0.63 (0.18, 0.85)

3.3.4 Viral pathogen exposure

Viruses were the second most dominant pathogen group that were detected at 530 occasions (**Figure 3.2**). The most detected viruses were enteroviruses, followed by adenovirus, norovirus and sapovirus (**Figure 3.6**). Rotavirus, the most common pathogen causing diarrhoea disease in children was detected in only 8 samples (2%) (**Figure 3.1**). Of the 530 detected viruses, 3% were astrovirus.

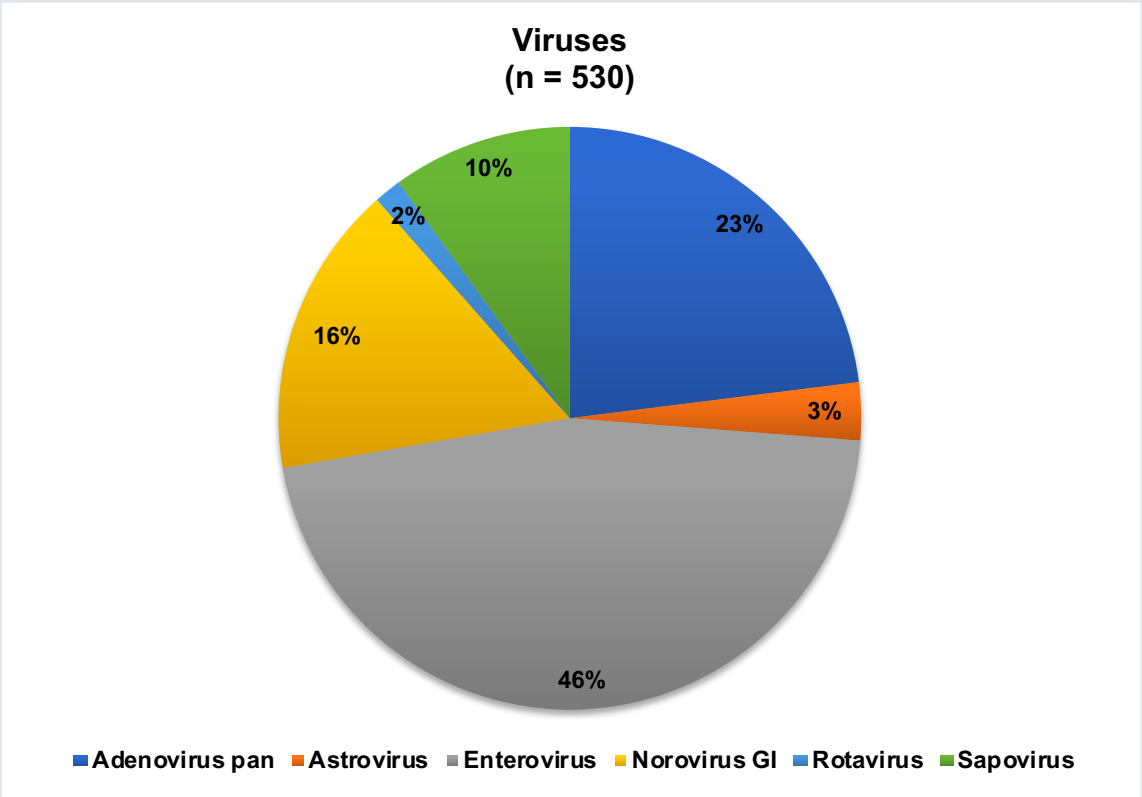

Figure 3.6: Proportions of viral pathogens detected from 6 to 18 months old children.

There was no significant difference in the detection of all viral pathogens by age. Just like *E. coli,* more than 50% of the children were exposed to enterovirus at all time points (**Figure 3.7**). The proportion of sapovirus detection increased from 13 months old although this was not statistically significant (P value = 0.063). There was intra and inter individual variability in the detection of all viral pathogens except enterovirus.

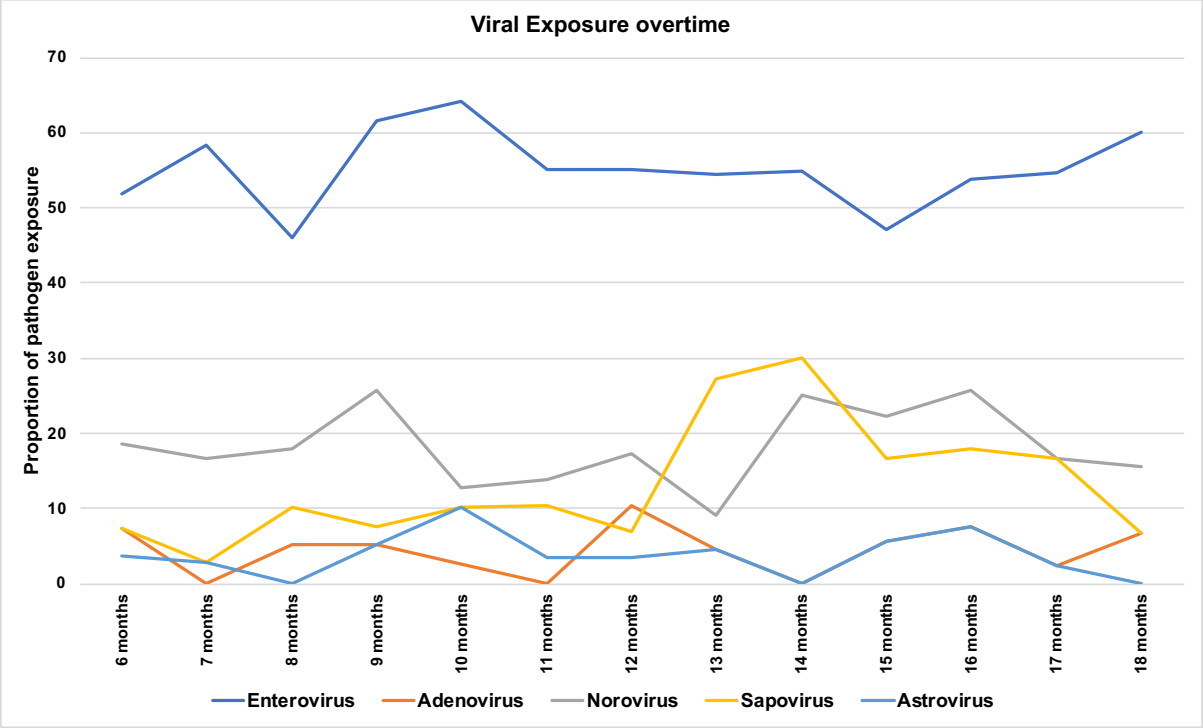

Figure 3.7: Viral pathogen exposure patterns over time

3.3.4.1 Association between viral pathogen exposure, clinical symptoms and social economic characteristics

Statistical analysis to determine the association between enterovirus, astrovirus, rotavirus, norovirus and adenovirus, with clinical features and water source demonstrated no significant differences between positive and negative episodes. Sapovirus was associated with a history of vomiting and drinking water using both univariable and multivariable analysis (**Table 3.5**). Exposure to sapovirus was 1.11 times more likely to have had a history of vomiting (multivariable P value = 0.006). sapovirus positive episodes were 4.16 times more likely to be associated with not drinking boiled water (multivariable P value = 0.006). Figure 3.8 show the association between sapovirus and clinical symptoms and social economic status.

Table 3.5: Association between sapovirus and clinical symptoms and social economic status (n pos = 52)

Feature	Univariable	Multivariable	OR for multivariable (95% CI)
History of diarrhoea (61)	0.82	0.068.	2.34 (0.05, 1.12)
Current diarrhoea (42)	0.064	0.151	1.97 (0.78, 4.98)
History of vomiting (21)	0.005**	0.006**	1.11 (1.98, 61.81)
Current vomiting (18)	0.203	0.699	1.30 (0.34, 5.00)
History of cough (86)	0.109	0.986	1.01 (0.39, 2.59)
Current cough (86)	0.409	0.492	1.28 (0.63, 2.62)
History of antibiotics (84)	0.088.	0.390	1.51 (0.59, 3.90)
History of antimalarials (6)	0.988	0.987	3.93 (0, inf)
Current fever (10)	0.818	0.596	0.56 (0.07, 4.60)
Water source: Tap water (316) Borehole (40) River (22) Bottled water (7)	 0.496 0.161 0.985	 0.700 0.411 0.991	 8.06 (0.27, 2.41) 4.06 (0.05, 3.30) 1.36 (0, inf)
Boiled drinking water Yes (116)	 0.018*	 0.006**	 4.16 (0.22, 0.78)

3.3.5 Protozoa and Helminths exposure

Ten percent of the detected pathogens were protozoa. There were 168 occasions of protozoa detected during the one year study period. *Cryptosporidium* (50%) and

Giardia (40%) were the most detected parasites (**Figure 3.8**). Only 1 helminth was detected.

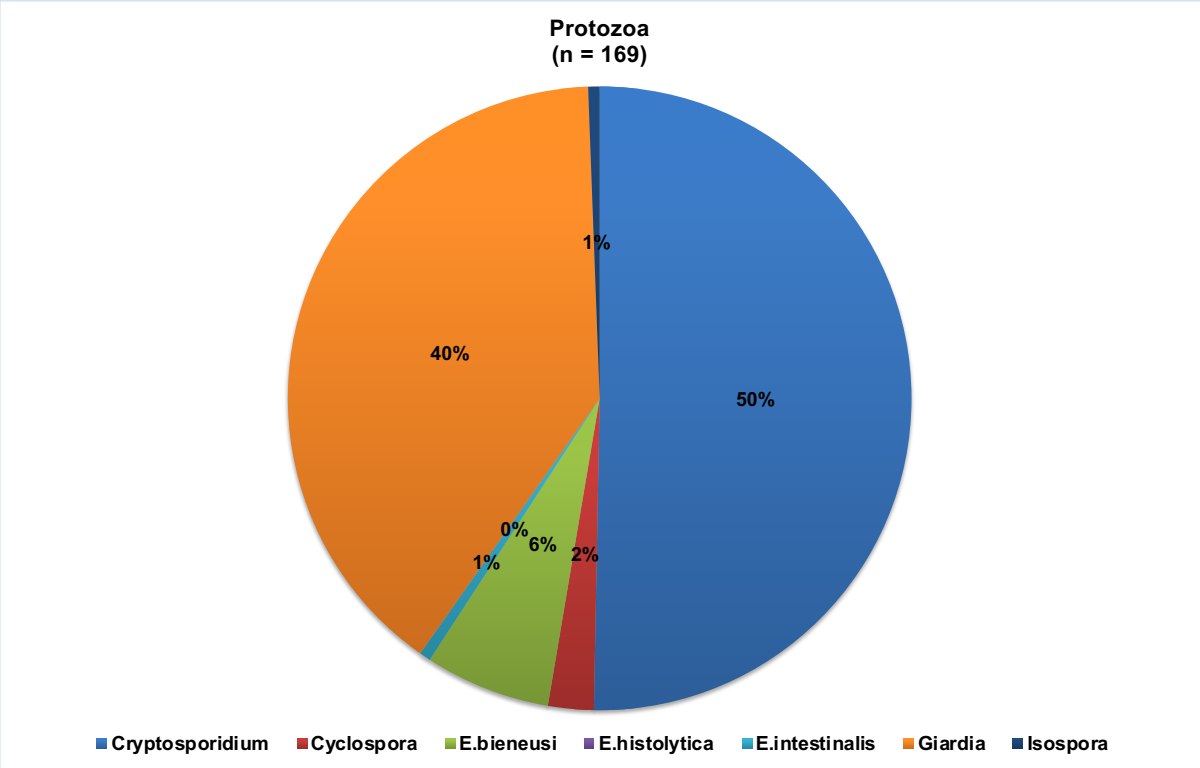

Figure 3.8: Proportions of protozoa detection from children aged 6 to 18 months old

There was no significant difference in *Cryptosporidium* (P value = 0.984), *Giardia* (P value = 0.983) and *E. bieneusi* (1.000) detection over time. Although not significant, the trend showed that *Cryptosporidium* was more likely to be detected in samples collected from participants aged 7 – 13 months old with a peak appearing at 12 months old (**Figure 3.9**). There was a reduction in *Cryptosporidium* exposure from 14 to 18 months. Similarly, although not significant, the trend showed that *Giardia* detection was low between 6 and 13 months and started to rise from 14 months and continued though out the remaining follow-up visits (**Figure 3.9**).

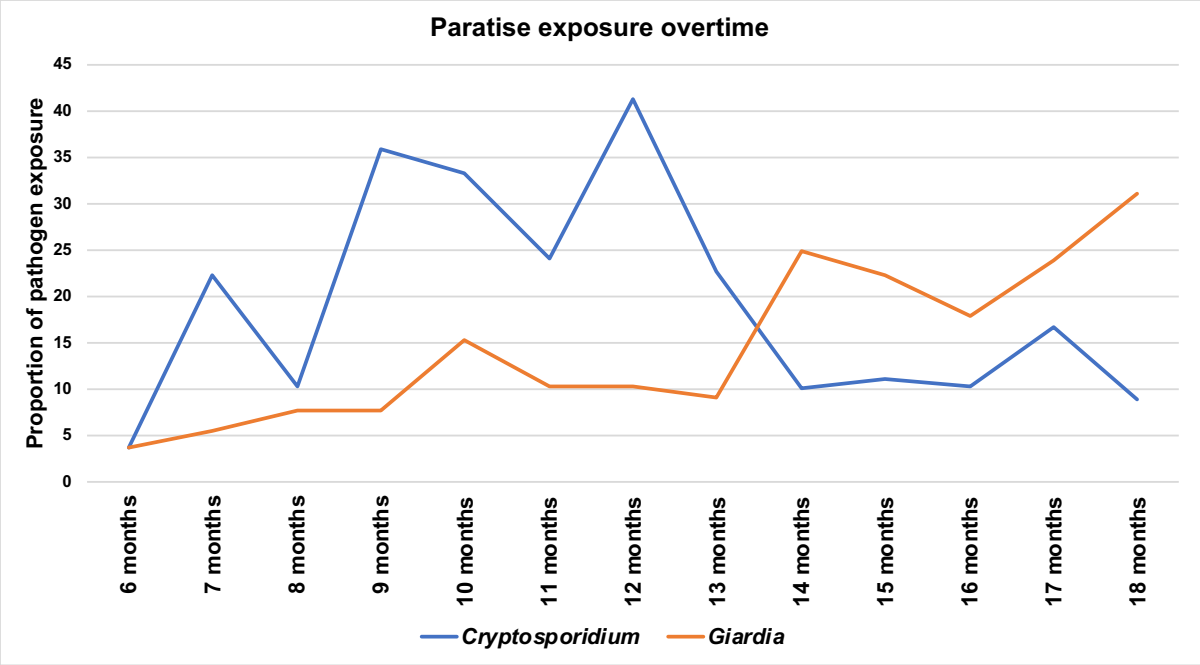

Figure 3.9: Longitudinal protozoa exposure patterns form the 2 commonly detected protozoa.

3.3.5.1 Association between parasite exposure, clinical symptoms and social economic characteristics

Giardia detected was associated with current diarrhoea and not with all other clinical symptoms, using both univariate (P = 0.015) and multivariable analysis (P = 0.0008) (**Table 3.6**). A *Giardia* positive sample was 4.18 times more likely to be detected when a participant had current diarrhoea than a sample that was *Giardia* negative. *Giardia* was also associated with water source (P = 0.046, 0.038). There was however no association using multivariable analysis.

Cryptosporidium was significantly associated with current diarrhoea, current vomiting and water source (**Table 3 .6**). A positive *Cryptosporidium* sample was 2.4 times more likely to be detected when a child had an episode of diarrhoea. Water source remained significant in multivariable analysis. Current vomiting was not significantly associated with *Giardia* when tested using multivariable analysis.

Table 3.6: Association between parasite exposure and clinical symptoms and social economic status

	Giardia (n pos = 66)			***Cryptosporidium (n pos = 83)***		
Feature	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**	**Univariable**	**Multivariable**	**OR for multivariable (95% CI)**
History of diarrhoea (61)	0.451	0.5773	1.29 (0.53, 3.15)	0.089	0.144	1.75 (0.83, 3.73)
Current diarrhoea (42)	0.015*	0.00084***	4.18 (1.81, 9.68	0.003**	0.029*	2.40 (1.09, 5.27)
History of vomiting (21)	0.463	0.3561	1.98 (0.46, 8.48)	0.231	0.569	1.43 (0.42, 4.90)
Current vomiting (18)	0.325	0.0675	1.32 (0.02, 1.16)	0.030*	0.416	1.61 (0.51, 5.04)
History of cough (86)	0.535	0.6817	1.22 (0.46, 3.20)	0.763	0.674	8.33 (0.36, 1.95)
Current cough (86)	0.967	0.6383	1.19 (0.58, 2.43)	0.229	0.38	1.31 (0.71, 2.41)
History of antibiotics (84)	0.114	0.0608	0.35 (0.12, 1.05)	0.852	0.863	9.23 (0.39, 2.20)
History of antimalarials (6)	0.829	0.6705	1.63 (0.17, 15.46)	0.981	0.987	2.26 (0, inf)
Current fever (10)	0.984	0.9831	3.05 (0.0, inf)	0.905	0.696	0.72 (0.14, 3.76)
Water source: Tap water (316)	0.046* 0.108	0.06 0.15	0.14 (0.02,1.08)	0.038* 0.340	0.027 0.662	2.27 (1.09, 4.71)

Borehole (40) River (22) Bottled water (7)	0.038*	0.49	0.22 (0.03, 1.72) 3.48 (0.62, 19.52)	0.602	0.922	7.54 (0.21, 2.68) 1.12 (0.12,10.54)
Boiled drinking water Yes (116)	0.875	0.49	1.27 (0.64, 2.54)	0.857	0.824	9.36 (0.52, 1.67)

3.3.6 : Co-detection of multiple pathogens in symptomatic and asymptomatic children

Studies that have used molecular methods to determine the etiology of diarrhoea have previously shown that most participants whether cases or controls, had co-detection of multiple pathogens. It was hypothesized that this study cohort, which mostly comprised of asymptomatic children, would also have co-detection of multiple pathogens, even though participants were asymptomatic for most of the study period. As hypothesized, there was a very high rate of multiple pathogen co-detection per sample. A total of 87.76% of the samples had more than 1 pathogen detected with a median of 3 (range of 0 to 7) pathogen targets being detected from a sample (**Figure 3.10**). There was no significant correlation between number of co-detections and gender (p = 0.756), and all clinical and demographic characteristics except for current diarrhoea (p =0.006) and current cough (0.035) (**Table 3.7**). Multivariable analysis of these variables did not yield any significant association with number of co-detection.

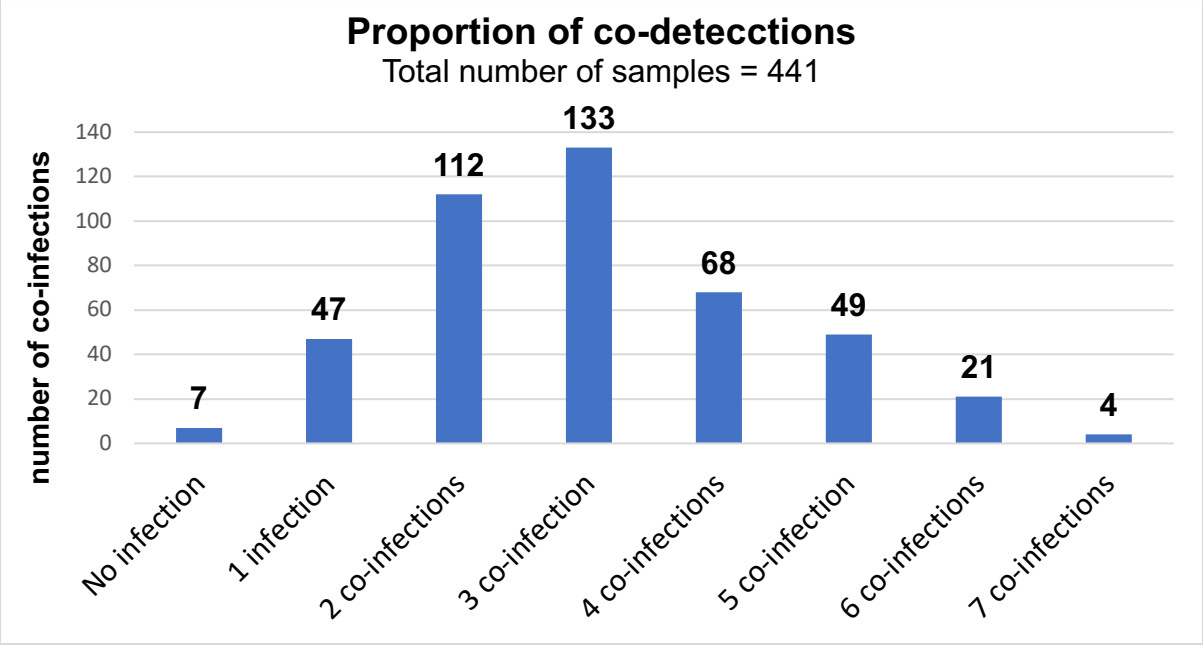

Figure 3.10: Number of samples with co-detected pathogens

Table 3.7: Correlation between number of co-detection and clinical and demographic characteristics

	Number of co-detections		
Feature	**Number of cases**	**Chi square**	**P-value**
History of diarrhoea	36	6.207	0.515
Current diarrhoea	18	19.679	0.006*
History of vomiting	8	8.695	0.275
Current vomiting	88	12.380	0.089
History of cough	47	5.234	0.631
Current cough	36	15.09	0.035
History of antibiotics	57	3.779	0.805
History of antimalarials	5	4.782	0.687
Current fever	4	8.010	0.332
Water source: Tap water Borehole River Bottled water	 237 14 9 0	19.548	0.55
Drinking water Yes No	 84 176	 5.993	 0.540

3.4 Discussion

This study has longitudinally described the enteric pathogen exposure patterns in children from 6 months to 18 months of age. These children, who were asymptomatic at the time of recruitment (6 months old) and for most of the study period, were exposed to multiple enteric pathogens at each time point during the one year followed up period. Previous studies in Malawi have looked at exposure to enteric pathogens in a cross-sectional manner in children presenting to the hospital with diarrhoea or other forms of morbidity such as malnutrition (Iturriza-Gómara et al., 2019; Versloot et al., 2018). Information on the type of enteric pathogens that asymptomatic children from high density urban areas are exposed to and the pattern of exposure was lacking.

Over the 1 year follow-up period, some participants developed clinical symptoms such as diarrhoea, but they remained asymptomatic for most of the study period. Enteric pathogen detection in the absence of clinical symptoms have now been widely reported in LMIC, especially by studies that use molecular pathogen detection methods. The consequences of asymptomatic pathogen exposure, especially in Malawian children, are not very well established and require more research. Studies have however linked asymptomatic pathogen exposure to poor growth outcomes such as stunting and reduced cognitive development (Bartelt et al., 2017; MAL-ED, 2014; Network Investigators, 2018; Platts-Mills et al., 2017; Walker et al., 2007). The high burden of asymptomatic pathogen exposure is also being implicated in reduced oral vaccine effectiveness (Hajela, Chattopadhyay, Nair, & Ganguly, 2020; Hosomi & Kunisawa, 2020). These poor outcomes have been linked to pathogen mediated gut microbiota dysbiosis (de Jong, Olin, & Pulendran, 2020). Asymptomatic pathogen

detection may arise due to shedding before or after a symptomatic exposure, chronic carriage or continuous transmission from contaminated environments (McMurry et al., 2020). Pathogen shedding from asymptomatic populations may significantly contribute to pathogen transmission in the community/ environment (M. M. Levine & Robins-Browne, 2012). Understanding the true burden of asymptomatic exposure will inform the design of public health interventions.

A wide range of pathogens were detected in children regardless of whether they were symptomatic or asymptomatic. The commonest detected bacterial pathogens were EAEC, EPEC, ETEC, *Campylobacter, Shigella, B. fragilis* and *Salmonella.* Studies using similar pathogen detection methods have shown EAEC and *Shigella/*EIEC as the principal bacterial pathogen in both symptomatic and asymptomatic populations. Although EAEC and *Shigella* were not significantly associated with clinical symptoms such as diarrhoea and vomiting, their role in other morbidities such as malnutrition, should be investigated. A Bangladesh study aimed at determining the association between enteric pathogen exposure and malnutrition in 6 to 23 months old malnourished children demonstrated that EAEC, *Campylobacter*, heat-labile ETEC, *Shigella*/EIEC, norovirus genogroup I and *Giardia* were associated with malnutrition (Platts-Mills et al., 2017). In their study, they found that malnutrition was mostly associated with bacterial and not viral pathogens. The link between specific bacterial enteric pathogen exposure (symptomatic or asymptomatic) and malnutrition has not been established for the Malawi population which has a huge burden of malnutrition cases.

A study investigating the etiology of diarrhoea in diarrhoea in-patient children at QECH in Blantyre demonstrated that adenovirus, *Cryptosporidium*, ETEC, EPEC and *Shigella*/ EIEC were associated with diarrhoea (Iturriza-Gómara et al., 2019). In the current study, the presence of EPEC, *Campylobactor, Giardia* and *Cryptosporidium* was associated with diarrhoea. Another study looking at pathogen-specific burden of diarrhoea showed that norovirus, rotavirus*, Campylobacter,* astrovirus *and Cryptosporidium* were mainly associated with diarrhoea in the first year of life while *Campylobactor,* norovirus, rotavirus, astrovirus and *Shigella* were associated with diarrhoea in the 2nd year of life (Platts-Mills et al., 2018). This shows that there is a wide range of enteric pathogens that have the potential to cause diarrhoeal diseases. Deliberate efforts should therefore be made to consider both symptomatic and asymptomatic enteric pathogen exposure pattern in understanding the short and long term consequences of enteric pathogen exposure and when designing public health interventions that aim at dealing with diarrhoeal infections.

Although *Salmonella* exposure was not common and was not associated with any clinical symptoms, about 50% of children were exposed to *Salmonella* for at least one time point within the one year study period. Malaria and malnutrition are the main risk factors for *Salmonella* infection in under 5 year old children (Feasey et al., 2015; M. A. Gordon, 2011). In this study population, only 8 episodes of history of taking antimalarials were reported and acute malnutrition based on MUAC was rare. Low prevalence of Salmonella risk factors, such as malaria and malnutrition, may have contributed to the low detection of *Salmonella* at different time points. *Salmonella* exposure in this study may not be linked to clinical symptoms, but previous studies have shown a possible beneficial impact of subclinical or

asymptomatic enteric *Salmonella* exposure, in that it may lead to T cell and antibody mediated immune response (Tonney S. Nyirenda et al., 2014). Asymptomatic *Salmonella* exposure, just as with all other pathogens may, however, also be a risk for symptomatic or invasive infection and pathogen transmission.

Despite having a very high prevalence of enterovirus exposure, the descriptive analysis conducted in this study did not show any statistically significant associations with any clinical symptom. Quantitative analysis of the data would take into account pathogen load and may produce different results. Just as with EAEC, the impact that enterovirus may have on the incidence of other enteric pathogens, oral vaccine effectiveness and poor growth outcomes should be studies in Malawi and the whole SSA. Enterovirus detection has been associated with reduced rotavirus IgA, failure to seroconvert and rotavirus diarrhoea in children vaccinated against rotavirus (Taniuchi et al., 2016).

The overall prevalence of rotavirus was very low in this cohort. This could be attributed to high immunization coverage (Bar-Zeev et al., 2016). This is in agreement with a case control study recruiting Malawian children under 5 year of age from 2012 to 2015. They reported a low prevalence of 1.5% in controls but the prevalence in diarrhoea cases still remained high - 34.6% (Iturriza-Gómara et al., 2019).

Rotavirus vaccine (Rotarix) was introduced in Malawi in 2012 with 92% immunization coverage. A reduced vaccine effectiveness of 31.7% in the second year of life has however been reported. This is in comparison to 70.6% vaccine effectiveness reported in the first year of life. The vaccine has resulted in a reduction

in rotavirus prevalence and hospitalisations in the under 12 months old infants (Bar-Zeev et al., 2016). Low prevalence rates in the control population has also been reported in other case control studies. The MAL-ED study conducted an intensive birth cohort community surveillance study in Africa, Asia and Africa. They also reported low rotavirus prevalence rate from non diarrhoeal compared to diarrhoeal stools (Platts-Mills et al., 2015).

The prevalence of enteric pathogens or diarrhoeal infections is high in LMIC where levels of poor sanitation and use of untreated water is common (GBDa, 2018). Malawi still remains one of the poorest countries in the world (Bank, 2020). Although the absolute numbers of people living in urban settings is low, the rate of urbanization in Malawi is increasing very quickly (Office, 2019). Urbanization has an impact on the epidemiology of infectious diseases and global health. Some urbanization specific risk factors for infectious diseases are poor sanitation and untreated drinking water (Boyce, Katz, & Standley, 2019). This study identified some enteric pathogens (sapovirus, *Campylobacter*, *B. fragilis, Giardia* and *Cryptosporidium*) that were associated with using boreholes as the main source of water and not drinking of boiled water.

3.5 Conclusion

This study has shown that Malawian children are exposed to a wide array of enteric pathogens, very early in life. Bacterial and viral pathogens were detected throughout the study period with intra and inter individual variations observed. In this population, EAEC and enterovirus exposure was very common. Helminth exposure was uncommon in this young population. An sequential pattern of timing of detection was observed between *Cryptosporidium* and *Giardia*, with *Cryptosporidium* being highly detected between 6 and 12 months, while an increase in *Giardia* detection was observed from 12 to 18 months. Multiple co-detection of pathogens was common. Most of this exposure was however not associated with clinical symptoms and drinking water but water source.

The significance of this high burden of pathogen exposure in a population that was mostly asymptomatic should be further studied. *E. coli* and enterovirus exposure may not be clinically significant but needs to be further studied for the immunological implication it may have in vaccine response in this low-income population and also its role on other enteric pathogen infections. The relationship between overall asymptomatic pathogen exposure and malnutrition (stunting and waiting), reduced vaccine effectiveness should be established. These outcomes should be taken into account when designing public health interventions such as sanitation measures.

Chapter 4: Gastrointestinal tract microbiota changes Associated with Enteric Pathogen exposure in Malawian children

4.1 Introduction

Being open to the outside environment, the gut is continuously exposed to different pathogens, making it a common site for infection. Like most SSA, children in Malawi are exposed to a wide range of enteric pathogens, which are a common cause of morbidity and mortality (Colston et al., 2020; J. H. Kim, Cheong, & Jeon, 2018; MAL-ED, 2014). It has been demonstrated in **Chapter 3** that, on average, a Malawian child is exposed to 3 pathogens at one point in time. Most of these exposure events are asymptomatic. This asymptomatic enteric pathogen exposure reported in Malawian children has also been reported in children from other SSA countries like South Africa (MAL-ED, 2014), Tanzania (MAL-ED, 2014), Kenya (Kotloff, 2017), Mozambique (Kotloff, 2017) and Zambia. Besides diarrhoea which is a common direct and immediate effect of enteric infection, enteric pathogen exposure has also been linked to health abnormalities that can be indirect and long-lasting. These abnormalities include malnutrition/ stunting.

Besides enteric pathogen exposure, gut microbiota dysbiosis during early life has also been linked to long-term health challenges (F. Turroni, Milani, Ventura, & van Sinderen, 2021). Factors such as enteric pathogen exposure and antibiotic usage can cause gut microbiota dysbiosis (Socha-Banasiak, Pawłowska, Czkwianianc, & Pierzynowska, 2021). Gut microbiota dysbiosis due to enteric pathogen exposure or infection could significantly affect a child's development and disease prognosis. Changes in gut microbiota composition may create a conducive environment for primary or secondary infection, and it also promotes pathogen replication.

Here we describe changes in gut microbiota composition with exposure to different enteric pathogens. The relationship between enteric pathogens and gut microbiota dysbiosis can be two ways; enteric pathogens can lead to gut microbiota dysbiosis, and gut microbiota dysbiosis can promote enteric pathogen infection. This study only considered a situation where pathogen exposure of the gut may lead to dysbiosis.

4.1.1 Hypothesis

Enteric pathogen exposure in Malawian children is associated with changes in the gut microbiota profiles.

4.1.2 Objectives

1. Describe the longitudinal changes in the general gut microbiota composition of Malawian children 6 to 18 months old
2. Describe differences in gut microbiota composition with environmental conditions that may predispose children to enteric infection
3. Describe differences in gut microbiota composition with nutritional status
4. Describe differences in gut microbiota composition with bacterial, viral and parasitic enteric pathogens

4.2 Methods

Detailed methods used in sample collection and processing and data processing and analysis for the data being presented here can be found in Chapter 2.

4.2.1 Study type and study site

Samples for the data reported in this chapter were collected from children recruited in a longitudinal cohort study conducted at Zingwangwa Health Centre. Children were recruited at six months and followed up for one year. Details are in **Chapter 2, section 2.1.**

4.2.2 Ethical approval

Ethical approval for this study was obtained from the College of Medicine Research Ethics Committee (COMREC – P.01/13/1327). Full details are in **Chapter 2, section 2.2**.

4.2.3 Enteric pathogen detection

Sixty healthy children were recruited at six months of age and were followed monthly over one year. The table below indicates the samples used in this chapter, showing participant samples and age at the time of sample collection for samples. Data used here are only for samples with both TAC and sequenced data. Details can be found in **Chapter 2, section 2.5.2.**

4.2.4 DNA extraction, library preparation and sequencing for microbiota profiling

Qiagen Fast stool Minikit with a bead-beating step was used for stool DNA extraction. DNA extraction was done within seven days of sample collection. Libraries were prepared targeting the V1V2 region of the 16S rRNA gene. Libraries were sequenced using the Illumina miseq sequencing platform. Both library preparation and sequencing were done at Wellcome Trust Sanger Institute. Details for this can be found in **Chapter 2, section 2.5.4.**

4.2.5 Microbiota sequenced data analysis

Illumina sequenced data were assessed for quality using FastQC. All bioinformatic analysis was conducted using Qiime2. Following the quality check, forward and reverse reads were truncated at 284 and 260. DADA2 within Qiime2 was used to truncate the forward and reverse reads at positions 284 and 260, respectively, and sequence variant calling. The taxonomy of SVs was assigned using the green genes database. Details can be found in **Chapter 2, section 2.8.**

4.2.6 Data management and statistical analysis

Qiime2 artefacts for taxonomy and SVs were exported into csv files. The csv files were then exported into R statistical package for statistical analysis. A phyloseq object was created by combining a biom file as a matrix, a taxonomy file as a matrix and a metadata file as a data frame. Alpha diversity was computed using Observed OTU, Chao1, Fisher and Shannon diversity metrics. Shapiro-WILK normality test was used to test for normality of the Alpha diversity metrics. For Alpha and beta diversity, non-parametric statistical methods were used because the alpha and beta

diversity values were not normally distributed. For Alpha diversity, pairwise group comparisons were made using Wilcoxon rank-sum test, while multiple group comparisons were made using the Kruskal-Wallis test. Bray-Curtis distances were computed for beta diversity analysis, and a non-metric multidimensional scaling analysis was done. Specifically, to test for significance in associations between participant characteristics, environmental condition and pathogen exposure with beta diversity, PERMANOVA analysis with 999 permutations by using the Adonis function in a vegan R package (Dixon, 2003) was done. Tests that detected significant associations were subjected to age-adjusted regression analysis to determine the impact of age as a fixed effect in the regression analysis. Differences between groups were calculated using the Wilcoxon rank-sum or Kruskal-Wallis test.

Gut microbiota compositional differences for particular pathogens were assessed by comparing the gut microbiota in participants with positive detection against negative detection.

4.3 Results

4.3.1 General description of 16S rRNA sequence reads

Four hundred and five samples were sequenced on four lanes. The total number of sequence reads were 44,386,952 with a median of 106,768 (IQR of 80,846 to 132,939) reads. All the 405 raw sequence reads went through a post-sequencing quality filtering step. Forward reads were of higher quality than the reverse reads. The total number of sequence reads obtained after quality filtering and removing chimeras were 33,603,559, with a median of 78,561 (IQR of 60,735 to 100,472) reads (**Figure 4.1**). Sequence variant calling identified 10,237 amplicon sequence variants. Of the 10,237, 2919 were singletons, 2664 were identified as chloroplast or mitochondria. Before further analysis, singletons and chloroplast/ mitochondria taxa were pruned from the dataset. Three samples with a sequencing depth of greater than 200,000 were also pruned (**Figure 4.2**). The data had 98.49% sparsity.

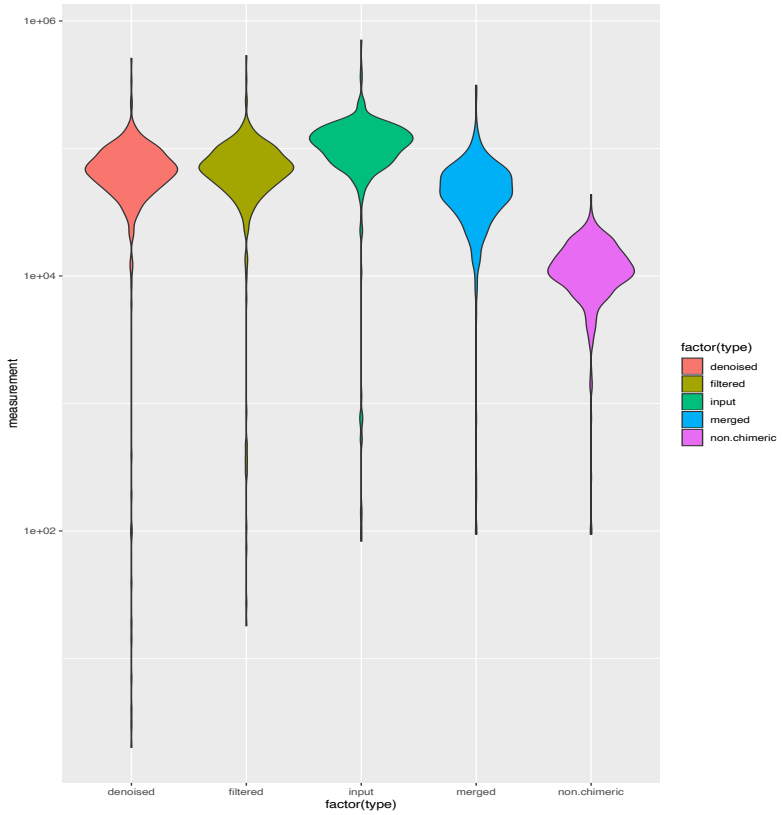

Figure 4.1: Number of sequence reads.

Raw sequence reads (input) went through different quality filtering steps to obtain reads used for analysis (non-chimeric). The specific filtering steps were filtering and denoising, merging the reads and

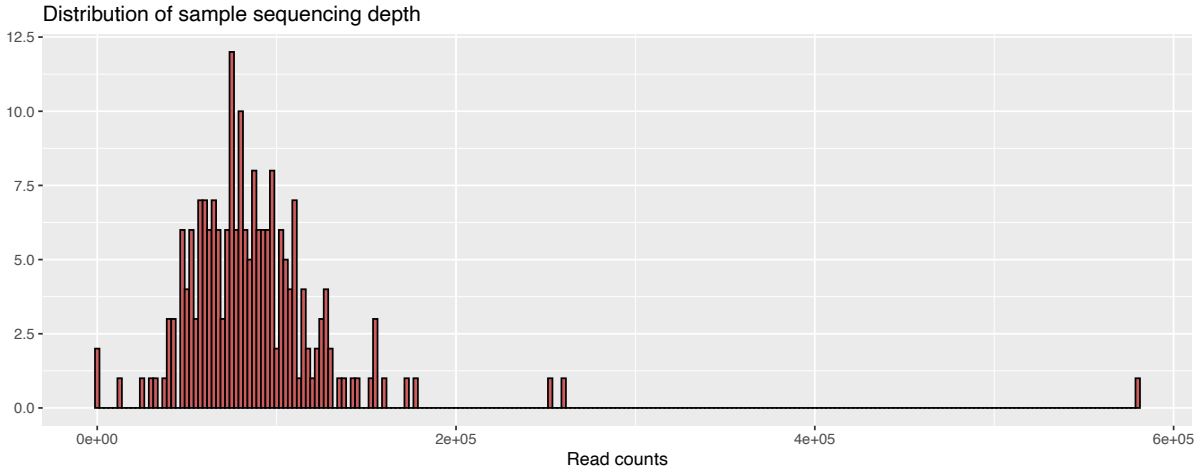

Figure 4.2: Distribution of sequencing depth.

4.3.2 Quality control assessment

A comparison of the control reads included in the four lanes was made to determine if there were any differences in sequencing for the different lanes. The study used a mock community, positive *Staphylococcus Aureas* and negative controls. For each particular lane, the negative control, S. Aureus and Mock community controls contributed less than 0 %, 1.23% and 1% of the total reads, respectively (**Table 4.1**). There were no significant differences in the number of control sequence reads by lanes.

Table 4.1: The number of sequence reads for controls used in the study

Number of sequence reads for controls (%)				
	Lane-22729	**Lane-22751**	**Lane-22818**	**Lane-22820**
Neg	819 (0.01)	681 (0.01)	517 (0)	517 (0)
S. Aureus	220,918 (1.56)	169,009 (1.26)	140,069 (0.98)	NA
MockCommA	NA	148,601 (1.1)	134,231 (0.94)	137,130 (0.96)

4.3.3 General description of the gut microbiota of Malawian children from 6 to 18 months of age

The 10,237 ASV identified comprised of 17 phyla with Actinobacteria, Firmicutes, Proteobacteria and Bacteroidetes, making more than 95% of the total reads. Two hundred and twenty-nine genera were identified. Of these, 12 genera contributed greater than 1% in abundance and were found in more than 10% of the samples. These represented 86.13% of the total sequence reads. In descending order, the 12 most abundant genera were *Bifidobacterium, Streptococcus, Lactobacillus, Prevotella, Olsenella, Escherichia_Shigella, Bacteroides, Veillonella,*

Faecalibacterium, Collinsella, Clostridium and *Shigella (***Figure 4.3***)*. Twenty-one genera contributed between 0.2% to 1% of the gut microbiota abundance, while 196 genera contributed less than 0.1% (**Figure 4.4**). These 196 different genera contribute 5% of the total sequence reads.

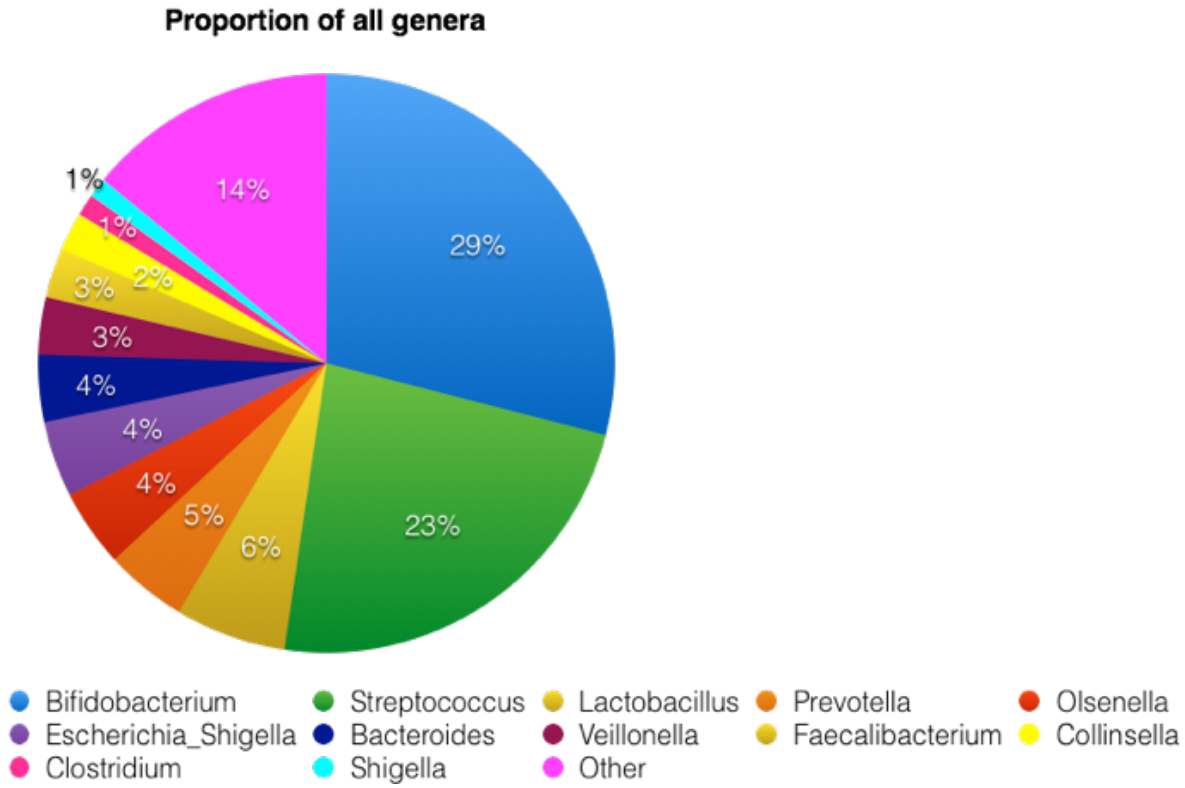

Figure 4.3: Relative abundance of the gut microbiota at the genus level.

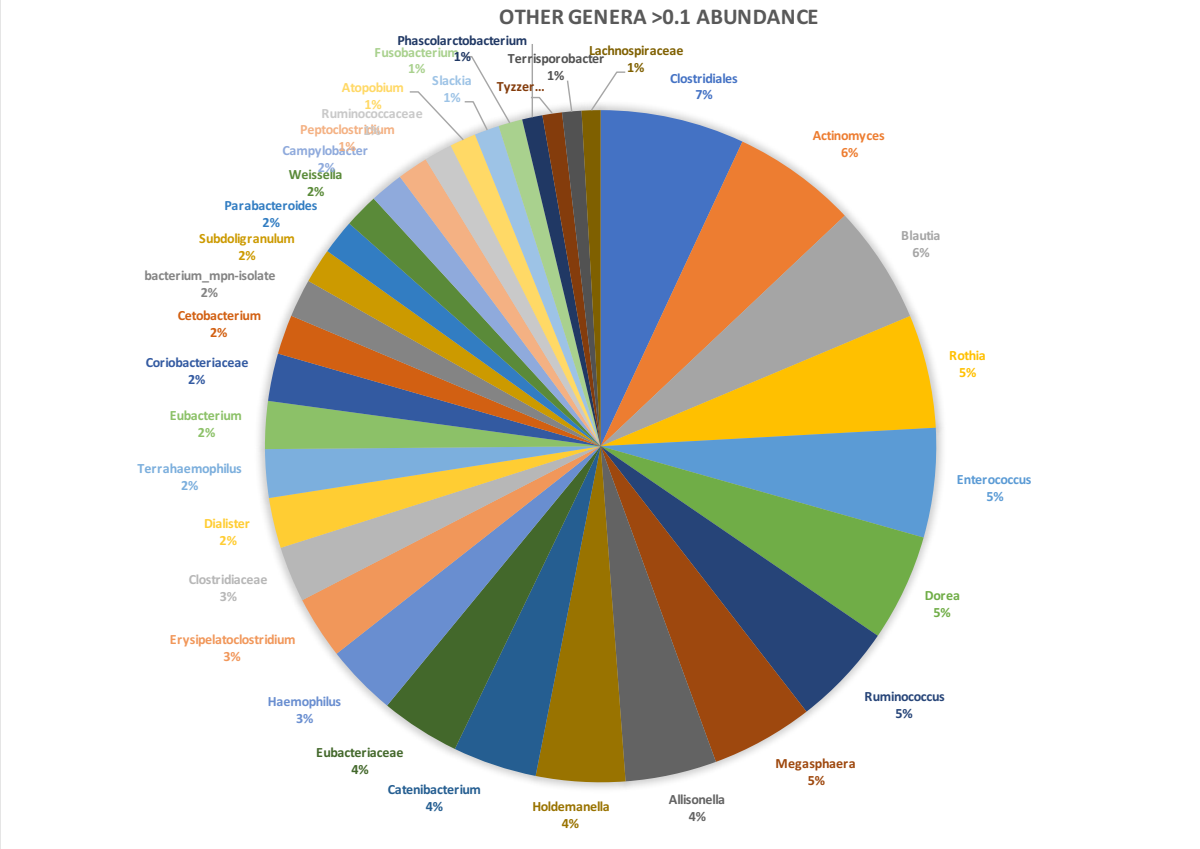

Figure 4.4: Relative abundance of the gut microbiota at the genus level. Gut microbiota composition with age and gender

To determine the alpha (within-sample diversity) diversity within samples, Observed OTU, Chao1, Shannon and Fisher diversity metrics were computed. The distribution of the alpha diversity metrics indicated a normal distribution for Shannon and a right-skewed diversity for Observed OTU, Chao1 and Fisher metrics (**Figure 4.5**). The overall median (IQR) for Observed OTU, Chao1, Fisher and Shannon index was 130.5 (100 -171.8), 130.5 (100 -171.8), 15.464 (11.358 - 21.078 and 2.957 (2.472 - 3.369), respectively.

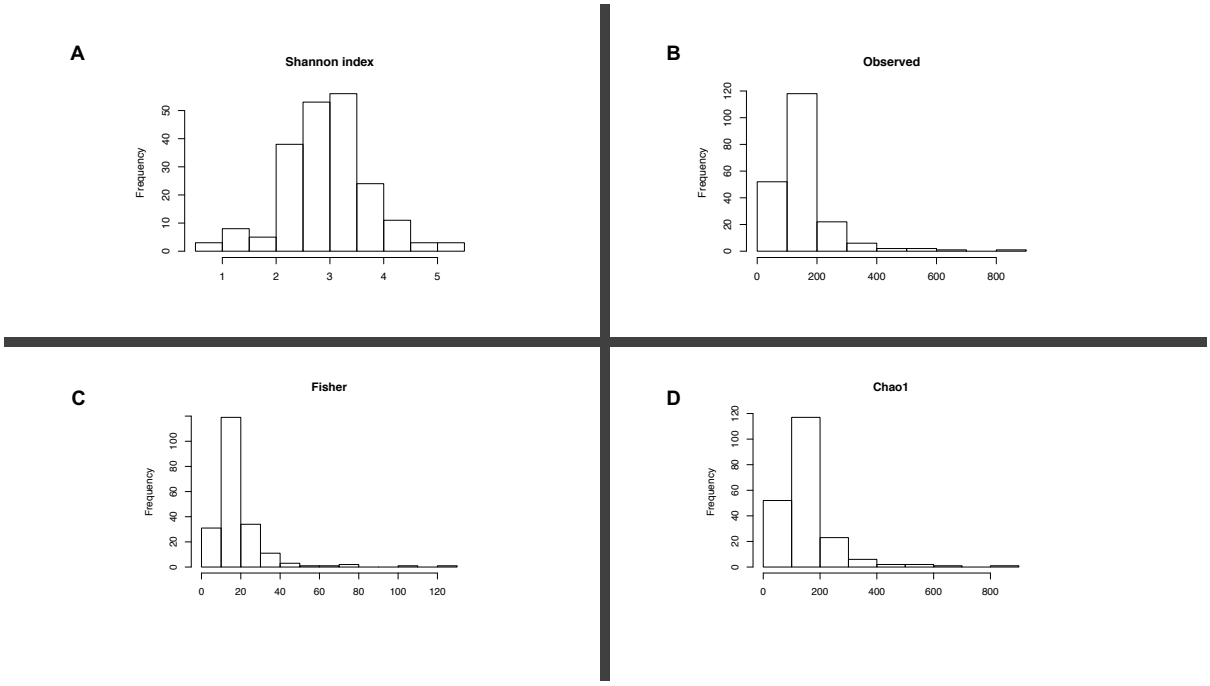

Figure 4.5: Distribution of alpha diversity metrics.

The distribution of alpha diversity metrics tested in the study plotted to determine if the data is normally distributed or not. A shows normally distributed Shannon index, B, C, and D indicates a right-skewed distribution for Observed OTU, Fisher and Chao1 index.

Significant differences were observed in alpha diversity with age. All the measured Alpha diversity metrics; observed OTU, Chao1, Fisher, and Shannon significantly increased with each subsequent month. **Table 4.2** shows the median (SD) values for observed OTU, Chao1, Fisher and Shannon diversity metrics. Significant differences were also observed in the four calculated diversity metrics when participants were grouped into two age groups; 6 – 12 months and 13 - 18 months (**Figure 4.6**). Participants in the 13 - 18 months age group had significantly higher alpha diversity values than the 6 to 12 months age group. For the 6 to 12 months age group, the median (SD) values for Observed OTU, Chao1, Fisher and Shannon were 126.00 (89.77), 126.25 (89.85), 15.12 (13.02) and 2.93 (0.75), while for the 13 to 18 months age group, the median (SD) values were 249 (128.57), 249.17 (128.71), 28.59 (19.22) and 3.79 (0.85) respectively.

Table 4.2: Distribution of Shannon diversities showing the median and standard deviation of the diversity values

Age	**Observe OTU (SD)**	**Chao1 (SD)**	**Fisher (SD)**	**Shannon (SD)**
Six months	99.00 (53.86)	99.00 (53.88)	11.07 (7.09)	2.83 (0.82)
7 months	130.50 (55.43)	130.75 (55.42)	15.26 (7.53)	2.98 ()0.75
Eight months	122.00 (57.25)	122.00 (57.26)	14.98 (7.92)	2.92 (0.68)
9 months	122.00 (133.99)	122.00 (134.13)	14.41 (19.53)	2.94 (0.86)
10 months	134.22 (122.01)	134.22 (122.09)	15.75 (12,76)	3.03 (0.72)
11 months	147.76 (98.22)	147.76 (98.32)	19.03 (14.00)	3.11 (1.11)
12 months	183.99 (43.45)	183.99 (43.52)	24.93 ()	3.33 (0.65)
13 months	220.00 (68.22)	220.00 (68.24)	26.02 (8.04)	3.45 (0.62)
14 months	233.00 (118.79)	333.53 (119.31)	29.30 (19.39)	3.71 (1.32)
15 months	252.50 (153.00)	252.61 (153.07)	34.20 (23.44)	4.29 (0.52)
16 months	259.00 (101.21)	259.00 (101.27)	35.93 (22.33)	4.53 (0.22)
17 months	301.01 (99.73)	301.01 (100.22)	37.56 (22.44)	4.88 (1.01)
18 months	320.00 (120.00)	320.00 (121,22)	43.21 (12.22)	4.93 (0.72)

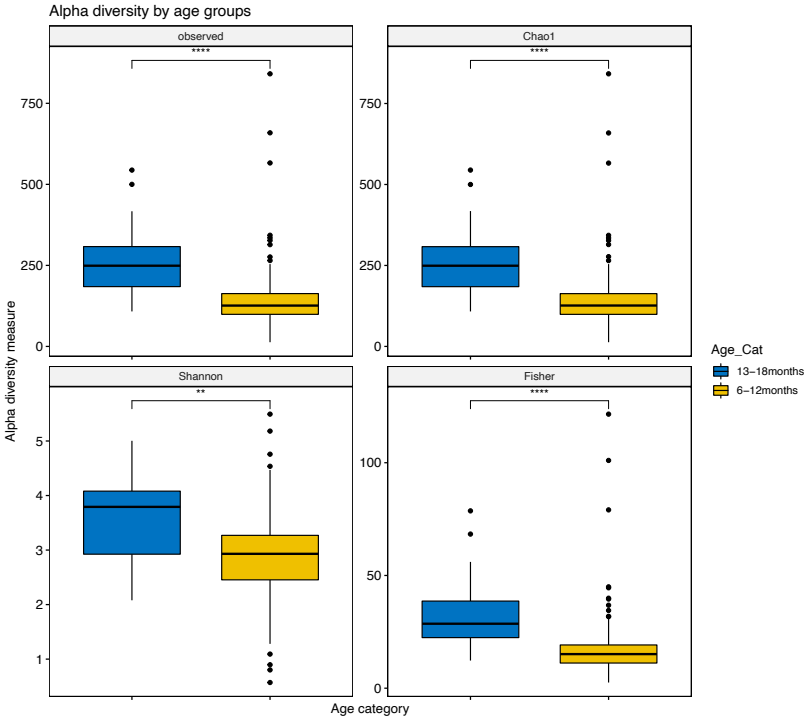

Figure 4.6: Alpha diversity by age group.

Significant differences were detected in Observed OTU, Chao1, Shannon and Fisher diversity metrics between 6 – 12 months old and 13 -18 months old age groups.

To determine the differences in gut microbiota composition between age groups and age categories, Bray-Curtis distances were computed using Adonis. PCoA plots were used to visualise differences in clustering (**Figure 4.7**). Age, analysed for each month (R2 = 0.06, P = 0.001) and as categorized in 6 to 12 and 13 to 18 months age groups (R2 = 0.01, P = 0.001), showed significant differences. At the phylum level, the relative abundance of Actinobacteria was higher, while the relative abundance of Bacteroidetes was lower in the younger age group (P = 0.01). At the genus level, the relative abundance of *Bifidobacterium* was higher in the young age group (P = 0.01).

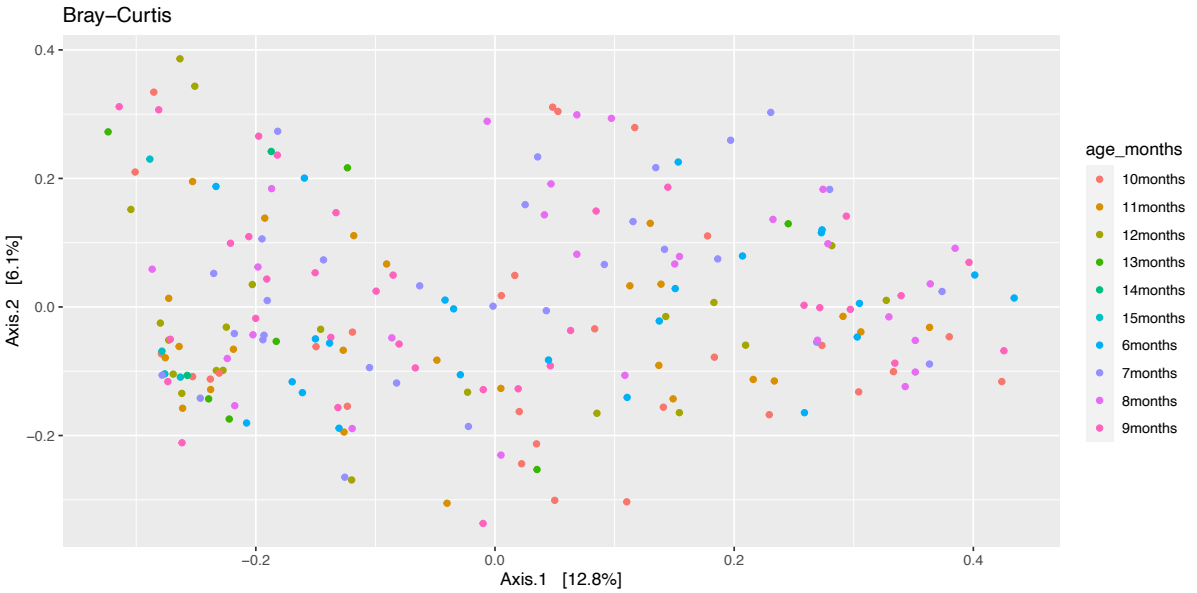

Figure 4.7: Beta diversity by age in months.

PCoA plots for Bray-Curtis dissimilarly distance computed by age in months and analysed using PERMANOVA with 999 permutations. Axis 1 explains 12.8% of the difference, while axis 2 explains 6.1%.

No significant differences were observed in Observed OTU (P = 0.13), Chao1 (P = 0.082), Fisher (P = 0.112) and Shannon (P = 0.721) diversity metrics with gender when calculated using the non-parametric Wilcoxon Rank Sum Test (**Figure 4.8**). Although not statistically significant, alpha diversity for males was slightly higher than for females. For males, the median (SD) computed for Observed OTU Chao1, Fisher and Shannon's diversity were 142.00 (25.26), 142 (87.00), 16.57 (12.21) and 3.02 (0.71), respectively. While for females, the median (IQR) computed for Observed OTU, Chao1, Fisher and Shannon's diversity were 127.00 (106.86), 127.00 (106.92), 15.19 (15.63) and 2.89 (0.82), respectively.

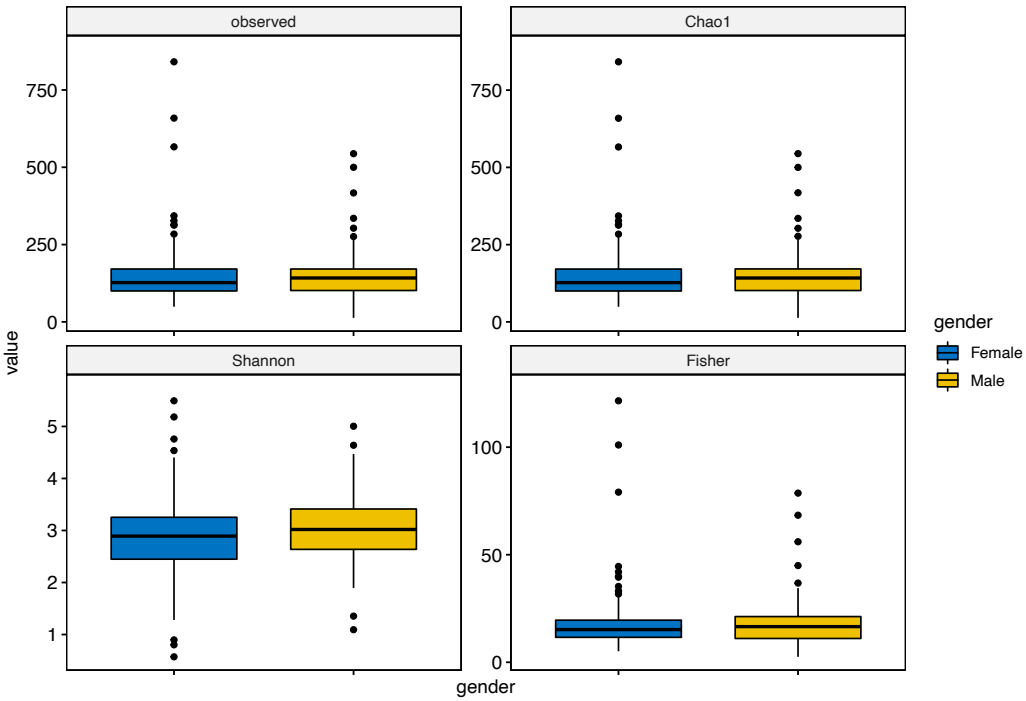

Figure 4.8: Alpha diversity with gender.

No significant differences were detected in Observed OTU, Chao1, Shannon and Fisher diversity metrics between males and females.

Bray Curtis distances were then calculated to determine differences in microbiota composition by sex. Significant differences were observed by PERMANOVA with 999 permutations (R2 = 0.012, P = 0.002). However, the significant difference observed in beta diversity by sex was lost after adjusting for age (**Figure 4.9**). Age-adjusted linear regression analysis did not detect any association between sex and gut microbiota composition (P = 0.823).

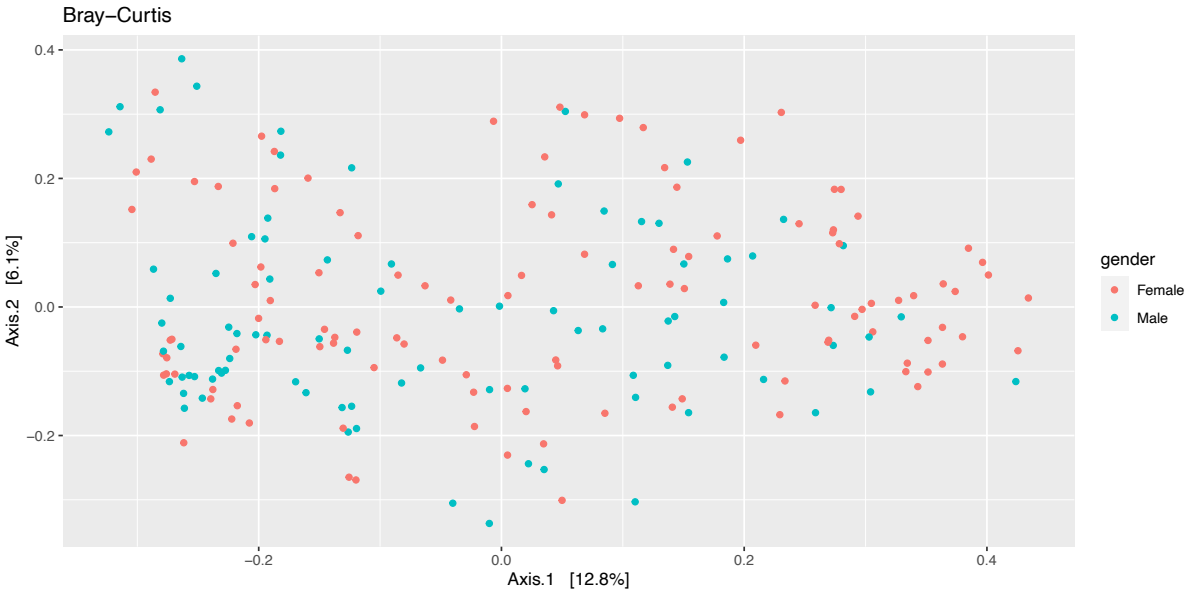

Figure 4.9: Beta diversity by age in months.

PCoA plots for Bray-Curtis dissimilarly distance computed by sex and analysed using Permanova with 999 permutations.

4.3.4 Gut microbiota composition with clinical characteristic

The effect of clinical characteristics; diarrhoea: analysed as current, history or any diarrhoea, vomiting: analysed as current, history or any vomiting, cough: analysed as current and history of cough, current fever, antibiotic usage and usage of antimalarial on gut microbiota composition and diversity were tested using both alpha and beta diversities. There were no significant differences in gut microbiota composition and diversity based on all the tested clinical characteristics as per the calculated alpha and beta diversity metrics. **Figure 4.10** and **Figure 4.11** differences in alpha diversity and associated PCoA plots for diarrhoea and vomiting respectively. The effect of acute malnutrition on gut microbiota composition detected no significant differences for both alpha and beta diversity.

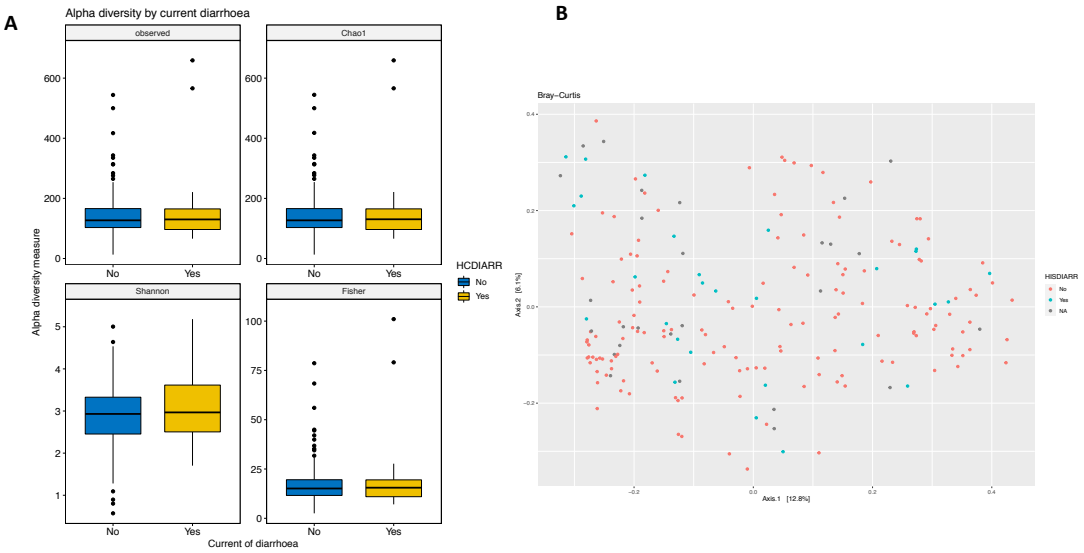

Figure 4.10: Alpha and beta diversity with current diarrhoea.

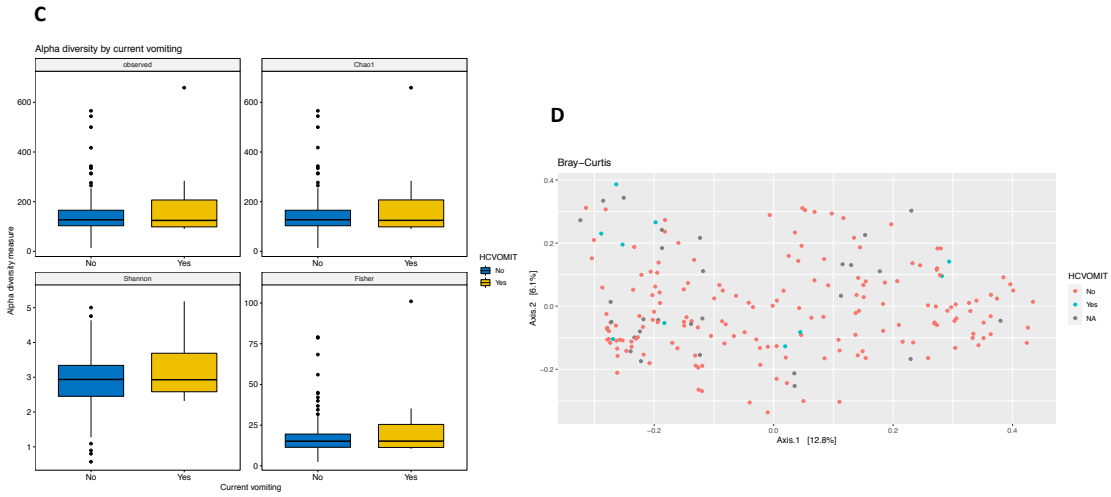

Figure 4.11: Alpha and beta diversity with current vomiting.

4.3.5 Gut microbiota composition with water source and boiled drinking water

The effect of water source and use of boiled water on gut microbiota composition and diversity was also tested using alpha and beta diversity. Using the non-parametric Wilcoxon Rank Sum Test, no significant differences were observed in gut

microbiota composition and diversity with either water source and drinking boiled water for both alpha and beta diversities (Figure 4.12).

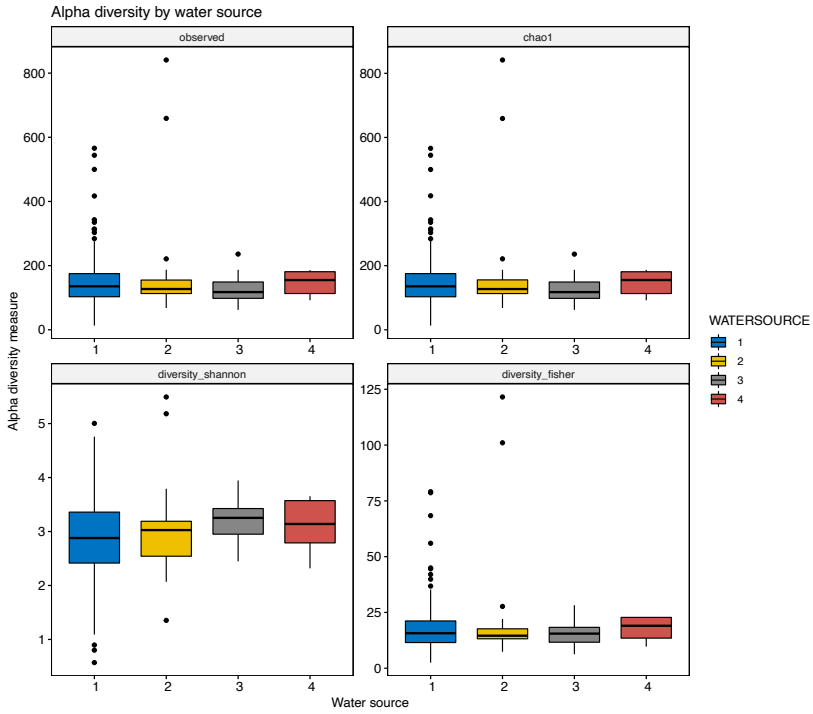

Figure 4.12: Alpha diversity with water source.

4.3.6 Gut microbiota composition with Bacterial pathogen exposure

The effect of common bacterial pathogens that Malawian children aged 6 to 18 months are exposed to on the gut microbiota composition and diversity was also tested using both alpha and beta diversities. Tested bacterial Pathogens included EAEC, EPEC, ETEC, *B. fragilis*, *Campylobacter*, *Shigella*/EIEC and *Salmonella*. For EAEC, Significant differences were detected in Observed OTU (P = 0.031) and Chao1 (P = 0.034) but not in Shannon (P = 0.062) and Fisher (P = 0.071) diversity metrics (**Figure 4.13**). This difference was however lost after adjusting for age. For *B. fragilis*, significant differences in Fisher (P = 0.01) and Shannon (0.01) but not in

Observed OTU (P = 0.062) and Chao1 (P = 0.057) diversity metrics. As with EAEC, significance was lost after adjusting for age. No significant differences were observed in alpha diversity for EPEC, ETEC, *Campylobacter*, *Shigella*/EIEC and *Salmonella.* The statistical analysis of the computed Bray – Curtis distances using Permanova with 999 permutations did not detect any significant differences in gut microbiota composition and diversity between participant exposed to EAEC (R^2 = 0.004, P = 0.062), EPEC (R^2 = 0.006, P = 0.059), ETEC (R^2 = 0.003, P = 0.081), *B. fragilis* (R^2 = 0.007, P = 0.07), *Campylobacter* (R^2 = 0.008, P = 0.072), *Shigella*/EIEC (R^2 = 0.001, P = 0.063) and *Salmonella* (R^2 = 0.007, P = 0.073) and those not exposed (**Figure 4.14**). Age adjusted regression analysis to determine the association between EAEC and *B. fragilis* exposure and gut microbiota composition did not detect any significant association (P = 0.072 and 0.061 respectively).

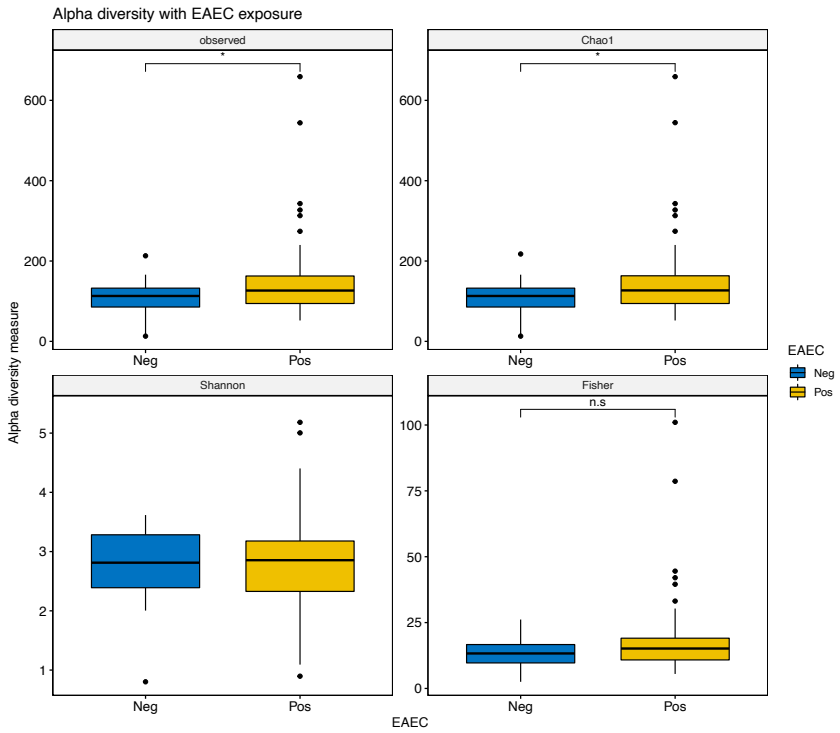

Figure 4.13: Alpha diversity with EAEC exposure.

Significant differences in Observed OTU and Chao1 indices and not in Shannon and Fisher indices

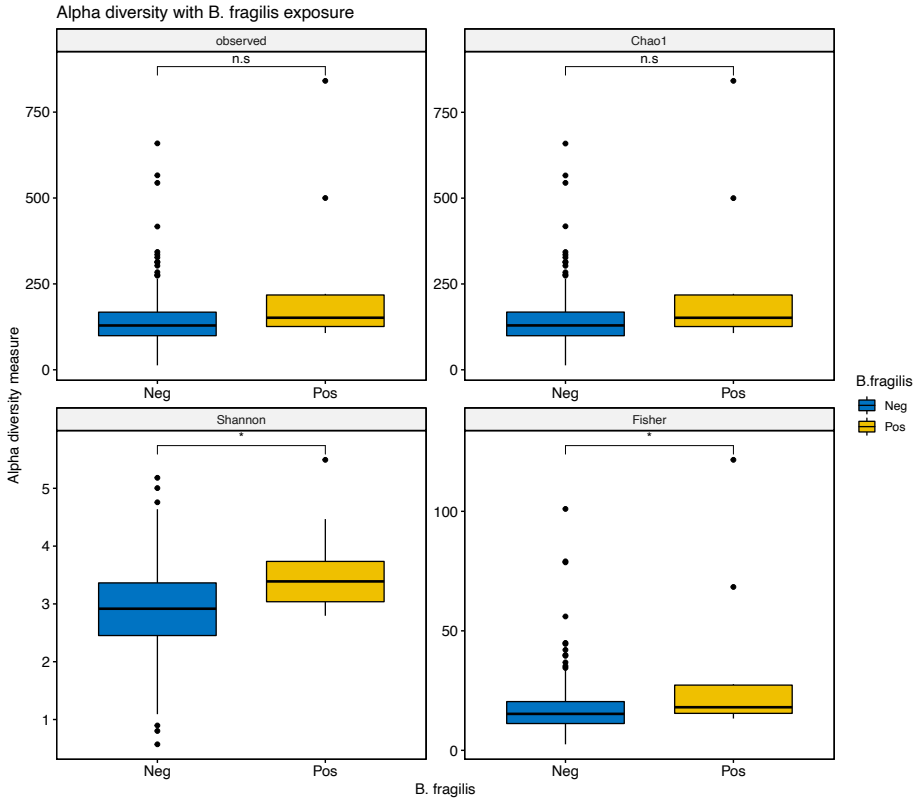

Figure 4.14: Alpha diversity with B. fragilis exposure.

There were significant differences in Shannon and Fisher indices but not in Observed OTU and Chao1 index.

4.3.7 Differences in gut microbiota composition with viral pathogen exposure

The effect of viral pathogens, enterovirus, norovirus, adenovirus and sapovirus, on the gut microbiota composition was also tested. No significant differences in alpha and beta diversities were detected for all viral pathogens. **Figure 4.15** show the alpha diversity measures for sapovirus.

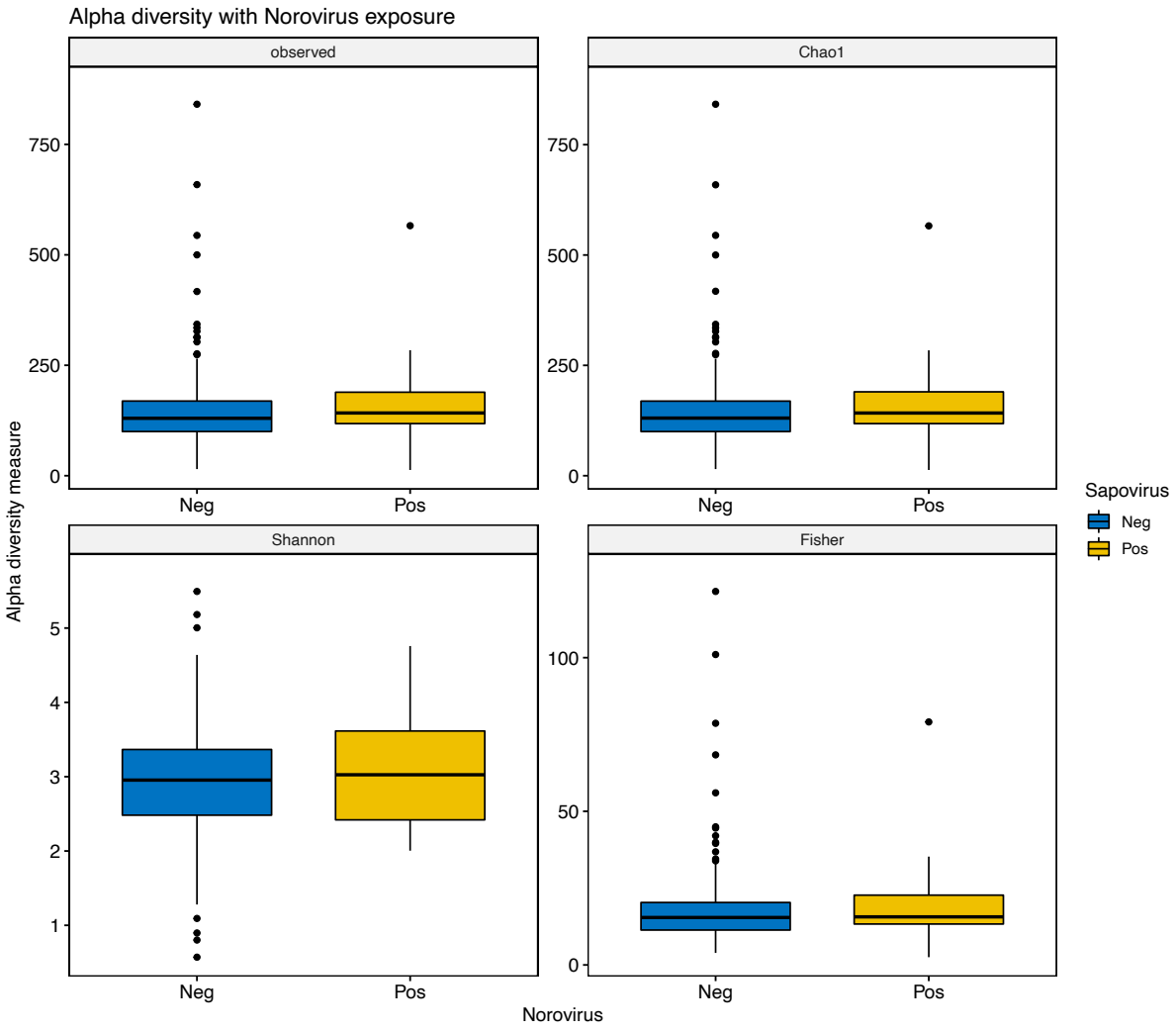

Figure 4.15: Alpha diversity with sapovirus exposure. No significant differences were detected.

4.3.8 Differences in gut microbiota composition with parasitic pathogen exposure

The effect of *Giardia* and *Cryptosporidium* on gut microbiota exposure was also tested. Significant differences were detected in alpha diversity with *Giardia* exposure but not with *Cryptosporidium (***Figure 4.16***)*. The gut microbiota of children with *Giardia* exposure had relatively high Observed OTU, Chao1, Fisher and Shannon diversity values compared to those not exposed to *Giardia*. For *Giardia* exposed children, the median (SD) values for Observed OTU, Chao1, Fisher and Shannon diversity were 189.00 (98.11), 190.00 (98.20), 26.53 (9.32) and 3.17 (0.98). For

Giardia unexposed children, the median (SD) values for Observed OTU, Chao1, Fisher and Shannon diversity were 156.00 (76.22), 156.00 (76.30), 20.12 (8.97) and 2.81 (0.79). The significant difference was also observed when beta diversity was computed for *Giardia* (R^2 = 0.009, P = 0.005) (**Figure 4.17**). However, age-adjusted regression analysis did not detect any significant association between *Giardia* and exposure and gut microbiota composition (P = 0.062). No significant difference was detected in beta diversity with *Cryptosporidium* exposure (R^2 = 0.006, P = 0.212).

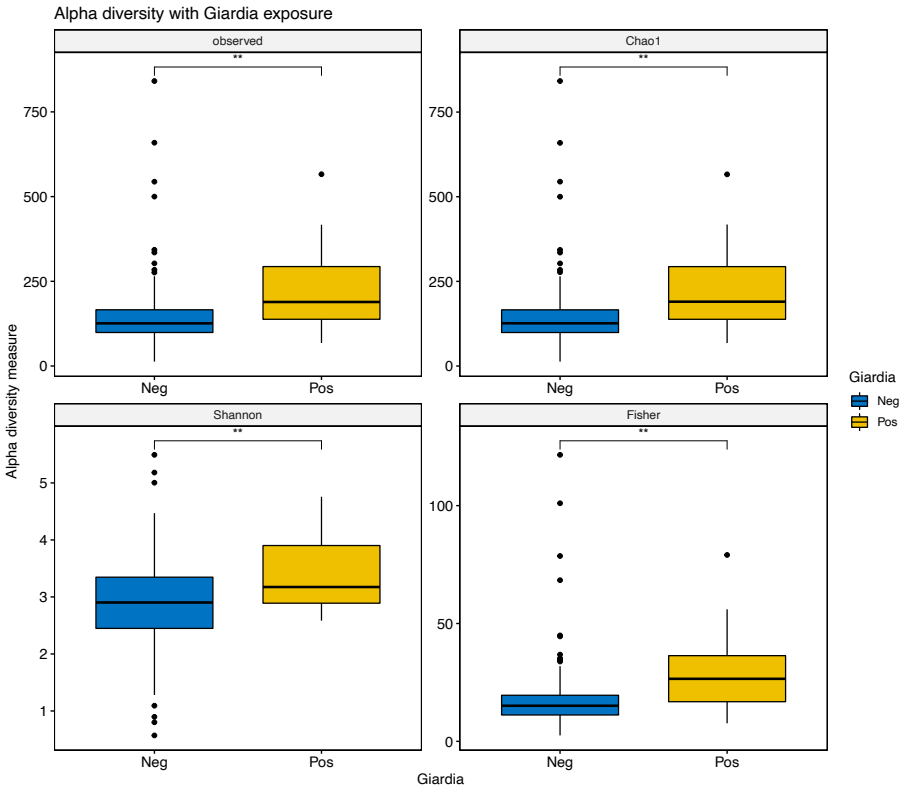

Figure 4.16: Alpha diversity with Giardia exposure.

Significant differences were detected in Observed, Chao1, Shannon and Fisher metrics.

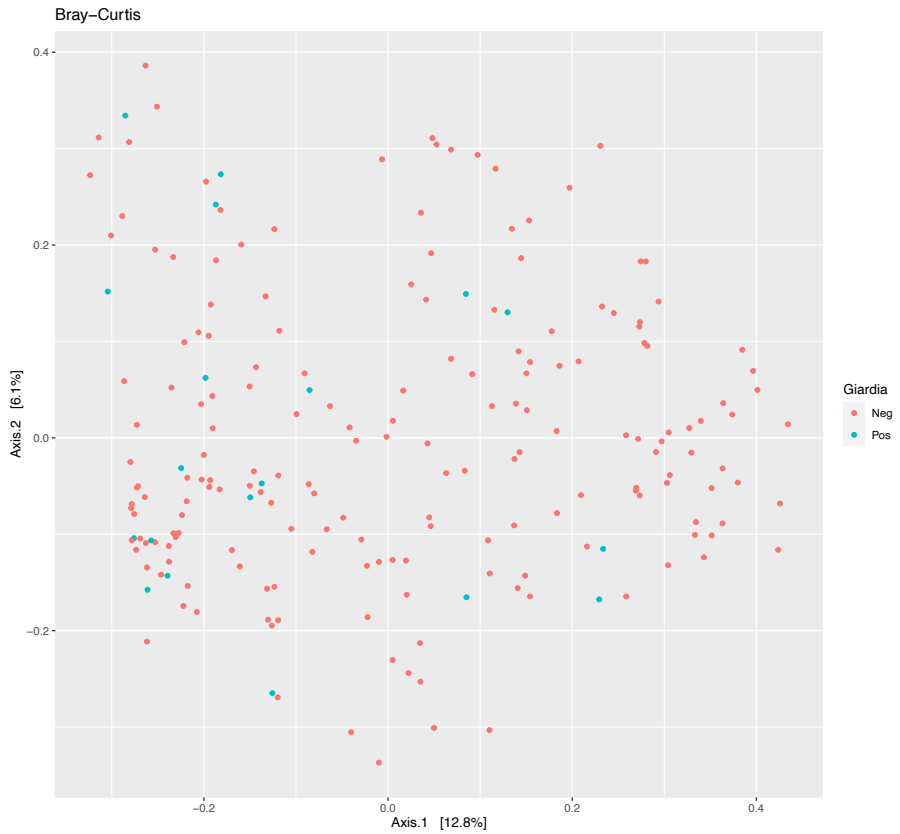

Figure 4.17: Beta diversity with Giardia exposure.

PCoA plots for Bray-Curtis dissimilarly distance computed by sex and analysed using Permanova with 999 permutations. Significant differences were detected in gut microbiota composition with Giardia exposure.

4.4 Discussion

Pathogen colonisation and infection of the gut are major factors that influence gut microbiota composition. This study has provided a general description of the changes in gut microbiota composition and diversity in 60 Malawian children from 6 to 18 months of age assessed differences in gut microbiota composition with clinical characteristics and markers of social-economic status. It then describes longitudinal differences in gut microbiota composition in these 60 healthy Malawian children exposed to different enteric pathogens over 12 months.

This study has agreed with finding from previous studies that have reported that the gut microbiota matures with age. The abundance and diversity of the gut microbiota significantly increased with age. Also, in agreement with previous studies, the relative abundance of some gut microbiota components such as *Bifidobacterium* decreased with age. In contrast, other components, mainly those belonging to the Bacteroidetes phylum, increased with age. Overall, *Bifidobacterium* was the most dominant gut microbiota member. *Bifidobacterium* thrives on breastmilk derived carbohydrates. More than 80% of children in this study were still breastfeeding by the end of the follow-up period. After weaning, the introduction of other food drives the shift in gut microbiota composition. Age-related microbiota changes have been widely reported, with older children having a more diverse and rich microbiota than younger children (Beller et al., 2021; Carter, Olm, & Sonnenburg, 2022; Kamng'ona et al., 2019; Kortekangas et al., 2020). Gut microbiota composition and diversity increase with age until age 3, when it stabilises and becomes more adult-like (Carter et al., 2022; F. Turroni et al., 2021). Gut microbiota abundance and diversity was assessed by computing alpha and beta diversities for each of the 12 individual

months and using age grouped into two categories (6 to 12 months and 13 to 18 months). In agreement with findings from previous studies, gut microbiota abundance and diversity increased over time for most of the individual months. They were very notable when the 13 to 18 age group was compared to the 6 to 12 age group. The study has demonstrated that sex affects gut microbiota composition. This was shown using Bray Curtis beta-diversity, but not alpha diversity metrics tested in this study. Sex did not significantly affect the gut microbiota when the analysis was adjusted by age.

Giardia, *B. fragilis* and EAEC were the only pathogens that demonstrated a significant effect on gut microbiota composition. For *B. fragilis* and EAEC, the difference was only observed when a cross-sectional analysis of alpha diversity metrics was conducted. Significance was lost for both alpha and beta diversities once the analysis adjusted for age. *Giardia* was the only pathogen that demonstrated significant differences in alpha and beta diversity measures. This was, however, also lost once age adjustments were made. Enterovirus, norovirus, adenovirus, and sapovirus did not significantly affect gut microbiota composition and diversity. Similarly, no significant differences were detected in gut microbiota composition with diarrhoea, vomiting, cough, antibiotic use, antimalarials, acute malnutrition, water source, and drinking boiled water. The loss in significance for *Giardia, B. fragilis* and EAEC after adjusting for age in months could be explained by loss of statistical power due to most pathogens' low pathogen positivity rate. This was because a healthy cohort was used and because of the study's longitudinal nature.
EAEC and *B. fragilis* affected gut microbiota composition while EPEC, ETEC, *Campylobacter, Shigella*/EIEC and *Salmonella* did not. EAEC and *B. fragilis*

exposure affects gut microbiota composition based on alpha diversity results. In Chapter 3, EAEC exposure was highly prevalent but was not associated with any clinical characteristics such as diarrhoea, fever, and vomiting. Despite not having any association with clinical symptoms, the significant association of EAEC with changes in the diversity of gut microbiota could have immediate or long-term direct or indirect impacts on a child's health. The interaction between EAEC and viral pathogens in the gut has been shown to promote gut microbiota dysbiosis. A study that used EAEC, EPEC, rotavirus and norovirus to determine the effect of bacterial and viral interaction on the gut microbiota showed that the presence of EAEC increased gut microbiota dysbiosis compared to viral pathogens alone (Mathew et al., 2019). In addition to gut microbiota changes, the study also reported that co-infection with the pathogenic E. *coli* worsened the clinical conditions of the children. EAEC and *Giardia* co-infections have been reported to enhance malnutrition in weaned C57B1/6 male mice (Bartelt et al., 2017). Gut microbiota dysbiosis due to these highly prevalent enteric pathogens could also create an environment that could promote the establishment and proliferation of other pathogens that can cause clinical disease. The contribution of EAEC and enterovirus mediated gut dysbiosis to stunting and low vaccine effectiveness should be studied.

Although this study did not reveal an effect of most bacterial enteric pathogens on the gut microbiota, clinical or mouse model studies that have investigated the impact of specific bacterial pathogens on the gut microbiota have reported some significant differences (Gillis et al., 2018; Haak et al., 2020; Mathew et al., 2019). An observational study comparing gut microbiota composition and function in typhoid fever patients (TF), non-typhoid febrile patients (N-TF) and healthy controls

demonstrated a decrease in Actinobacteria and enrichment of Proteobacteria in TF patients when compared to the healthy control group. The study showed that *Streptococcus* was enriched in TF at the genus level. Still, there was a reduction in beneficial anaerobic bacteria, including *Ruminoccocccus* and *Faccalibacterium,* compared with N-TF (Haak et al., 2020).

Although no significant effect in the gut microbiota was observed with viral exposure, some studies have shown some significant but temporal effects on the gut microbiota with viral pathogen exposure (A. H. Kim et al., 2020; Mathew et al., 2019). There has been a growing interest in understanding low oral vaccine effectiveness reported in developing countries and the gut microbiota. A study that aimed to determine the influence of enterovirus and gut microbiota on vaccine effectiveness concluded that exposure to enteric viruses influences vaccine effectiveness more than the gut microbiota (Praharaj et al., 2019). Lack of association between the gut microbiota and oral vaccine effectiveness was also reported in another study in India (Parker et al., 2018).

Based on the computed alpha and beta metrics, *Giardia* exposure significantly affected gut microbiota composition. This was, however, lost after adjusting for age. Common parasitic pathogens, *Giardia* and *Cryptosporidium,* have been reported to affect gut microbiota composition and diversity (Riba et al., 2020; Toro-Londono, Bedoya-Urrego, Garcia-Montoya, Galvan-Diaz, & Alzate, 2019). *Giardia* infection is more prevalent in young children than in adults, but it is very common in children from developing countries where the levels of sanitation are low (Dunn & Grider, 2020). In the cohort reported in **Chapter 3,** the relative abundance of Giardia

exposure was lower in children less than 13 months old. A steady increase in exposure was observed from 13 months of age. In Malawi, children at this age are very active and are left to be playing outdoors where they could be exposed to unhygienic conditions and that they are also eating supplementary food. Although *Giardia* infection is a common cause of diarrhoea in the developed world, Giardia seems to protect against diarrhoea in children living in areas associated with poor sanitation, such as the SSA. A study that aimed at understanding mechanisms behind the protective nature of Giardia by co-infecting human adenocarcinoma cells with Giardia and EPEC reported that there was a reduction in bacterial invasion, colitis, weight loss, the appearance of blood in stool and softening of faecal matter (Manko-Prykhoda et al., 2020). Gut microbiota dysbiosis was reported in Colombian children with *Cryptosporidium* and *Giardia* infection than uninfected children (Toro-Londono et al., 2019). The loss and lack of significant association between *Giardia* and *Cryptosporidium* exposure observed in this study are likely due to the inadequate statistical power, as the exposure was relatively rare.

Although this study focused on changes in the gut microbiota with pathogen exposure, the relationship between the gut microbiota and pathogens is two-way; 1) Pathogens affect gut microbiota composition and function, 2) the gut microbiota controls pathogens through colonisation resistance. The gut microbiota always defends the host against pathogen colonisation using direct and indirect colonisation resistance mechanisms. Symptomatic clinical infection occurs when the pathogen has overpowered the gut due to high pathogen load and virulence factors or other host factors. This study cohort comprised healthy participants who had a few episodes of clinical symptoms during the follow-up period. However, the use of

molecular detection methods allowed for the detection of pathogens even at a very low pathogen load. Healthy gut microbiota is expected to control such low pathogen load. This could also explain the lack of significant effect with moth of the tested pathogens. Using a quantitative analysis would help determine the impact of pathogen exposure on the gut microbiota while factoring in pathogen load. The scope of this study did not account for that.

4.4.1 Strengths and weaknesses

This is the only study that has attempted to have a longitudinal understanding of how enteric pathogens that Malawian children are exposed to would affect gut microbiota composition using a cohort of healthy children. Understanding the effect of asymptomatic enteric pathogens exposure on gut microbiota composition will help unravel and answer some of the health challenges in SSA, such as low oral vaccine effectiveness.

This study had several limitations. The loss to follow-up of participants and samples reduced the statistical power for longitudinal data analysis. Although the study was clearly adequately powered to describe longitudinal changes with age, it was not adequately powered to show longitudinal changes with pathogen exposure. This could be because there are no changes with pathogen exposure or that there are changes but are more subtle. The change in gut microbiota composition could be insignificant, or it could be of more functional importance than numerical. This could be resolved by increasing the sequencing depths or the number of samples to detect a smaller effect size.

Besides challenges with statistical power, this study focused on enteric pathogen exposure in the gut as defined by pathogen detection in stool samples using molecular detection and not diarrhoeal disease. The results would be different if we focused on enteric pathogen infection. Since molecular methods are highly sensitive and can detect very minimal pathogen load, the amount and impact of the pathogens in the stool samples may not have been enough to affect the overall gut microbiota composition significantly. Pathogen exposure may also not result in a clinically important disease state. The virulence or pathogen load in pathogen infection may need to be significant to cause gut microbiota dysbiosis (Lertsethtakarn et al., 2018). It is, therefore, possible to observe a substantial difference in gut microbiota composition and function when the focus was on pathogen infection, even when there were no significant differences observed when the focus was on pathogen exposure. Seasonality was not accounted for in the analysis. There were different seasons experienced during the study period, and recruitment was spread over several months, meaning the impact of season was not consistent in the cohort. Seasonality is associated with pathogen exposure and infection and may also affect gut microbiota composition due to differences in environmental conditions and food uptake (Kirolos et al., 2021; Thindwa, Chipeta, Henrion, & Gordon, 2019).

The association between changes in the gut microbiota and enteric pathogen exposure were more observational, where the gut microbiota was assessed based on pathogen presence and absence. However, this field is progressing to a more functional and mechanical way of study, adding more value and nuance. This, however, uses data generated through shotgun metagenomic sequencing, which was not done in this study. The study used data generated using 16S taxonomic profiling,

which is only sensitive up to the genus level and not species level. It is important to look at species level differences since some specific species of organisms like *Bifidobacterium* have been shown to promote health.

4.5 Conclusion

This study has described the microbiota pattern over time in 60 children. It also examined the longitudinal association of pathogen exposure with changes in microbiota composition. It showed that *Giardia, B. fragilis,* and EAEC affects the gut microbiota composition of Malawian children. There was no significant association between most enteric pathogens tested in this study, clinical characteristics and markers of economic status. Although the study was clearly adequately powered to describe longitudinal changes with age, it was not adequately powered to show longitudinal changes with pathogen exposure.

Chapter 5: Gut Microbiota Profiling of Healthy Exclusively Breastfed Malawian Children and the Genomic Characterisation of Bifidobacterium Isolated from their Stool

5.1 Introduction

Bifidobacterium is an important member of the gut microbiota throughout the entire human life. It is, however, predominant during early life (O'Toole & Claesson, 2010; Selma-Royo et al., 2019). It constitutes approximately 80% of the gut microbiota in exclusively breastfed children, while in adults, *Bifidobacterium* accounts for between 3% to 6% of the normal microbiota (D. H. Taft et al., 2018; F. Turroni et al., 2018; F. Turroni et al., 2021). *Bifidobacterium* is regarded as a “beneficial” member of the gut microbiota because of the important roles it plays in carbohydrate metabolism, maintaining gut physiology, immune system modulation, and colonisation resistance (Aw & Fukuda, 2019; F. Turroni et al., 2018). To better understand the beneficial properties of *Bifidobacterium*, genomic studies have been done to understand the genetic composition and link that to different beneficial properties of *Bifidobacterium*.

Although several studies that aimed at identifying, characterising and understanding the mechanisms behind the beneficial effects of *Bifidobacterium* to the human host have been done, such information is not available for SSA countries, including Malawi, where the burden and impact of infectious diseases are the highest. Although studies in SSA have described *Bifidobacterium* as a member of the gut microbiota, studies aimed at understanding its genomic characteristics are rare (Cheung et al., 2016; Kortekangas et al., 2020; Lackey et al., 2019; Smith et al., 2013; D. Taft et al., 2019). Based on studies conducted in other parts of the world, the genomic understanding of *Bifidobacterium* has provided valuable information on

their characteristics, interaction with other microbiota components and the host environment (Duranti, Longhi, Ventura, van Sinderen, & Turroni, 2020; Henrick et al., 2021; Li et al., 2021; Longhi, van Sinderen, Ventura, & Turroni, 2020; Tarracchini et al., 2021). Through genomic studies, it is established that *Bifidobacterium* abundance in infancy is promoted by the presence of genes or gene clusters responsible for carbohydrate metabolism, particularly those responsible for Human Milk Oligosaccharides (HMO) degradation (Bode, 2012; Thomson et al., 2018). The metabolism of HMOs by *Bifidobacterium* produces fatty acids as the by-product. Fatty acid are important for colonisation resistance.

In Malawi, the strains of *Bifidobacterium* found in exclusively breastfeeding infants and their genomic characteristics are not known. This represents a real bottleneck if we are to understand their beneficial properties for the next stage of therapy development. Importantly, understanding the HMOs metabolism clusters encoded in *Bifidobacterium* isolated from Malawian children and their antimicrobial properties will provide insights into their colonisation resistance potential against important enteric pathogens in Malawi and the SSA region. The work being reported in this chapter was, therefore, aimed at identifying *Bifidobacterium* isolated from exclusively breastfed Malawian infants and describing genes responsible for HMO digestion.

5.1.1 Hypothesis

Bifidobacterium in the gut microbiota of Malawian exclusively breastfed infants may be distinct from those seen in other regions of the world, and they may possess distinct HMO gene clusters essential for colonisation resistance.

5.1.2 Study objectives

1. Recruit 30 healthy, exclusively breastfeeding children under 18 weeks old and isolate *Bifidobacterium* from their stool.
2. Use microbiome profiling on the stools of the study participants, focusing on the abundance of *Bifidobacterium*.
3. Use Illumina whole-genome sequencing to characterise the *Bifidobacterium* isolates from objective 1, focusing on human milk oligosaccharide digesting enzyme.

5.2 Methods

5.2.1 Study type and site

This was a single site cross-sectional pilot study conducted at Ndirande Health Centre in Blantyre, Malawi. Details are as indicated in **Chapter 2, section *2.1***.

5.2.2 Consenting and recruitment of healthy exclusively breastfed children

Mothers attending routine under-five clinics at NHC were sensitised about the study. Only participants that consented to take part in the study were recruited. Details are as shown in **Chapter 2, section *2.3.2***.

5.2.3 Stool sample collection

When possible, stool samples were collected from nappies by well-trained study nurses. Otherwise, mothers collected the stool samples at home and brought them to the study clinic within four hours. Samples were then transferred to MLW for *Bifidobacterium* isolation and stool DNA extraction. Details are as shown in section *2.3.2.*

5.2.4 Isolation and confirmation of Bifidobacterium

Stool samples were diluted and cultured anaerobically in a Ruskinn anaerobic chamber, using MRS selective media with 50mg/ litre of L-cysteine and Mupirocin. Five *Bifidobacterium* presumptive colonies were collected and subcultured until pure colonies were achieved. We used 16S rRNA targeted Sanger sequencing and a standard biochemical assay -Fructose 6 Phospoketolase (F6PK) assay to confirm that the isolates are *Bifidobacterium*. Sanger sequence reads were blasted against

the NCBI database. The performance of 16S rRNA Sanger sequencing and F6PK *Bifidobacterium* isolate confirmatory tests was also compared. Details are as indicated in **Chapter 2, section *2.3.2.***

5.2.5 Bacteria DNA extraction and sequencing

For microbiota profiling, DNA extraction was performed on stool samples using MPBio fast DNA kit for soil according to the manufacturer's instruction with an added bead-beating step. Libraries were prepared targeting the V1V2 region of the 16S rRNA gene using PCR, cleaning and quantification of the PCR products and pooling of the products. Sequencing was done using Illumina sequencing.

For whole-genome sequencing, manual DNA extraction on *Bifidobacterium* isolates was done using Phenol Chloroform. Eurofins did library preparation and whole-genome sequencing of the isolates. Details are as shown in **Chapter 2, section *2.3.2***.

5.2.6 Microbiome data analysis – 16S rRNA sequence data

16S rRNA sequence reads were processed and analysed using Quantitative Insights into Microbial Ecology (Qiime) – 2 (Bolyen et al., 2019). Quality checking was done using the Babraham FASTQC application version: 0.11.8. Quality filtering and sequence variant calling were done using the Dada2 package within QIIME2. Silva reference database was used to assign taxonomy to the sequence variants. Details on how microbiome data was analysed are as indicated in **Chapter 2, section *2.5.2***.

5.2.7 Whole genome sequenced data

Unicyler and Prokka were used for genome assembly and annotation, respectively. Maximum Likelihood trees were generated using RaxML with 999 bootstraps and viewed on iTOL. Average Nucleotide Identity and Kraken were used to identify the Malawi genomes. ANI with default BLAST+ settings was used to calculate the Average Nucleotide Identity of the Malawi genomes by blasting the Malawi genomes against 88 publicly available *Bifidobacterium* genomes. Kraken was also used for taxonomic classification of the Malawi genomes by using MiniKraken2_v1_8GB database.

Thirty-five publicly available *B. longum* genomes representing *B. longum* subspecies were retrieved from NCBI and used to provide a broader context of the Malawi *B. longum* genomes.

HMO genes present in the Malawi *Bifidobacterium* genomes were identified by blasting the Malawi genomes against a custom database of 90 HMO genes. Details on how WGS data was analysed are as indicated in **Chapter 2, section *2.5.2.***

5.2.8 Statistical analysis

Variations in microbiota profiles were assessed by determining alpha and beta diversity measures. Statistical assessment of categorical variables was done using permutational multivariate analysis of variance (PERMANOVA) test using Pseudo-F statistical method with 4,999 permutations. Qiime2R, Phyloseq and ggplot2 R-packages were used.

Welch two-sample unpaired Student's t-test was performed to determine differences in the general participant characteristics and clinical characteristics between genders. The student t-test and Kruskal Wallis test was also used to compare HMO digesting genes present between and amongst *B. longum* subspecies clusters. R (version 3.6.1) and RStudio (version 1.2.1335) were used.

5.2.9 Ethical approval

Ethical approval was granted by the College of Medicine Research and Ethics Committee (P.10/17/2296) and the University of Liverpool Research and Ethics committee (Reference number: 2867). Details are as indicated in **Chapter 2, section *2.2***.

5.3 Results

5.3.1 Characteristics of study participants

Thirty participants were recruited from 9 to 26 January 2018. Of the 30 participants, 14 (46.7%) were females. There were no significant differences in the participant characteristics by sex (**Table 5.1**). The participants' median age was 2.45 months, with a range of 0.53 months to 3.61 months. All participants were healthy and exclusively breastfeeding at the time of recruitment. Three (10%) participants reported to have had a cough in the past two weeks. Only one participant was born through a caesarean section.

Table 5.1: General participant clinical characteristics

Parameter	All	Female	Male
Screened (%)	87 (100)	48 (55.2)	39 (44.8)
Recruited (%)	30 (100)	14 (46.7)	16 (53.3)
Median age (IQR)	2.45 (1.61 - 2.78) months	2.46 (0.53 - 3.61) months	2.43 (1.38 – 3.78) months
Exclusive breastfeeding	30 (100)	14 (100)	16 (100)
Cough	3 (10)	1 (7.1)	2 (12.5)
Temperature median (IQR)	36.45 (35.00 -37.10) ^{o}C	36.4 (35.0 -36.9) ^{o}C	36.5 (35.0 – 37.1) ^{o}C
Weight median (IQR)	6.00 (3.02 - 7.00) Kg	5.9 (3.2 – 6.8) Kg	6.2 (4 – 7) Kg
Height median (range)	60.00 (49.00 – 66.00) cm	59 (49 – 66) cm	60.5 (52 – 65) cm
MUAY median (range)	13.00 (10.00 – 16.00) cm	13 (10 – 15) cm	13 (10 – 16) cm
Weight for height	65.9 (53.2 -72.8)	64.3 (53.2 – 72.8)	67.2 (56.0 – 71.0)

5.3.2 Quantitative analysis of the gut microbial community in exclusively breastfed Malawian children

Microbial community profiling was carried out to investigate the composition of the microbiome and the amount of *Bifidobacterium* present. Analysis was conducted on a total of 2,461,042 sequence reads obtained after 16S rRNA gene sequencing of DNA extracted from the 30 collected stool samples. The median number of reads was 79,708, with a range of 68,164 to 100,990. After quality filtering, a total of 1,702,499 sequenced reads with a median of 55,714 (range of 36,393 to 70,958) were used for sequence variant calling. A total of 820 sequence variants were identified.

The infant gut microbiota was composed of eight phyla with Actinobacteria, Proteobacteria, Firmicutes and Bacteroides as the main phyla (**Figure 5.1**). Actinobacteria accounted for the highest median percent of reads, appearing in all the 30 samples with a median relative frequency of 35.76 (range of 11.76 to 73.55) %. Proteobacteria were the second most relatively abundant phyla, available in 29 samples with a median relative frequency of 27.97 (range of 0 to 63.71) %. Firmicutes were also present in all the samples with a median relative frequency of 15.62 (range of 4.07 to 42.25) %. Bacteroides were available in 14 samples with a relative frequency range of 0 to 74.68%. Epsilonbacteraeota, Patescibateria, Fusobacteria and Cyanobacteria were detected at lower levels and only observed in 4, 2, and 2 samples, respectively.

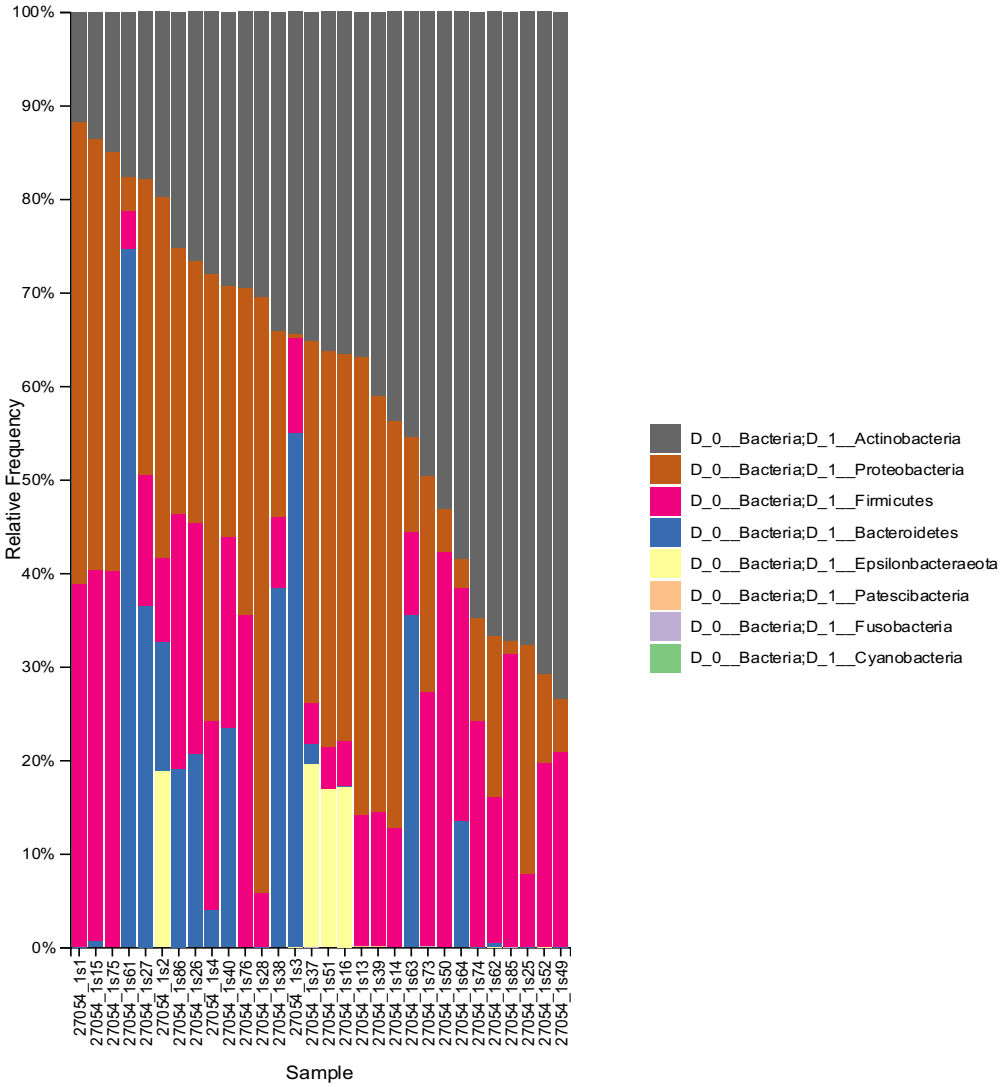

Figure 5.1: Gut microbiota composition at phylum level

The relative frequencies of each of the eight phyla is shown for each participant stool sample

At genus level, the most relatively abundant genus was *Bifidobacterium*. It was identified in all the 30 participants with a median relative abundance of 35.32% (range of 11.481% to 73.274%). The gut microbiota of these Malawian infants had a high abundance of *Enterobacteriaceae, Escherichia-Shigella, Streptococcus,* and *Veillonella* with a median relative abundance of 6.93 (range of 0 to 45.299) %, 7.355

(range of 0 to 37.991)%, 2.583 (range of 0 to 38.787)% and 1.751 (range of 0 to 17.099)% respectively. Species-level analysis was not done because 16S rRNA data is not sensitive enough in identifying organisms at the species level, thus only allowing genus-level analysis (Lackey et al., 2019; Schriefer et al., 2018).

As an exploratory analysis to determine if there are differences in the microbiota profiles based on participant characteristics, ability to isolate *Bifidobacterium* (positive or negative *Bifidobacterium* culture result) and based on the *B. longum* clusters identified, I conducted alpha diversity analysis using observed, Chao1 and Shannon, but no significant differences were observed with age (**Figure 5.2b**), gender (**Figure 5.2a)**, *Bifidobacterium* culture results and species type/ cluster identified based on WGS results (**.**

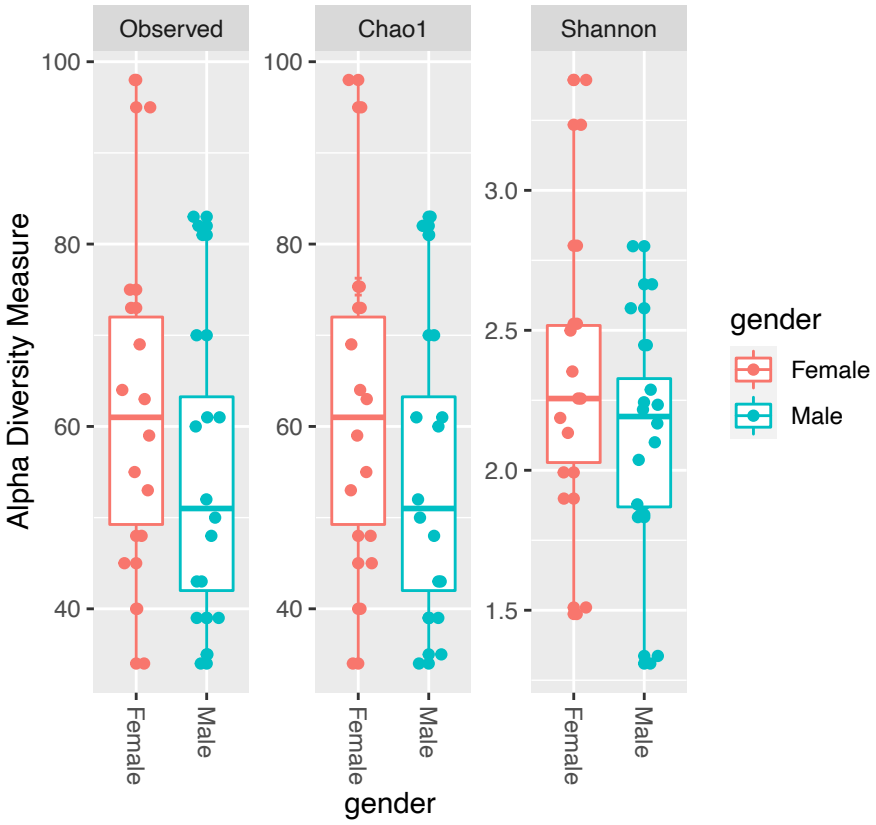

Figure 5.2a: Alpha diversity by sex. Alpha diversity was assessed using Observed, Chao1 and Shannon diversity matrices.

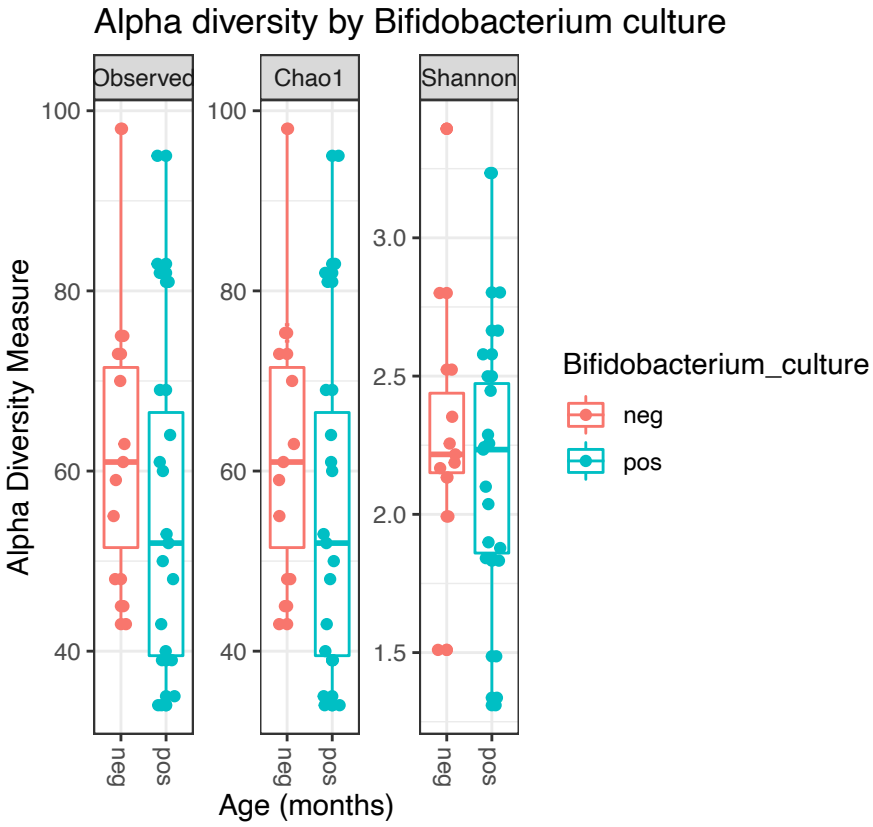

Figure 5.2b: Alpha diversity by Bifidobacterium culture. Alpha diversity was assessed using Observed, Chao1 and Shannon diversity matrices.

Beta diversity analysis using Jaccard and weighted and unweighted Bray-Curtis index, but no significant differences were observed with age, gender (**Figure 5.3a)**, *Bifidobacterium* culture results (**Figure 5.3b)** and species type/ cluster identified based on WGS results. Figure 3 is a representation of the Beta diversity results obtained by *Bifidobacterium* culture.

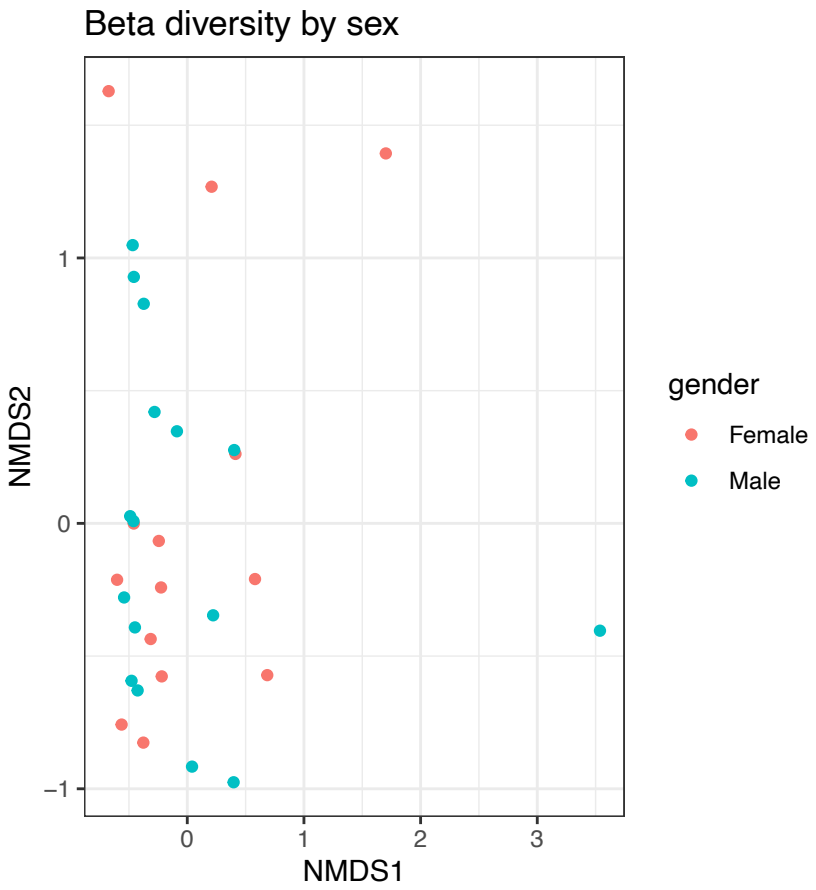

Figure 5.3a: Beta diversity by Sex - weighted Bray Curtis.

Beta diversity was assessed using Jaccard and weighted and unweighted Bray-Curtis diversity matrices for Bifidobacterium culture results.

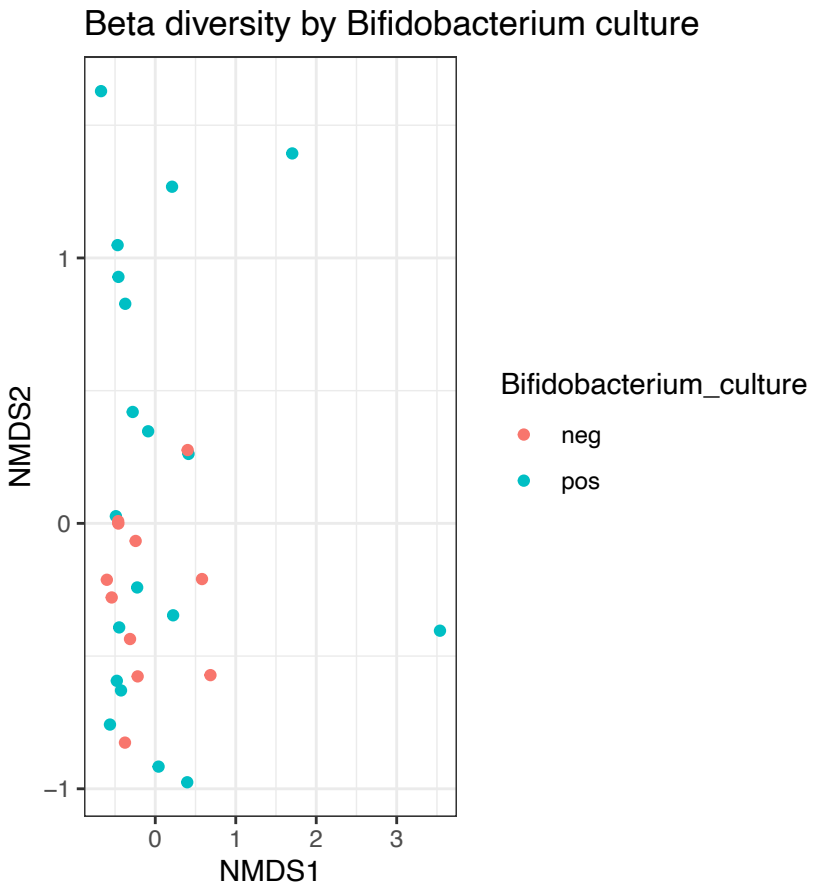

Figure 5.3b: Beta diversity – weighted by Bifidobacterium Culture Bray Curtis.

Beta diversity was assessed using Jaccard and weighted and unweighted Bray-Curtis diversity matrices for Bifidobacterium culture results.

5.3.3 Identification and general features of the Malawi Bifidobacterium genomes

The general characteristics of the *Bifidobacterium* genomes from Malawian infants were consistent with previously published *Bifidobacterium* genomes. The median total assembled length was 2.60 Mbp with a range of 2.3 to 2.71 Mbp. The longest genome was 5.4 Kbp and was obtained from *B. longum*. The median G + C content was 59.63%, with a range of 56.80% to 62.44%. *De novo* assembly generated between 22 and 149 contigs. B. pseudocatenulatum had the fewest contigs while B. longum had the most with a median of 22.67 and 43 contigs respectively. All

genomes had 2 rRNA and 1 tmRNA. The median number of tRNA was 57, with a range of 54 and 81. **Table 5.2** shows the genomic characteristics of all the Malawi genomes. Detailed Malawi genome Characteristics a.

Table 5.2: Identity and genomic characteristics of the Malawi Bifidobacterium

Species	Number of genomes	Subspecies	Median total length- Kbp (min – max)	Median contigs (min – max)	Median GC content (min – max)	Median tRNA (min - max)
B. bifidum	2	-	2.3 – 2.4	29, 47	62.4 (62.4 - 62.4)	56.5 (55 - 58)
B. breve	2	-	2.4	29, 34	58.9	54
B. longum	60	All	2.5 (2.4 - 5.4)	43 (25 – 149)	59.8 (59.5 – 60.4	58 (55 – 74)
	5	Cluster A	2.4 (2.4 - 2.5)	49.2 (34 - 58)	60.3 (60.3 - 60.4)	56.2 (56 - 57)
	11	Cluster B	2.7 (2.4 - 5.4)	47 (30 - 149)	59.8 (59.5 - 61.6)	55 (55 -74)
	44	Cluster C	2.5 (2.4 - 2.6)	43 (25 - 144)	59.6 (59.5 - 60.1)	58 (56 - 60)
B. pseudocatenulatum	3	-	2.4	22.67 (22 - 24)	56.8	56

Whole-genome sequencing data were available for 67 of 69 isolates. Using Kraken and ANI analysis, 60 isolates were identified as *B. longum at the* species level with three distinct clusters, which were designated as Cluster A, B and C. Two genomes were identified as *B. bifidum*, 2 as *B. breve* and 3 as *B. pseudocatenulatum* (**Figure 5.4** and **Table 5.2**). The two samples which did not provide WGS data were not

included in downstream analysis. However, these samples were initially identified as *B. longum* when their Sanger sequence reads were blasted against the NCBI database.

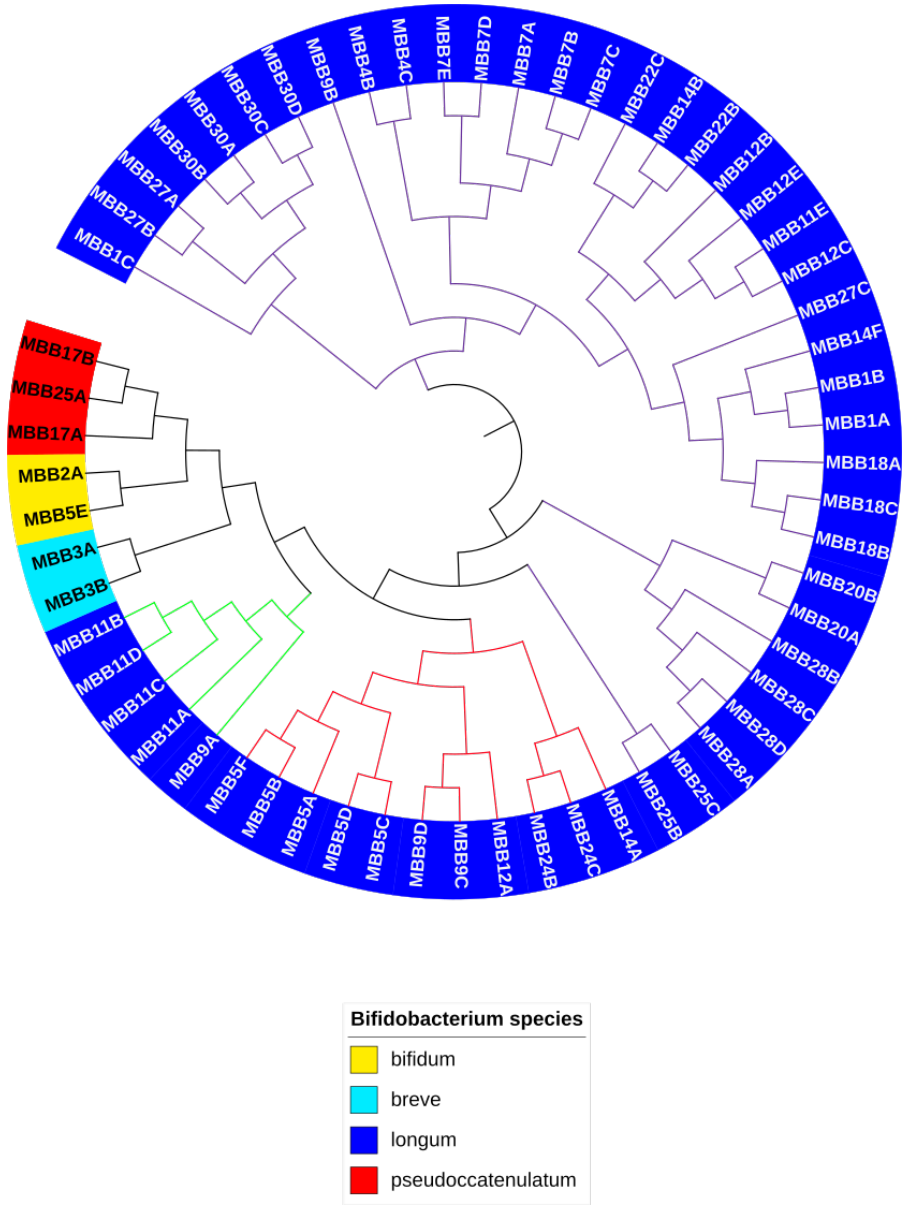

Figure 5.4: A dendrogram showing the phylogenetic distribution of Bifidobacterium isolated from exclusively breastfed Malawian children.

Three of the 60 genome assemblies were classified as B. pseudocatenulatum, two as B. breve, two as B. bifidum and 53 as B. longum. B .longum had three distinct clusters classified here as Cluster A (green branches), Cluster B (red branches) and Cluster C (purple branches).

Sixty samples were included in the whole genome sequence analysis. Sixteen of the 20 participants with WGS data were colonized by the same *Bifidobacterium* species/ cluster but 4 participants had two different *Bifidobacterium* species/ cluster isolated from their stool samples.

5.3.4 General features of *Bifidobacterium* longum genomes

Considering that *B. longum* was the most dominant species, these genomes were further examined at a sub-species level by placing them into an international context. To do this, 35 publicly available genomes belonging to subsp *longum, infantis* and *suis* were retrieved from the National Center for Biotechnology Information (NCBI) database (herewith referred to as global genomes) and were used to identify *B. longum* genomes from Malawian infants at the subspecies level. The Malawi *B. longum* in clusters A and B could not be confidently identified as belonging to a single subspecies because the genomes clustered together with global *B. longum* subsp. *longum*, *suis* and *suillum*. In contrast, cluster C genomes clustered with *B. longum* subspecies *infantis* only. Genomes in Cluster C was therefore identified as *B. longum* subsp *infantis.*

Notable genomic differences we observed by geography between the Malawi and global genomes (**Figure 5.5**). Global and Malawian genomes generally formed distinct clusters. Global genomes were also clustering by the specific country where the strain was isolated from. Genomes for strains isolated in the UK (prefix LH) and Switzerland (prefix BIB/BIC) clustered together. This clearly shows that infants from different geographic populations appear to be colonised by *Bifidobacterium* with distinct genomic characteristics.

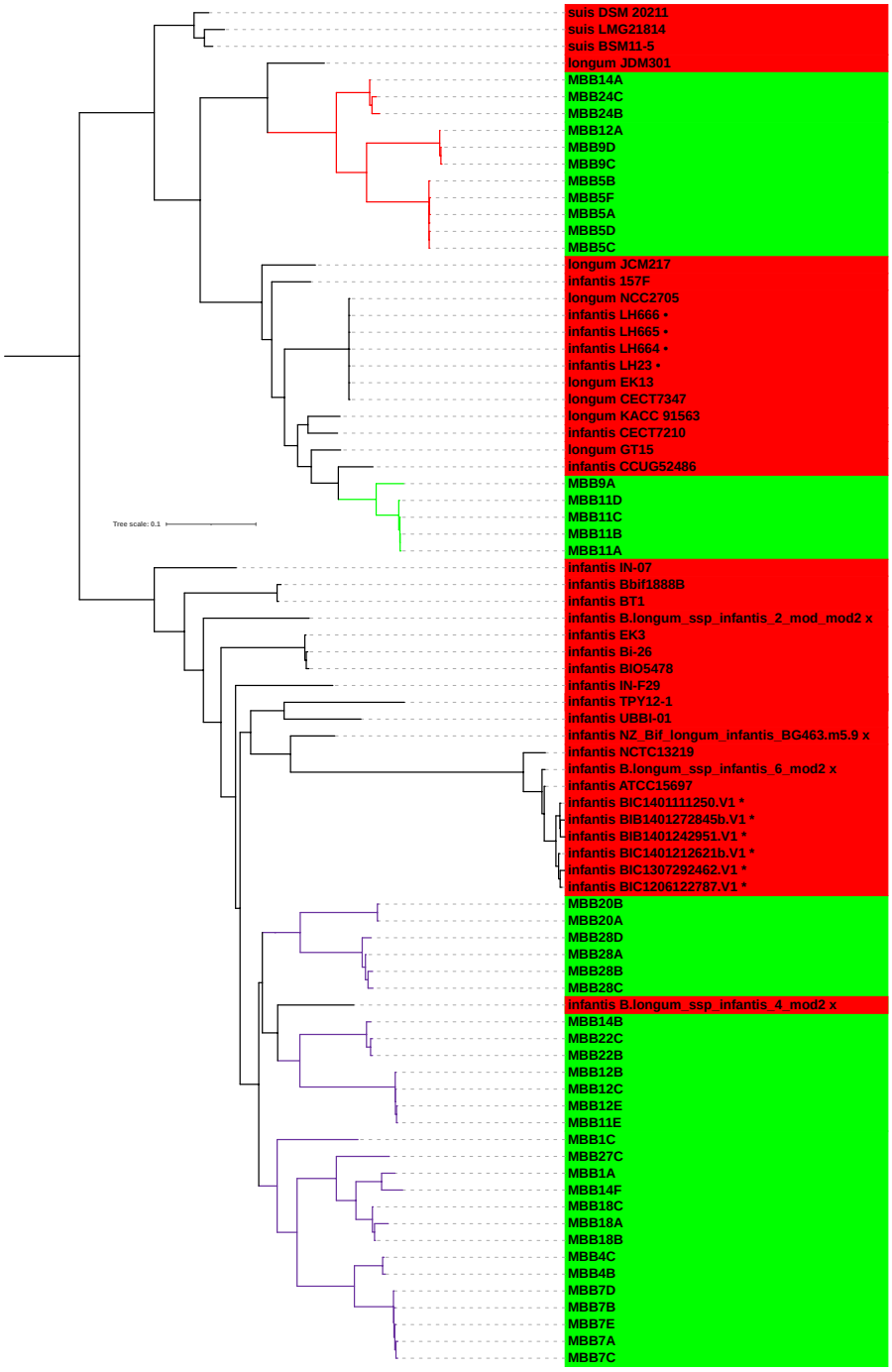

Figure 5.5:: Phylogenetic Distribution of global and Malawi Bifidobacterium longum genomes.

*Global Bifidobacterium longum genomes are shown in red, while Malawi Bifidobacterium longum genomes are shown in green. Malawi Bifidobacterium longum clusters A, B and C are shown using green, red and purple branches. Genomes from the UK, Switzerland and Bangladesh are marked with a •, * and x*

Some *Bifidobacterium* species have been shown to possess colonisation resistance properties against some bacterial pathogens (Thomson et al., 2018).

Bifidobacterium's ability to digest HMOs have been shown to contribute to this beneficial property. To determine the HMO digesting genes, present in the *Bifidobacterium* genomes from Malawian infants, the Malawi genomes were BLASTed against 90 HMO digesting genes provided by the Hall laboratory (L. Hall, Quadram Institute, personal communication). Significantly higher numbers of HMO digesting genes were identified in *B. longum* than, *B. breve, B. bifidum* and *B. pseudocatenulatum* with a median of 36, 25, 9 and 9 HMOs respectively (**Figure 5.6**). The median HMO digesting genes present in *B. longum*, *B. bifidum*, *B. breve* and *B. pseudocatenulatum* were 36, 9, 25 and 9, respectively (**Figure 5.7**).

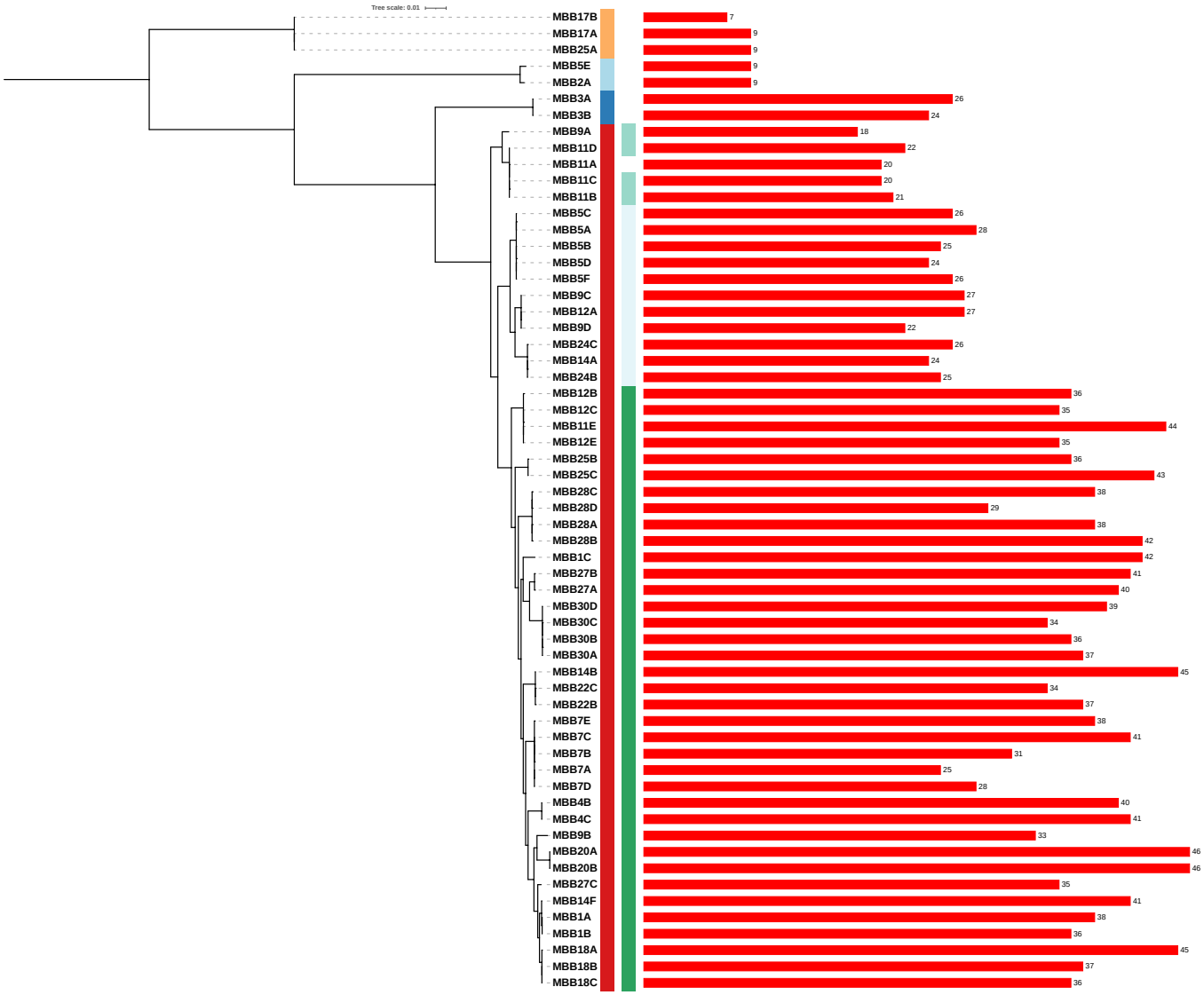

Figure 5.6: A dendrogram with the number of HMO genes present in Malawi Bifidobacterium genomes.

The first colour strip shows the Bifidobacterium species with red, blue, light blue and yellow standing for longum, breve, bifidum and speudocatenulatum. The second colour strip shows B. longum subsp with pale green, faint green and green showing Cluster A, B and C.

Table 5.3: Number of HMO metabolising genes present in Malawi Bifidobacterium genomes

Species/ subspecies (number of genomes)	**Number of HMO genes – median (% of identified against 90 reference HMO genes)**		
	Min (%)	**Median (%)**	**Max (%)**
B. bifidum (2)	9 (10)	9 (10)	9 (10)
B. breve (2)	24 (26.67))	25 (27)	26 (28.88)
B. speudocatenulatum (3)	0 (0)	9 (10)	9 (10)
B. longum (53)	18 (20)	36 (40)	46 (51.11)
B. longum cluster			
B. longum – cluster A (11)	22 (24.44)	26 (28.88)	28 (31.11)
B. longum – cluster B (5)	18 (20)	20 (22.22)	22 (24.44)
B. longum – cluster C (38)	*25 (27)*	*38 (42.22)*	*46 (51.11)*

Comparing the number of HMO digesting genes present in the Malawi *Bifidobacterium* by *B. longum* clusters demonstrated that Cluster C, which we confidently identified as containing *B. longum* subspecies *infantis* genomes, showed significant differences among the three clusters (Kruskal Wallis statistics 32.38, P-value = 0.0001). Group by group comparison using student t-test also demonstrated significant differences between the groups (**Figure 8.8**). The median HMO digesting genes present in *B. longum* Cluster A, B and C were 26, 20 and 38, respectively. The number of genomes for *B. bifidum, B. breve,* and *B. pseudocatenulatum* was too small to test for statistical significance.

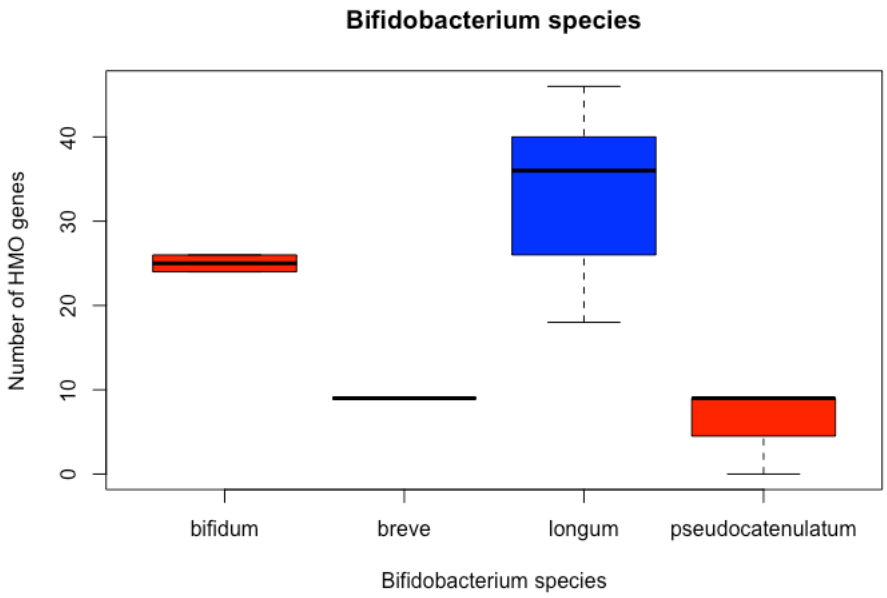

Figure 5.7: HMO digesting genes present among Bifidobacterium species isolated from exclusively breastfeeding Malawian infants.

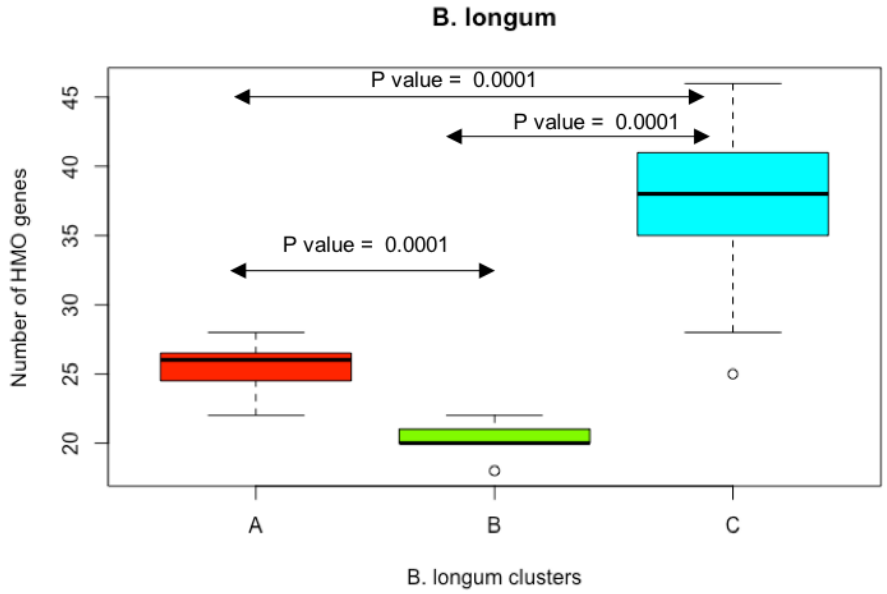

Figure 5.8: HMO digesting genes present in the Malawi B. longum clusters.

5.4 Discussion

Through this study, we provided insights on the genomic composition of *Bifidobacterium* isolated from Malawian infants by identifying the *Bifidobacterium* species and the HMO digesting genes in this population. By analyzing the gut microbiota profiles of Malawian children, the study has confirmed that *Bifidobacterium* is predominant in exclusively breastfeeding infants.

Four *Bifidobacterium* species; *B. bifidum, B. breve, B. longum,* and *B. pseudocatenulatum* were identified using whole genome analysis. About 90% of the isolated *Bifidobacterium* were identified as *B. longum* with three distinct clusters A, B and C. Most of the B. longum genomes were Cluster C. *B. longum* has four subspecies; *longum*, *infantis*, *suis* and *suillum* (Paola Mattarelli & Biavati, 2018; Tarracchini et al., 2021). Cluster C was specifically identified as belonging to the subspecies *infants (B. longum* subspecies *infantis)* because it clustered well with global *B. longum* subsp *infantis* genomes. *B. longum* subspecies *infantis* was first isolated from infant stool and is known to thrive on complex sugars found in breastmilk (P. Mattarelli, Bonaparte, Pot, & Biavati, 2008; Tarracchini et al., 2021). Some of the *Bifidobacterium* genomes in cluster A were identified as belonging to *B. longum* subspecies *suillum* when Sanger sequenced reads were blasted against the NCBI database. Kraken, however, failed to identify any *suillum* subspecies because the database did not have a *suillum* reference genome. This is the first isolation of B. suillum from Malawian infants. This is a novel identification and there is currently no reference strains. Ideally, further validation would be done using additional strains from global collections, and, this highlights the need for large scale studies of the

microbiomes of understudied populations and the role of microbiome in health and disease.

The most prevalent *Bifidobacterium*, *B. longum* subspecies *infantis* (*B. longum* Cluster C) had the highest number of HMO digesting genes. HMOs are not digested by human digestive enzymes but by gastrointestinal bacteria belonging to the genus *Bifidobacterium*. Some *Bifidobacterium* species, such as *B. longum* subspecies *infantis*, are better at digesting HMOs than others (Bode, 2012; Thomson et al., 2018). The significant differences observed in the number of HMO digesting genes amongst the different species and subspecies support this. As previously reported that, the *B. longum* subspecies *infantis* genome has genes specifically adapted to the digestion of the complex sugars in breastmilk; I found that Cluster C, which was identified as a cluster composed of only *B. longum* subspecies infants, had the highest number of HMOs. Another study that genomically compared the ability of *B. longum* subspecies longum and *B. longum* subspecies *infantis* in digesting HMOs also reported that subspecies *infantis* possessed more of the HMO digesting genes than subspecies *longum*. (Li et al., 2021). *B. longum* subspecies *longum* has more plant carbohydrate digesting genes which may be important through the weaning process (Kujawska et al., 2020). Human milk oligosaccharides have specifically been shown to play an important role in colonisation resistance by, among others, preventing pathogenic organisms from attaching to the gut epithelial cells through blocking the surface sugar on the epithelial cells and promoting gut microbiota diversity (Thomson et al., 2018).

The relative abundance of *Bifidobacterium* for this very young and exclusively breastfeeding infant cohort, was high in most of the participants, with 73.28% being the highest. High *Bifidobacterium* relative abundance of up to 90% has been reported (F. Turroni et al., 2018; Yatsunenko et al., 2012). A study describing the gut microbiota of Finnish and Malawian infants aged six months reported a relative *Bifidobacterium* abundance of 70.8% in the Malawian cohort. This was significantly different from the Finnish cohort, with a relative abundance of 46.8% (Grześkowiak et al., 2012). The difference in the relative abundance of *Bifidobacterium* in Finnish and Malawian children was explained by early feeding habits. Most Malawian children were breastfed while Finish children were formula fed. Breastmilk promotes the proliferation of *Bifidobacterium* (Murphy et al., 2017; D. Taft et al., 2019; F. Turroni et al., 2021). Breastmilk contains complex sugars that *Bifidobacterium* thrives on. These complex sugars are not found in formula milk; hence the relatively lower *Bifidobacterium* abundance reported in children fed on formula milk (Bode, 2012; Yasmin et al., 2017). Other infant gut microbiota studies conducted in Malawi have also reported a high abundance of *Bifidobacterium*, although most studies focused on the infant from six months old (Cheung et al., 2016; Kamng'ona et al., 2020; Kortekangas et al., 2020). In Mangochi, a rural district in Malawi, a study reported *Bifidobacterium* relative abundance of 54.6% and 25.3% at 6 and 18 months, respectively (Cheung et al., 2016). The relative abundance of *Bifidobacterium* in highest in exclusively breastfed children and wanes overtime. This explains the high relative abundance of *Bifidobacterium* reported for this cohort and the low relative abundance reported for children in Mangochi at 6 and 18 months in the study done in Mangochi.

There were no significant differences in alpha and beta diversities computed based on sex, *Bifidobacterium* culture results (whether positive or negative), and identity of *Bifidobacterium species*. The fact that this cohort comprised of infants within a very narrow age bracket, all were exclusively breastfeeding and only one was born through caesarean section may explain the lack of significant differences in alpha and beta diversity. The fact that 90% of the isolated *Bifidobacterium* belonged to the species *B. longum* could explain the lack of statistical significance in alpha and beta diversities based on *Bifidobacterium* culture results. It is unlikely that *Bifidobacterium* belonging to the same species would differentially impact the overall gut microbiota composition since they are likely to genomically share a lot in common and phenotypically behave similarly in the intestinal environment.

Except for 4 participants, most participants who had multiple *Bifidobacterium* isolate, had the same *Bifidobacterium* species isolated. This could mean that Malawian infants are colonized by one *Bifidobacterium* species at a time. This might not be the case considering the challenges faced in the initial attempt to culture *Bifidobacterium,* which led to using frozen samples for isolation. Infant stool has been shown to harbour multiple *Bifidobacterium* species/ variants (Francesca Turroni et al., 2012). To assess the actual composition and diversity of *Bifidobacterium* in the gut microbiota of Malawian infants, advanced molecular techniques such as shotgun metagenomics should be used instead of culture-based methods or 16S rRNA sequencing. Studies utilizing advanced molecular methods have reported that infants harbour multiple *Bifidobacterium* strains (Francesca Turroni et al., 2012). It is, therefore, possible that infants in this cohort also harboured multiple *Bifidobacterium* strains. Shotgun metagenomic studies would help provide in-depth knowledge of

Bifidobacterium composition, diversity and function in the Malawian population (Milani et al., 2020; Ranjan et al., 2016; Ventura, Milani, Turroni, & van Sinderen, 2021).

Incomplete annotation of *Bifidobacterium* genomes in the databases resulted in Cluster A and Cluster B genomes clustering together with a mixture of *B. longum* subsp *longum*, *B. longum* subsp *infantis* and *B. longum* subsp *suis* genomes. The identity of Cluster A and Cluster B genomes was therefore inconclusive. The NCBI and Kraken databases both did not have a reference genome for *B. longum* subspecies *suillum*. There is need for proper annotation of the available *Bifidobacterium* genomes and identification of reference-type strains for the different *Bifidobacterium* species and subspecies to better understand the population of *Bifidobacterium*.

5.4.1 Strengths and weaknesses

This is the first study to isolate and genomically characterise *Bifidobacterium* from Malawian infants. I have genomically shown that the Malawi *Bifidobacterium* is different from the global *Bifidobacterium*. Different species and subspecies have different numbers of HMO digesting genes, with *B. longum* subspecies *infantis* having more of these genes. In this study, I also isolated and identified a unique *Bifidobacterium* identified as *B. longum* subspecies *suillum* and as part of *B. longum* Cluster 3. The study's primary limitation is that it only recruited a small number of participants from a very narrow age group. In addition, isolation of *Bifidobacterium* from frozen samples may have affected the *Bifidobacterium* population that were

isolated. Using 16S rRNA sequencing data to profile the gut microbiota also limited analysis of the gut microbiota data to genus and not species level.

5.4.2 Future direction

Specific *Bifidobacterium* species or populations with the right HMO digesting genes could be used as therapeutics against important conditions. In the gut, *Bifidobacterium* abundance and diversity change with age and diet. A longitudinal or cross-sectional study recruiting participants with a broad age group is needed to further understand *Bifidobacterium* abundance, diversity, and genomic properties in the first year of life, and the potential protective effect of *Bifidobacterium* colonisation in pathogen colonisation resistance. To understand the effect of diet on genomic and phenotypic properties of *Bifidobacterium*, studies recruiting both mothers and infants should be conducted where breastmilk and mothers gut microbiota composition could be correlated to the infant gut microbiota and genomic and phenotypic properties of *Bifidobacterium*. Advanced molecular methods such as shotgun metagenomic sequencing should be used to generate an in-depth understanding of *Bifidobacterium* communities present in Malawian children. Genotypic and phenotypic studies need to be done to understand further the mechanisms by which this *Bifidobacterium* and breastmilk metabolism helps control infection. The next chapter describes an initial exploration of the possibility that Malawian bifidobacterial might have the functional capacity to inhibit the growth of an important Malawian pathogen, *S.* Typhimurium.

5.5 Conclusion

This study has shown that Malawian exclusively breastfeeding children are colonized by different *Bifidobacterium* species and subspecies but *B. longum* subspecies *infantis* is predominant. *Bifidobacterium* genomes from Malawi have also been shown to be genomically different from global genomes. The study has also reported inter-species and inter-subspecies variation in the HMO digesting genes in the Malawi *Bifidobacterium* genomes with *B. longum* subspecies *infantis* having the highest number of HMO digesting genes.

Chapter 6: Exploiting Anti-infection Properties of the Early Life Microbiota Member Bifidobacterium Against Invasive Salmonella Typhimurium – An Exploratory Study

6.1 Introduction

Salmonella Typhimurium is the leading cause of iNTS infections in SSA. In Malawi, 35.9 % of 29,183 blood cultures done at MLW from 1998 to 2016 were positive for *S.* Typhimurium (Musicha et al., 2017). Another study in Malawi reported that about 50% of Malawian children had eNTS in a period of one year (Tonney Stophen Nyirenda, 2015). Most of the iNTS and eNTS is attributed to *S.* Typhimurium. Similar findings are reported in **Chapter 3,** where 46% of children aged 6 to 18 months had culture or qPCR positive eNTS, most of which were also attributed to *S.* Typhimurium. As highlighted in the **Introduction chapter section 5.1.3**, eNTS exposure is potentially harmful as it can cause iNTS. The prevalence of iNTS starts to rise once supplementary feeding is introduced and in Malawi, it peaks at 14 months (Tonney S. Nyirenda et al., 2014). The low prevalence of iNTS in early life may be attributed to different factors including protection from maternal serum antibody transferred transplacentally and gut microbiota associated factors.

The composition of the gut microbiota; organisms present, diversity, or richness, determine the establishment or elimination of invading pathogenic organisms. Using mice, Claudia Eberl et. al showed that *E. coli* Mt1B1 provides colonisation resistance to *S.* Typhimurium by limiting carbon sources. This was, however, only possible when E. coli was part of a microbial community and not in isolation (Eberl et al., 2021). A study aimed at identifying organisms with inhibitory properties against *S.*

Typhimurium by testing supernatants from 973 bacterial isolates identified *Mitsuokella, Escherichia/Shigella, Anaerovibrio, Selenomonas,* and *Streptococcus* as having inhibitory properties against *S.* Typhimurium (U. Y. Levine, Bearson, & Stanton, 2012).

Bifidobacterium, a member of the gut microbiota is important for colonisation resistance. *Bifidobacterium* has been shown to have immunomodulatory impacts and has also been shown to possess antibacterial and antiviral properties (Henrick et al., 2021; Lim & Shin, 2020). A reduced *Bifidobacterium* relative abundance is associated with an increase in iNTS prevalence (Selma-Royo et al., 2019). One known key to *Bifidobacterium* mediated colonisation resistance is *Bifidobacterium's* ability to digest HMO (Lawson et al., 2020; Thomson et al., 2018). Digestion of HMO by *Bifidobacterium* produces short-chain fatty acids that are important for colonisation resistance (Bode, 2012). In **Chapter 5**, I report that *Bifidobacterium* isolated from Malawian children is distinct from *Bifidobacterium* isolated from other parts of the world and that they have a wide range of HMO digesting genes. *Bifidobacterium* species can also inhibit pathogens in a number of ways: reducing the pH of the local environment, competing for nutrients, and via the production of highly specific anti-microbials, i.e. bacteriocins and bacteriocin-like inhibitory compounds (Alessandri et al., 2019; Lim & Shin, 2020). *Bifidobacterium* has been shown to possess pro-and anti-inflammatory cytokines that may help control Salmonella infection (Lim & Shin, 2020).

Bifidobacterium, a known health promoting gut microbiota member, is predominant in exclusively breastfeeding infants. In **Chapter 5**, I have shown that exclusively

breastfed Malawian children have a high level of bifidobacteria, with *Bifidobacterium longum* (*B. longum*) being the predominant isolate. It is, however, not known if the *Bifidobacterium* from Malawi possesses anti-infective properties against iNTS, specifically the ST313 *S.* Typhimurium. This chapter reports exploratory experiments to understand whether *Bifidobacterium* isolated from Malawian breastfeeding infants might have inhibitory properties against the growth of ST313 invasive *S. Typhimurium* strain (D23580), by using invitro competition assays.

6.1.1 Hypothesis

Some *Bifidobacterium* strains found in Malawian children have anti-infective activities that limit the *in vitro* growth of invasive *S.* Typhimurium ST313.

6.1.2 Overall objective

- To explore the anti-infective activities of *Bifidobacterium* strains isolated from Malawian children that may limit the *in vitro* growth of invasive *Salmonella* Typhimurium ST313.

6.1.3 Specific objectives

1. To characterize the growth of invasive *Salmonella* Typhimurium and *Bifidobacterium* in liquid culture media
2. To assess the growth of invasive *Salmonella* Typhimurium ST313 when co-cultured with *Bifidobacterium*
3. To assess the growth of invasive *Salmonella* Typhimurium ST313 in *Bifidobacterium* supernatant

6.2 Methods

6.2.1 Bacterial strains used

Bifidobacterium isolates from Malawian exclusively breastfeeding infants - reported in Chapter 5, and an invasive ST313 *S.* Typhimurium isolate (Strain D23580) from a Malawian child were used. **Table 6.1** shows all isolates used during the exploratory experiments.

Table 6.1: Bifidobacterium isolates used for different exploratory experiment

Experiment	*Bifidobacterium* isolates used	Species/ Cluster	Measurement
***Bifidobacterium* growth in BHI and CDRIM media**	All *Bifidobacterium* isolates were tested (n = 60)	*B. longum* *B. breve* *B. bifidum* *B. speudocatenulatum*	OD values
***Bifidobacterium* and *S.* Typhimurium co-culture**	Two isolates were tested MBB11C MBB30D	 *B. longum* Cluster A *B. longum* Cluster C	CFU/ml
***S.* Typhimurium growth in *Bifidobacterium* supernatant**	5 isolates were tested	*B. longum* Cluster A *B. longum* Cluster C *B. breve* *B. bifidum* *B. speudocatenulatum*	OD values

6.2.2 Culture media for *Bifidobacterium* and *S.* Typhimurium growth

All *Bifidobacterium* isolates and *S.* Typhimurium were cultured in BHI and a chemically defined media (CDRIM) liquid media. For experiments requiring solid media, *Bifidobacterium* and *S.* Typhimurium were cultured on MRS and LB agar respectively. *Bifidobacterium* was cultured anaerobically using Ruskinn Concept Plus anaerobic chamber (5% carbon dioxide gas, 10% hydrogen gas and 10% nitrogen

gas) while *S.* Typhimurium was cultured both aerobically and anaerobically. Details on this experiment are described in **Chapter 2 section 2.7.1**

6.2.3 Co-culturing of Bifidobacterium and S. Typhimurium

Two *Bifidobacterium longum* isolates, one each from Cluster A (MBB11C), and Cluster C (MBB30D) were individually co-cultured anaerobically with D25380 using CDRIM liquid media for 24 hours, washed and co-cultured anaerobically. One ml of the culture suspension was taken at 0, 3, 6, 9, 12, 15, 18, 21, and 24 hours. Following Miles and Misra technique, *Bifidobacterium* and *S.* Typhimurium were sub-cultured onto MRS and LB agar media, respectively. Details of the co-culture experiment are in **Chapter 2 section 2.7.2.**

6.2.4 Growth of S. Typhimurium in Bifidobacterium cell-free supernatant

Using a single colony, representatives of *B. breve*, *B. speudocatenulatum, B. longum* Cluster A and Cluster C were anaerobically cultured in CDRIM liquid media for 24 hours. *Bifidobacterium* supernatant was obtained by filtering the *Bifidobacterium* culture using a Minisart® CA 0.45-micron non-pyrogenic filter. *S.* Typhimurium was cultured aerobically in a 24 flat bottom well plates with a loose, lid at 37°C for 12 hours in *Bifidobacterium* supernatant diluted in CDRIM: 1:1, 1:2, 1:4, 1:8, 1:16 and 1:32. Details on this experiment are described in **Chapter 2 section 2.7.3**.

6.3 Results

6.3.1 Growth patterns of *Bifidobacterium* and *Salmonella* in a nutrient-rich media and chemically defined media

To identify the type of media that would support the growth of both *Bifidobacterium* and *Salmonella,* all samples were cultured in a nutrient-rich media - Brain heart infusion (BHI), and a chemically defined media (CDRIM) and OD values were recorded every 15 minutes for 24 hours. BHI and CDRIM both supported the growth of all *Bifidobacterium* species but demonstrated different levels (**Figure 6.1**) and patterns of growth (**Figure 6.2**) in the two culture media. The OD values for *Bifidobacterium* strains cultured in CDRIM were significantly higher when compared to that of strains cultured in BHI. The median OD value for *Bifidobacterium* in CDRIM media at 10 hours and 30 hours was 0.16 (CI of 0.14 to 0.17) and 0.31 (CI of 0.27 to 0.33) respectively. The median OD value for *Bifidobacterium* cultured in BHI media at 10 and 30 hours was 0.11 (CI of 0.08 to 0.14) and 0.18 (CI of 0.16 – 0.21) respectively.

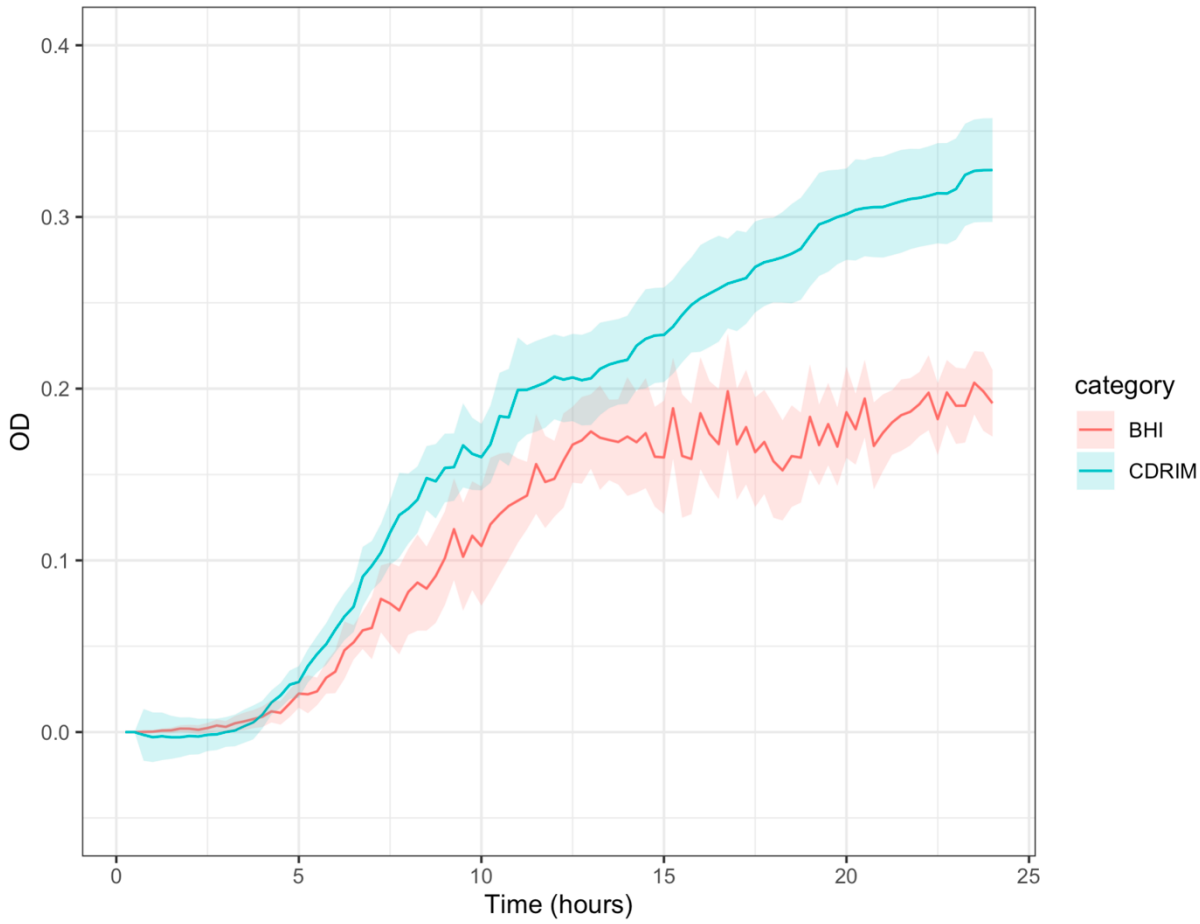

Figure 6.1: Anaerobic growth of Bifidobacterium in BHI and CDRIM culture media at different time points.

OD values for all Bifidobacterium strains cultured in CDRIM (green) and BHI (red) are shown at different time points. The shaded area indicates confidence intervals for triplicate readings from the same culture media and conditions.

The growth patterns of the isolates were also different when cultured in the two growth media. Assessing individual growth curves showed that *Bifidobacterium* strains cultured using CDRIM had typical sigmoid growth curves, while most of the *Bifidobacterium* strains cultured using BHI had atypical growth curves (**Figure 6.2**). Using CDRIM, a media that demonstrated a typical sigmoid growth curve, I observed that some *Bifidobacterium* strains reached the exponential growth phase faster than others. Most of the *Bifidobacterium* isolates reached the exponential phase anywhere between 4 to 14 hours from the time of culturing with a median time of 8

hours. Most *Bifidobacterium* isolates cultured in BHI were characterised by low OD values.

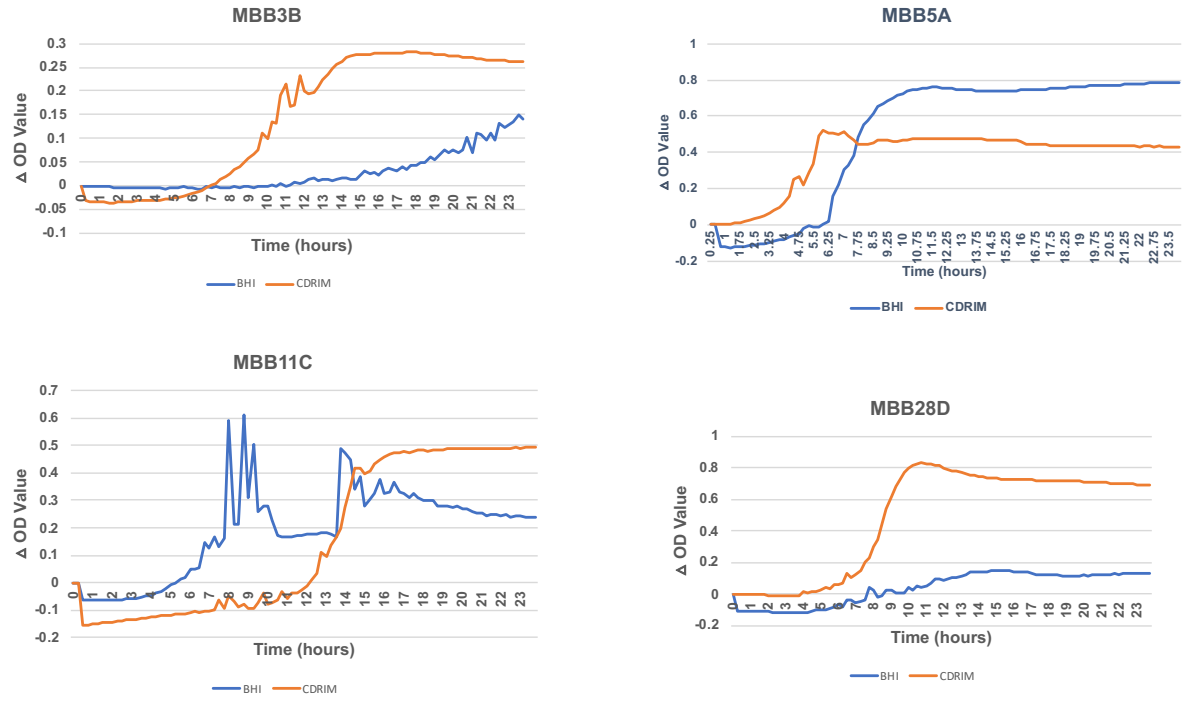

Figure 6.2: Growth patterns of Bifidobacterium in BHI and CDRIM culture media.

Individual growth curves of some of the Bifidobacterium isolates cultured in BHI and CDRIM were plotted over 24 hours incubation time. B. breve (MBB3B), B. longum Cluster B (MBB5A) , B. longum Cluster A (MBB11C), B. longum Cluster B (MBB5A) and B longum Cluster C (MBB28D)

Both BHI and CDRIM supported the growth of *Salmonella* in both aerobic and anaerobic conditions (**Figure 6.3**). There were no statistically significant differences in *S.* Typhimurium growth in aerobic conditions versus anaerobic conditions. However, *Salmonella* growth in anaerobic condition seem to be slower in reaching exponential growth and the stationary phase was lower than when it is cultured in aerobic condition.

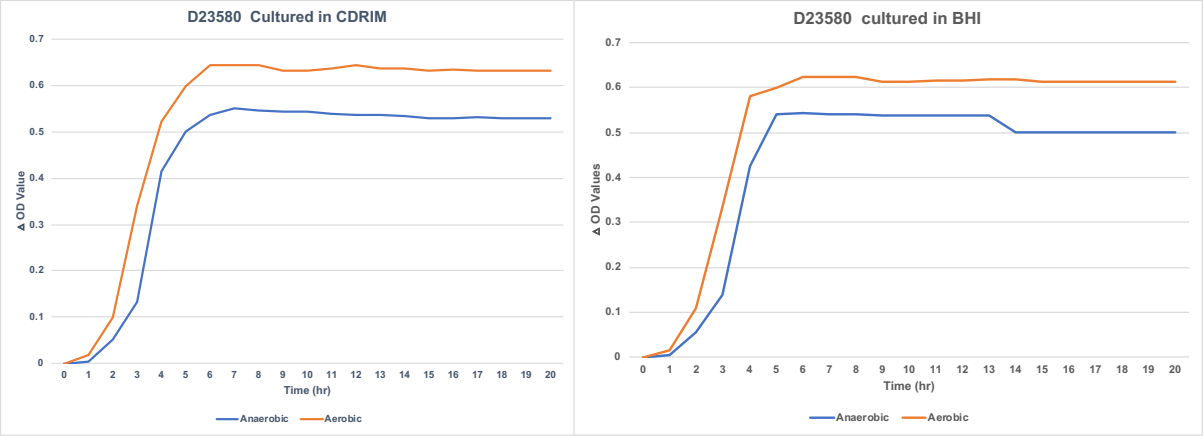

Figure 6.3: Growth of S. Typhimurium in CDRIM and BHI using both aerobic and anaerobic conditions.

6.3.2 The growth of *Salmonella* Typhimurium when co-cultured with *Bifidobacterium*

The growth patterns of *S.* Typhimurium D23580 in the presence of *Bifidobacterium* were then determined by co-culturing *S.* Typhimurium with *Bifidobacterium*. This experiment used one *B. longum* isolates from Cluster A (MBB11C) and Cluster C (MBB30D). Significant differences were observed between *S.* Typhimurium co-cultured with *B. longum* from cluster A (MBB11C) and the control. In the control experiment, *S.* Typhimurium reached the stationary phase at 3 hours while *Salmonella* co-cultured with *Bifidobacterium* from both clusters reached the stationary phase at 6 hours (**Figure 6.4**). A decline phase was observed after 9 hours for *Salmonella* co-cultured with MBB11C.

However, there were no statistically differences in the growth of *S.* Typhimurium when co-cultured with *B. longum* from Cluster C (MBB30D) compared to the control experiment after reaching the stationary phase.

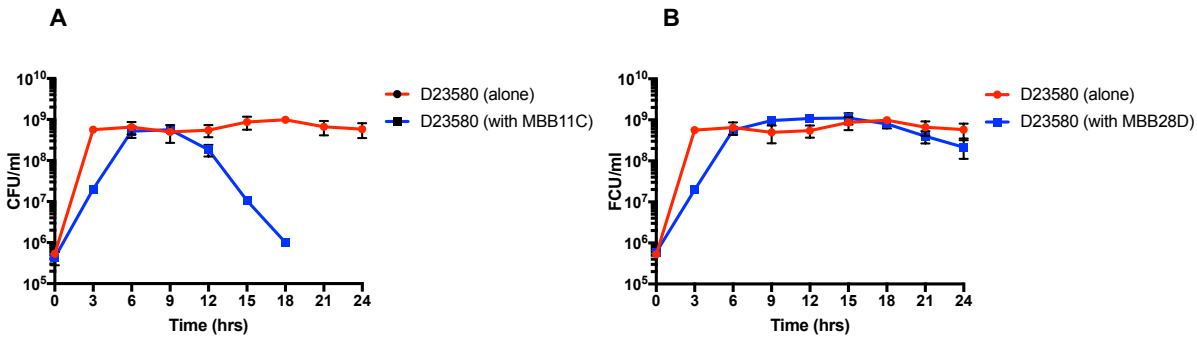

Figure 6.4: S. Typhimurium growth when co-cultured with Bifidobacterium

MBB11C - B. longum from Cluster A (A) and MBB28D – B. longum from Cluster C (B). Blue is for Salmonella growth in the presence of Bifidobacterium while red is S. Typhimurium alone.

The pH of the culture media was monitored throughout the culturing period. There was a decrease in the pH levels when D23580 was co-cultured with either MBB11C or MBB28D (**Figure 6.5).**

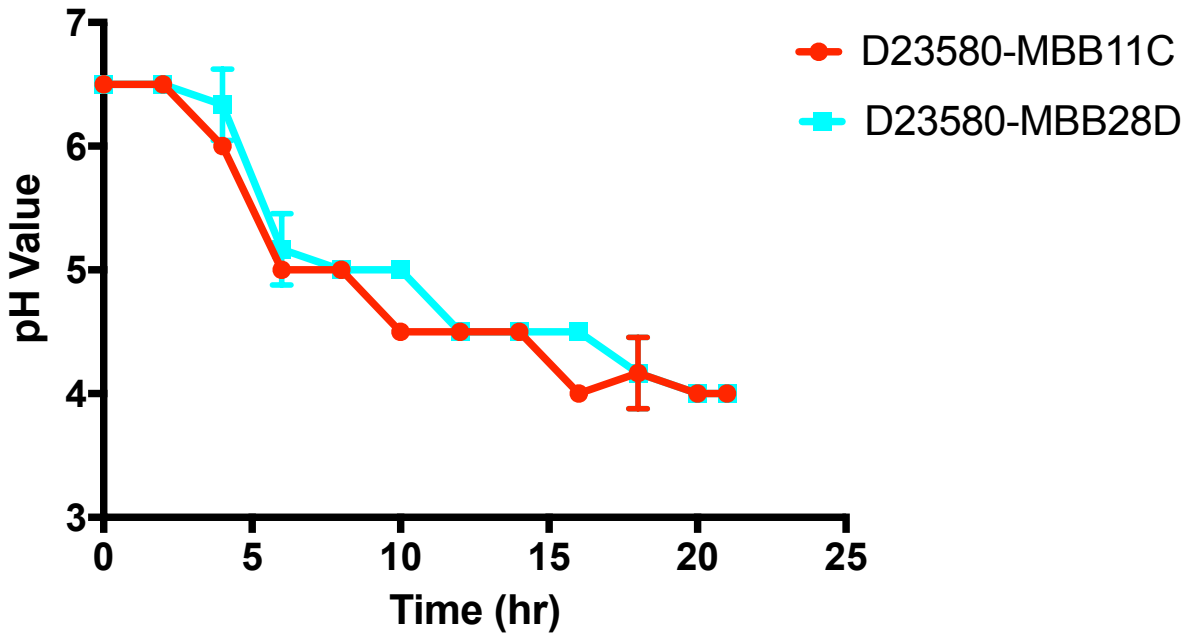

Figure 6.5: Impact of pH on Salmonella growth when S. Typhimurium and Bifidobacterium are co-cultured.

pH was adjusted to 6.5 at the beginning of the experiment. Any changes in pH value were monitored over the 24 hours culture period.

6.3.3 *S.* Typhimurium growth in *Bifidobacterium* supernatant

A dose-response effect was observed in all strains **(Figure 6.6).** Minimal dilution of 1:2 was however enough to completely abolished the effect on the growth of *S.* Typhimurium for most of the strains. The degree of inhibition was particularly significant for Cluster A where a dilute of 1:4 was shown to abolish the effect. Undiluted *Bifidobacterium* supernatant (1:1) for isolates from cluster A inhibited *S.* Typhimurium growth for the entire 24 hours incubation period. Diluted supernatant seemed to suppress growth at 1:2 dilution as observed by the low OD value at stationary phase. Higher dilutions (1:4, 1:8, 1:16, 1:32) did not inhibit the growth of *S.* Typhimurium. Undiluted *Bifidobacterium* supernatant from *B. longum* cluster C, *B. breve* and *B. pseudocatenulatum* did not inhibit *Salmonella* growth but seemed to slow down and suppress the growth of *S.* Typhimurium as observed by the time it

took to reach exponential phase and the OD value stationary phase. All diluted supernatants did not display any inhibitory properties, apart from a 1:2 dilution of *B. longum* cluster A.

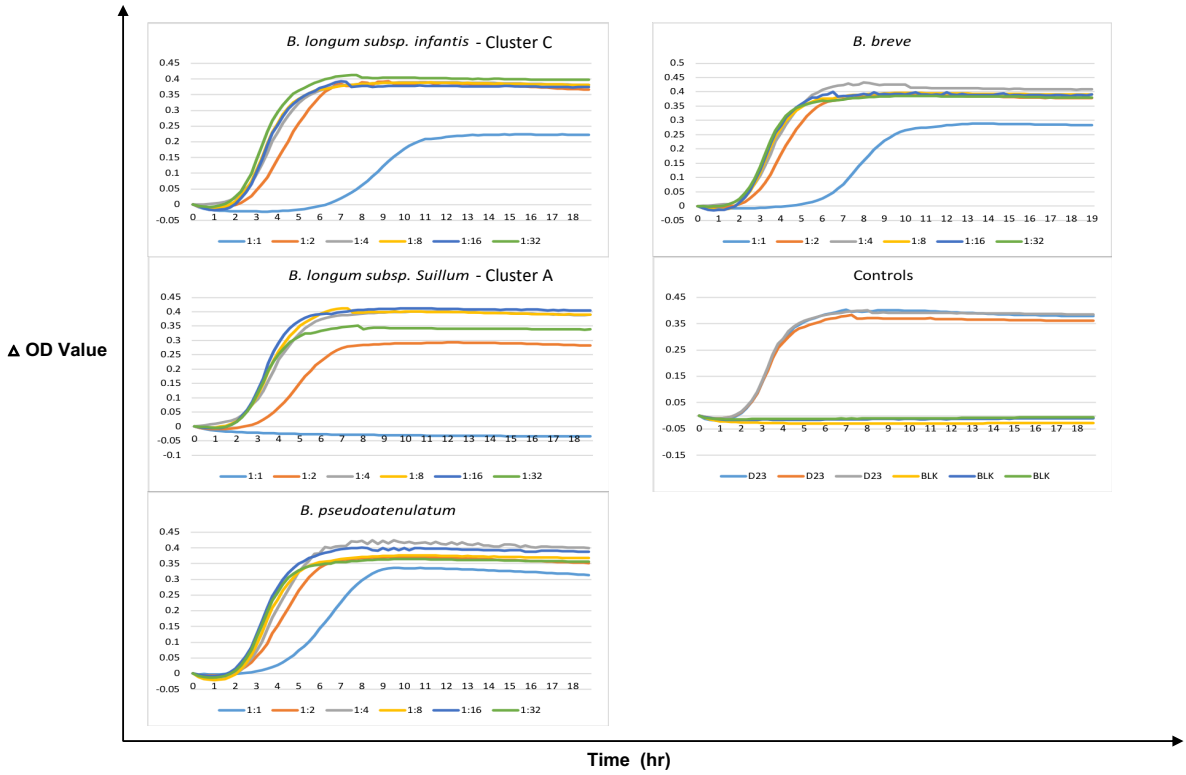

Figure 6.6: S. Typhimurium, D23580 growth in 1 in 2 serially diluted Bifidobacterium supernatant.

Supernatant for representative isolates for B. breve, B. pseudocatenulatum and B. longum Cluster A and C were used.

6.4 Discussion

This was an exploratory study that aimed to determine the anti-infective properties of *Bifidobacterium* isolated from Malawian infants against an invasive *S.* Typhimurium strain. The results have shown that some *Bifidobacterium* species isolated from Malawian infants possess properties that limit the growth of invasive *S.* Typhimurium, D23580. This effect was observed when *S.* Typhimurium, D23580, was co-cultured with *B. longum* from Cluster A or when *S.* Typhimurium was cultured in conditioned *Bifidobacterium* supernatant from Cluster A.

In the *Bifidobacterium* and *S.* Typhimurium co-culture experiment, *Salmonella* took longer to reach stationary phase. The *Bifidobacterium* strain belonging to *B. longum* Cluster A (MBB11C) demonstrated inhibitory properties against the growth of *S.* Typhimurium while the other strain belonging to *B. longum* Cluster C did not show any inhibitory properties after reaching the stationary phase. One of the direct gut microbiota colonisation resistant mechanisms involve direct gut microbiota and pathogen interactions, production of soluble substances like bacteriocins, competition for nutrition and acting as a barrier of the intestinal mucosa barrier (Ducarmon et al., 2019). The inhibition in *S.* Typhimurium growth by MBB11C suggest a direct colonisation resistant resistance mechanism that may involve direct *Bifidobacterium* and *S.* Typhimurium interaction.

A reduction in pH was observed in this study. One of the mechanisms that bifidobacteria uses to controls pathogens is by the production of acids which reduces the environmental pH thus making the environment not conducive for the survival of other organisms. This shows that the *Bifidobacterium* used in this experiment may

have produced some compounds that were acidic and thus reduced the pH. The production of SCFAs such propionate, butyrate and acetate, following *Bifidobacterium* – carbohydrate metabolism, is associated with colonisation resistance against *Salmonella* (Lawson et al., 2020; Tsugawa et al., 2020). Another study aimed at understanding gut microbiota colonisation resistance mechanisms against *S.* Typhimurium reported that a *Bacteroides* species, a bacterium producing short-chain fatty acid, propionate, mediated colonisation resistance against *S.* Typhimurium (Jacobson et al., 2018). In Jacobson et al. study, the disruption of intracellular and not extracellular pH is what mediated the inhibition of *S.* Typhimurium (Jacobson et al., 2018). This study measured extracellular pH and not intracellular pH. The inhibitory property observed in the *Bifidobacterium* and *S.* Typhimurium co-culture experiment can, therefore, be partly explained by the reduction in extracellular pH. Understanding the type and amount of acid produced may help unravel pH mediated colonisation resistance. The impact of acid production in the intracellular environment was beyond the scope of this study but it should be studied for a better understanding of pH mediated colonisation resistance. Besides the production of acids, the production of other secretary compounds may also be involved.

It is possible that the inhibition of growth occurred because of secretion of an inhibitory substance, or because of consumption of a scarce nutrient by bifidobacteria. In order to determine if there are secretary compounds involved, an experiment where the growth of S. Typhimurium in concentrated and diluted *Bifidobacterium* supernatant was conducted. Dilution might be expected to reverse inhibition if a secretory product is responsible for growth-inhibition, However, if

growth is inhibited by scarcity of a nutrient, dilution would be expected to increase the decree of inhibition. In this experiment, some dose-response effects were observed for all *Bifidobacterium* strains, and those belonging to *B. longum* Cluster A required the greatest supernatant dilution to reverse inhibition, suggesting the highest concentration of an inhibiting factor from these cluster A. The limitation in the growth of *S.* Typhimurium in *Bifidobacterium* supernatant, which is reversed by dilution suggests the production of soluble metabolites that have inhibitory properties against *S.* Typhimurium, rather than the depletion of nutrients that support the growth of *S.* Typhimurium. Otherwise, serial dilution would have made the inhibition greater. Bondue et al. support that soluble metabolites produced by some *Bifidobacterium* species inhibit *S.* Typhimurium growth by suppressing *S.* Typhimurium virulence genes (Bondue et al., 2016). The expression of *S.* Typhimurium and *Escherichia coli* (*E. coli)* virulence genes *ler* and *hilA*, were reduced when *S.* Typhimurium and *E. coli* were cultured in cell-free supernatant obtained from *B. bifidum* and *Bifidobacterium crudilactis (Bondue et al., 2016)*. In another study where *Bifidobacterium* cell-free supernatant was used to culture bacterial pathogens, undiluted supernatant for *Bifidobacterium* strains used was able to inhibit *Salmonella* growth.

Bacterial growth is influenced by the amount of moisture, pH, amount of Oxygen and source of food/media. The interaction between *Bifidobacterium* and *Salmonella* might be affected by these different growth conditions. The anaerobic conditions used in this study supported the growth of both *Bifidobacterium* and *S.* Typhimurium. *Salmonella* is a facultative organism and can therefore grow in both aerobic and anaerobic conditions. In anaerobic conditions, *Salmonella* uses fermentation for

energy metabolism. Aerobic metabolism is however ideal for *Salmonella* as it generates more ATP than fermentation. It is, therefore, not surprising that there was no significant difference in the growth of *S.* Typhimurium in aerobic and anaerobic conditions. *Salmonella* can survive on different carbon sources, including those from carbohydrates such as glucose and lipids (Diacovich, Lorenzi, Tomassetti, Méresse, & Gramajo, 2017). *S.* Typhimurium was able to grow on both BHI and CDRIM. BHI is a rich media that uses calf brain and beef heart infusions as carbon sources, while the CDRIM media used contained 20% glucose as a carbons source (aldrich, 2021). The CDRIM media also had selective supplements such as magnesium sulphate ($MgSO_4$), calcium chloride ($CaCl_2$), yeast extract, polypeptone and M9 to support *Bifidobacterium* growth.

The greatest inhibition has been observed from a strain that seem to be novel. Blasting the Sanger 16S rRNA sequence read identified this strain as being a *B. longum* subspecies *suillum*. *B. longum* subspecies *suillum* has previously only been isolated from swine. As indicated in Chapter 5, an in-depth genomic characterisation of this strains is needed in order to better understand its phenotypic properties. Although this strain was nor highly abundant in this study cohort, it is possible that this strain is favoured in the microbiota of Malawian kids. Since iNTS is an important illness Malawi, the gut microbiota of Malawian kids might have evolved in a way that allow colonisation of unique *Bifidobacterium* stains that would give better protection. Microbiome and *Bifidobacterium* characterisation studies using highly sensitive methods should be conducted to have an in-depth understanding of the organisms that colonise the gut of Malawian children, specifically health promoting gut microbiota members such as *Bifidobacterium*.

6.4.1 Strengths and limitations

The study's main strength is that it used locally-relevant *Bifidobacterium*, isolated from very young Malawian children, and *S.* Typhimurium ST313, a leading cause of bloodstream infection in Malawi and other SSA countries. Both organisms used in this study were isolated from Malawian children and are thus known to be locally relevant.

The main weakness of the study is that only preliminary and exploratory work is being presented. Culture and co-culture methods could be refined, individual experiments could be replicated on more than one occasion, and a greater range of isolates representing the different clades of local and global bifidobacteria could be tested. Inhibitory activity against a wider range of local and global *Salmonella* strains would also be interesting to explore. If these exploratory findings were upheld, then further work to understand the molecular mechanisms underlying this phenotypic phenomenon would be very valuable and might have important translational implications.

Although anaerobic cabinets such as the Ruskinn is the standard equipment used to culture *Bifidobacterium* and other anaerobic organisms, not all laboratories have them. Anaerobic cabinets are expensive, so are the gases used. The alternative to using anaerobic cabinets for culturing anaerobic microorganisms is the use of anaerobic jars and sachets. These are cheaper than anaerobic cabinets and can easily be used in laboratories that have ordinary incubation cabinets. The use of anaerobic jars in these experiments was however not feasible since the experiments

required multiple sampling during the incubation period. This would then introduce oxygen in the anaerobic jars. Anaerobic cabinet was therefore purchased to be used for repeat and follow-up studies.

There was therefore an intention to transfer these exploratory experiments, which were mostly done in the UK, to the laboratory in Malawi, in order to expand local capacity by establishing anaerobic culture methods and validate these findings more comprehensively. Brief descriptions of some of the experimental work that I had hoped to do, to improve the previous experiments and one further experiment to answer some of the emerging possible questions are given below. These experiments were meant to be done within the PhD study period using an anaerobic chamber that was to be bought and shipped to Malawi, but were not done due to COVID-19.

More broadly, *in-vitro* conditions are not necessarily a true reflection of the conditions in the gut. Different factors, including host factors, influence *Bifidobacterium* and *Salmonella* interactions in the gut. Although specific organisms such as *Bifidobacterium* have been shown to possess antibacterial properties, the gut microbiota colonisation resistance is also mainly aided by the rich diversity of organisms that work in synchrony to control pathogens infection and proliferation.

6.4.2 Experiments that will been conducted in Malawi to validate and extend these exploratory findings

Experiment 1: Growth of *Salmonella* in aerobic and anaerobic conditions

Rationale: As part of experimental transfer from the UK to Malawi**.** Previous experiment used S. Typhimurium only. Strains for non-invasive S. Typhimurium, S. Typhi and S. Enteritidis could be included. It might also be helpful to explore growth patterns at different pH measurements.

Objective: To characterise the growth of different *Salmonella* strains in aerobic and anaerobic conditions, and at different pH

Methodology

Aerobically and an anaerobically Culture invasive (n=1) and non-invasive *S.* Typhimurium (n=1), *S.* Typhi (n=1), *S.* and Enteritidis (n=1) strains in LB broth.

Required material:

Salmonella isolates

Aerobic and anaerobic cabinets

LB broth and agar

OD reader

Experiment 2: Growth of *Bifidobacterium* and *Salmonella* strains in BHI and CDRIM

Rationale: The results reported here were from one set of experiments. There is need to conduct the experiment in triplicate and using a range of strains. The growth curves in Figure 6.1 shows that at 24 hours, most of the *Bifidobacterium* in CDRIM culture had not reached stationary phase. Incubation time for all *Bifidobacterium* strains should therefore be increased to 72 hours.

Objective: To determine the growth of *Bifidobacterium* and *Salmonella* in BHI and CDRIM

Methodology:

Anaerobically culture *Salmonella* (n=4) for 12 hours and *Bifidobacterium* strains (n= 60 [all isolates]) for 72 hours.

Required materials:

Salmonella isolates: invasive (n=1) and non-invasive *S.* Typhimurium (n=1), *S.* Typhi (n=1), *S.* and Enteritidis (n=1) strains

All *Bifidobacterium* isolates: Malawian, type strains and common probiotic strains

BHI

CDRIM

Anaerobic cabinet

OD reader

Experiment 3: *Bifidobacterium* and *Salmonella* co-culture experiments

Rationale: Data reported from this experiment used two *B. longum* strains. The planned experiments would test all the *Bifidobacterium* isolates and four *Salmonella* strains. In addition to the Malawi *Bifidobacterium*, type strains for the species identified in Malawian children will be included in the test. Some of the *Bifidobacterium* strains that are commonly used as probiotics will be sourced and tested. In the reported results, *Bifidobacterium* and *Salmonella* were co-cultured at time point zero where both *Bifidobacterium* and *Salmonella* were on lag phase. The hypothesis, however, is that, different growth stages has different pathways activates and different metabolites released. The plan is therefore to add *Salmonella* strains when *Bifidobacterium* is at lag, exponential, stationery, and death growth phases.

Objectives:

1. To determine if *Bifidobacterium* from Malawian infants has anti-infective properties.
2. To compare anti-infective properties of Malawi *Bifidobacterium* with the anti-infective properties of *Bifidobacterium* type strains and common probiotic strains
3. To determine the effect of different *Bifidobacterium* growth stages on *Salmonella* growth
4. To monitor pH levels in the co-culture experiments

Methodology:

Co-culture *Bifidobacterium* and *Salmonella* at different growth phases of *Bifidobacterium*.

Required material:

All *Bifidobacterium* strains: Malawian, type strains and common probiotic strains

Four *Salmonella* strains: invasive (n=1) and non-invasive *S.* Typhimurium (n=1), *S.* Typhi (n=1), *S.* and Enteritidis (n=1) strains

Anaerobic and aerobic cabinets

CDRIM liquid media

MRS and LB agar

Experiment 4: *Salmonella* growth in *Bifidobacterium* supernatant

Rationale: Only a selected number of strains were tested. There is need to test all *Salmonella* and *Bifidobacterium* strains.

Objectives:

1. To determine if neat or serially-diluted *Bifidobacterium* supernatant from Malawian infants has anti-infective properties.
2. To compare anti-infective properties of Malawi *Bifidobacterium* supernatant with the anti-infective properties of supernatant from *Bifidobacterium* of type strains and common probiotic strains
3. To determine the anti-infective properties of supernatant collected from different *Bifidobacterium* growth stages on *Salmonella* growth

Methodology:

Culture *Salmonella* in neat and diluted *Bifidobacterium* supernatant collected at different growth phases of *Bifidobacterium*.

Required material:

All *Bifidobacterium* strains: Malawian, type strains and common probiotic strains

Four *Salmonella* strains: invasive (n=1) and non-invasive *S.* Typhimurium (n=1), *S.* Typhi (n=1), *S.* and Enteritidis (n=1) strains

Anaerobic and aerobic cabinets

CDRIM liquid media

MRS and LB agar

Filter paper

Experiment 5: Characterize short chain fatty acids and other metabolites for *Bifidobacterium* isolates that show inhibitory properties

Rationale: Data from the exploratory experiments has shown that Malawian strains of *Bifidobacterium* in particular have anti-infective properties against *S.* Typhimurium. This was shown in both the *Bifidobacterium* and *Salmonella* co-culture experiment as well as the experiment that used *Bifidobacterium* supernatant. One reported mechanism that *Bifidobacterium* uses to control pathogens is through the release of SCFA. To understand the reported anti-infective property, this experiment will characterize the SCFA released.

Methodology: Targeted mass spectrometry will be conducted to determine the SCFA that might be release during a *Bifidobacterium* and *Salmonella* co-culture experiment. For a broader view of the different metabolites that might be released, an untargeted mass spectrometry will be conducted. If resources are available, the use of proteomics will be explored.

Required material:

Supernatant from a co-culture experiment that has shown anti-infective properties.

Mass spectrophotometer

6.5 Summary

The preliminary and exploratory work reported here has demonstrated that some *Bifidobacterium* species isolated from Malawian children possess anti-infective activities against the invasive *S*. Typhimurium, D23580. *Bifidobacterium* anti-infection activities observed might be through direct *S*. Typhimurium – *Bifidobacterium* interaction or through the production of soluble metabolites that suppresses or inhibit the growth of D23580 or through the depletion of nutrients or other substances required for *Salmonella* growth. Before establishing these findings through further experimental work, there is need to transfer the UK experimental setup to Malawi and rigorously repeat the experiments.

CHAPTER 7: GENERAL DISCUSSION

7.1 Introduction

The gut microbiota is critical in ensuring good human health and wellbeing. Throughout the entire life span, the gut microbiota plays multiple roles that influence the health state of the host (Milani et al., 2020; Milani et al., 2017). Colonisation resistance against invading pathogens is one of the significant roles that the gut microbiota plays in ensuring good health (Antonini et al., 2019; Dinan & Cryan, 2017; Milani et al., 2020; Riba et al., 2020; Rowland et al., 2018). Gut microbiota dysbiosis during early life not only has an effect in the immediate but also in the long term. Infection, disease, and antibiotic usage are some of the drivers of gut microbiota dysbiosis. It is established that gut microbiota dysbiosis during the seeding and development stages of the gut microbiota is a risk factor for metabolic and immune disorders that occur later in life (Huh et al., 2012; Sevelsted, Stokholm, Bønnelykke, & Bisgaard, 2015). There is, therefore, a growing interest in understanding the gut microbiota in childhood and finding ways of modulating the gut microbiota in particular health conditions to ensure hosts immediate and long-term health and wellbeing.

The gut is non-sterile and frequently exposed to pathogens. In low-resource countries like Malawi, drinking untreated water and poor sanitation increases the risk of enteric pathogen exposure. This study aimed to determine the gut microbiota patterns associated with asymptomatic enteric pathogen exposure in the gut lumen and describe and characterise the early gut microbiota member, *Bifidobacterium*, among Malawian children. Specifically, a description of enteric pathogen exposure events and the relationships between gut microbiota composition and asymptomatic

exposure to the enteric pathogens among healthy Malawian children was made. *Bifidobacterium* was also isolated from healthy exclusively breastfed Malawian infants and characterised. Exploration of the anti-*Salmonella* properties of the isolated *Bifidobacterium* was also conducted

Multiple molecular laboratory methods were used in the study. A well-established semi-quantitative Taqman array card assay was used to detect multiple enteric pathogens from stool samples of children aged 6 to 18 months (J. Liu et al., 2013). Using Qiagen Fast stool Minikit, DNA was extracted from the same stool samples for 16S rRNA sequencing. A pilot study was also conducted where a cohort of 30 participants was used to study gut microbiota composition and *Bifidobacterium* characteristics in healthy exclusively breastfed children. Stool samples collected from these children were used to isolate *Bifidobacterium* and extract DNA for 16S rRNA sequencing. Whole-genome sequencing was done DNA from *Bifidobacterium* isolates. Illumina miseq sequencing was used for both studies, sequenced data was analysed using Qiime 2, and statistical analysis was done using R and RStudio. *Invitro* competitive assays by co-culturing *S.* Typhimurium and *Bifidobacterium* or in *Bifidobacterium* supernatant were also conducted to determine if *Bifidobacterium* from Malawian infants has anti-infective properties against *S.* Typhimurium.

7.2 Discussion of key findings

7.2.1 Malawian children are asymptomatically exposed to multiple enteric pathogens

In this study, a wide range of pathogens were detected in Malawian children aged 6 to 18 months, with an average of 3 detected pathogens per participant at a particular

time point. Globally, Malawi ranks 5th on the list of least developed countries (WHO, 2021). Poor sanitation and untreated drinking water are commonplace, mainly in densely populated urban areas (Cassivi, Tilley, Waygood, & Dorea, 2020; Kalumbi, Thaulo, MacPherson, & Morse, 2020). These environmental factors predispose children to different pathogens quite early in life. These pathogens can cause gastroenteritis, invasive disease, and in some cases, the exposure is asymptomatic. In this study, children were asymptomatically exposed to the pathogens with only a few episodes of clinical symptoms during the follow-up period. Asymptomatic exposure is a risk factor for gastroenteritis and invasive disease. The importance of asymptomatic exposure of pathogens such as E. coli and Enterovirus to growth status, susceptibility to infectious and non-infectious diseases in the immediate and long term, and vaccine effectiveness should be investigated.

7.2.2 Enteric pathogen exposure was not significantly associated with gut microbiota changes

The study findings support the fact that in young children, the diversity and richness of the gut microbiota increase with age and that *Bifidobacterium* is the dominant gut microbiota member. Significant associations were observed in alpha and beta diversity with *Giardia*, EAEC, and *B. fragilis* exposure when data were analysed in a cross-sectional manner. This significant association was lost when the age-adjusted analysis was done. There were no significant associations in alpha and beta diversity with participants characteristics, clinical symptoms, antibiotic usage, and markers of economic status; water source and drinking of boiled drinking water. The lack of significant association could mainly be attributed to the exposure being primarily asymptomatic and low statistical power due to the loss of participants during follow-

up and loss of samples. Data from acute enteric pathogen infection may demonstrate different results.

7.2.3 Exclusively breastfeeding Malawian infants have a high abundance of *Bifidobacterium* possessing a wide range of HMO digesting genes

Gut microbiota profiling of exclusively breastfeeding Malawian infants in this study demonstrated that the relative abundance of *Bifidobacterium* is higher than other intestinal organisms. There were no differences in alpha and beta diversities with age, sex, *Bifidobacterium* culture results, and *identity of Bifidobacterium species*. Most of the isolated *Bifidobacterium* were B. *longum* species. Genomes of all the identified *Bifidobacterium* species, *B. bifidum, B. breve, B. longum* and *B. Pseudocatenulatum,* possessed HMO digesting genes. There were, however, interspecies variations in the HMO digesting genes present.

7.2.4 *Bifidobacterium* found in breastfeeding Malawian infants seem to possess antibacterial properties against the invasive S. Typhimurium, D23580

The exploratory experiments conducted in this study have demonstrated that Malawi *Bifidobacterium* strains possess antibacterial properties against *S.* Typhimurium. Strains from *B. longum* Cluster A demonstrated the highest inhibitory properties. Culturing *S.* Typhimurium in *Bifidobacterium* supernatant has also shown that strains of *Bifidobacterium* from Malawian exclusively breastfeeding infants inhibit the growth of *S.* Typhimurium, suggesting that these *Bifidobacterium* strains produce soluble substances that have antibacterial properties against *S.* Typhimurium.

7.3 Study limitations

The longitudinal data set used to analyse the gut microbiota of Malawian children aged 6 to 18 months was not sufficiently powered to provide a conclusive answer on whether there are changes in the gut microbiota due to enteric pathogen exposure. Missing samples, the longitudinal nature of the data and the low positivity rate for some pathogens due to the asymptomatic nature of the study cohort were some of the main reasons that contributed to the loss of power. By focusing on just a few pathogens, understanding the microbiome-pathogen interactions would simplify analysis and draw clear conclusions as the study can be specifically powered for those few pathogens.

Studying the gut microbiota using 16S rRNA sequenced data limits one to only understanding the compositional differences. Another very important element of gut microbiome studies is the functional roles of the genes present in the organisms making up the microbiota. Metagenomic studies are therefore more informative as they look at both the composition and the function. Metagenomic studies are, however, more expensive than studies using 16S rRNA sequencing.

Isolating *Bifidobacterium* from frozen samples was one of the limitations of the BIFBAC study. Although freezing samples at -80^{o}C preserve most organisms, the ideal time to culture a sample to isolate an organism is when the sample is fresh. Isolating from a frozen sample may reduce the cultivability of some microorganisms. The conditions used to determine the anti-infective properties of *Bifidobacterium* against *S.* Typhimurium are not similar to the conditions in the gut. More experiments involving adjusting growth conditions or animal models would provide more insightful

information. Most experiments were done once. More replicates need to be done to get conclusive results. Only a few samples from the different *Bifidobacterium longum* subspecies were used in the competitive experiments. More samples should be used to have conclusive results.

7.4 Recommendations for further research

7.4.1 Recommendations for further research on enteric pathogen exposure and changes in gut microbiota profiles

1. Well powered studies focusing on one or a small number of pathogens should be done to clearly understand the role of gut microbiota in colonisation resistance against those particular pathogens.
2. Investigate the role of the gut microbiota in invasive disease. *Salmonella* is a very important cause of invasive disease in Malawi. Studies focusing on both typhoidal and non-typhoidal *Salmonella* should be conducted.
3. Investigate the association between oral vaccine effectiveness and gut microbiota composition and *Bifidobacterium* specifically
4. Studies employing shotgun metagenomics should be done to understand the functional aspect of the microbiota with a different disease state of interest

7.4.2 Recommendations for further research on *Bifidobacterium* as an essential health-promoting gut microbiota member

1. Investigate *Bifidobacterium* phenotypic and genotypic properties in children across a broad age group.
2. Describe *Bifidobacterium* phenotypic and genotypic properties in mothers and infant pairs

3. Describe the HMO digestive gene repertoire in the Malawi isolates
4. Describe *Bifidobacterium* strains or variants in a broad host range.
5. Explore how *Bifidobacterium* could be used in promoting vaccine effectiveness

7.5 Concluding remarks

Overall, this PhD research has contributed to our understanding of the vast repertoire of pathogens that Malawian children are exposed to. The impact of asymptomatic exposure observed here should be further investigated to understand further the impact that it may have on child growth and interventions such as oral vaccine effectiveness. This high burden calls for a multidimensional approach in controlling enteric infections. Harnessing the beneficial properties of the gut microbiota and its member in controlling enteric infections could be one arm of this multifaceted approach to fighting enteric infections. For the first time in Malawi, the study has described and genomically characterized a beneficial member of the gut microbiota, *Bifidobacterium*. The gut of Malawian exclusively breastfed children is rich in *Bifidobacterium,* mainly *B. longum*. The Malawi *Bifidobacterium* is genomically different from global *Bifidobacterium* and possesses HMO genes that vary based on species or *B. longum* clusters. Some Malawi *Bifidobacterium* strains demonstrated antibacterial properties against the invasive *S.* Typhimurium ST313 strain.

References

Aagaard, K., Ma, J., Antony, K. M., Ganu, R., Petrosino, J., & Versalovic, J. (2014). The Placenta Harbors a Unique Microbiome. *Science translational medicine, 6*(237), 237ra265-237ra265. doi:10.1126/scitranslmed.3008599

aldrich, S. (2021). Brain Heart Infusion broth. Retrieved from https://www.sigmaaldrich.com/MW/en/product/sial/53286?gclid=CjwKCAiA4veMBhAMEiwAU4XRrz7M0Tm1HwZz1SFj5gtCWH97u_joZx2-boZ34zKnFyOBx6kXLKndxhoCKSYQAvD_BwE

Alessandri, G., Ossiprandi, M. C., MacSharry, J., van Sinderen, D., & Ventura, M. (2019). Bifidobacterial Dialogue With Its Human Host and Consequent Modulation of the Immune System. *Front Immunol, 10*, 2348. doi:10.3389/fimmu.2019.02348

Andino, A., & Hanning, I. (2015). Salmonella enterica: survival, colonization, and virulence differences among serovars. *ScientificWorldJournal, 2015*, 520179. doi:10.1155/2015/520179

Andriessen, E. M., Wilson, A. M., Mawambo, G., Dejda, A., Miloudi, K., Sennlaub, F., & Sapieha, P. (2016). Gut microbiota influences pathological angiogenesis in obesity-driven choroidal neovascularization. *EMBO molecular medicine, 8*(12), 1366-1379. doi:10.15252/emmm.201606531

Antonini, M., Lo Conte, M., Sorini, C., & Falcone, M. (2019). How the Interplay Between the Commensal Microbiota, Gut Barrier Integrity, and Mucosal Immunity Regulates Brain Autoimmunity. *Front Immunol, 10*, 1937. doi:10.3389/fimmu.2019.01937

Ao, T. T., Feasey, N. A., Gordon, M. A., Keddy, K. H., Angulo, F. J., & Crump, J. A. (2015). Global burden of invasive nontyphoidal Salmonella disease, 2010(1). *Emerg Infect Dis, 21*(6). doi:10.3201/eid2106.140999

Apanga, P. A., Weber, A. M., Darrow, L. A., Riddle, M. S., Tung, W. C., Liu, Y., & Garn, J. V. (2021). The interrelationship between water access, exclusive breastfeeding and diarrhea in children: a cross-sectional assessment across 19 African countries. *J Glob Health, 11*, 04001. doi:10.7189/jogh-11-04001

Aw, W., & Fukuda, S. (2019). Protective effects of bifidobacteria against enteropathogens. *Microbial biotechnology, 12*(6), 1097-1100. doi:10.1111/1751-7915.13460

B Romond, M., Ais, A., Guillemot, F., Bounouader, R., Md, A., & Romond, C. (1998). *Cell-Free Whey from Milk Fermented with Bifidobacterium breve C50 Used to Modify the Colonic Microflora of Healthy Subjects* (Vol. 81).

Bäckhed, F., Ding, H., Wang, T., Hooper, L. V., Koh, G. Y., Nagy, A., . . . Gordon, J. I. (2004). The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci U S A, 101*(44), 15718-15723. doi:10.1073/pnas.0407076101

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., . . . Wang, J. (2015a). Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe, 17*(5), 690-703. doi:https://doi.org/10.1016/j.chom.2015.04.004

Bäckhed, F., Roswall, J., Peng, Y., Feng, Q., Jia, H., Kovatcheva-Datchary, P., . . . Wang, J. (2015b). Resource: Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life. *Cell Host Microbe, 17*, 690-703. doi:10.1016/j.chom.2015.04.004

Bank, W. (2020). The World Bank in Malawi. Retrieved from https://www.worldbank.org/en/country/malawi/overview

Baothman, O. A., Zamzami, M. A., Taher, I., Abubaker, J., & Abu-Farha, M. (2016). The role of Gut Microbiota in the development of obesity and Diabetes. *Lipids in Health and Disease, 15*, 108. doi:10.1186/s12944-016-0278-4

Bar-Zeev, N., Jere, K. C., Bennett, A., Pollock, L., Tate, J. E., Nakagomi, O., . . . Cunliffe, N. A. (2016). Population Impact and Effectiveness of Monovalent Rotavirus Vaccination in Urban Malawian Children 3 Years After Vaccine Introduction: Ecological and Case-Control Analyses. *Clin Infect Dis, 62 Suppl 2*(Suppl 2), S213-219. doi:10.1093/cid/civ1183

Bartelt, L. A., Bolick, D. T., Mayneris-Perxachs, J., Kolling, G. L., Medlock, G. L., Zaenker, E. I., . . . Guerrant, R. L. (2017). Cross-modulation of pathogen-specific pathways enhances malnutrition during enteric co-infection with Giardia lamblia and enteroaggregative Escherichia coli. *PLoS Pathog, 13*(7), e1006471. doi:10.1371/journal.ppat.1006471

Belkaid, Y., & Hand, T. (2014). Role of the Microbiota in Immunity and inflammation. *Cell, 157*(1), 121-141. doi:10.1016/j.cell.2014.03.011

Beller, L., Deboutte, W., Falony, G., Vieira-Silva, S., Tito, R. Y., Valles-Colomer, M., . . . Matthijnssens, J. (2021). Successional Stages in Infant Gut Microbiota Maturation. *MBio, 12*(6), e0185721. doi:10.1128/mBio.01857-21

Bender, J. M., Li, F., Martelly, S., Byrt, E., Rouzier, V., Leo, M., . . . Aldrovandi, G. M. (2016). Maternal HIV infection influences the microbiome of HIV-uninfected infants. *Science translational medicine, 8*(349), 349ra100-349ra100. doi:10.1126/scitranslmed.aaf5103

Birkhold, M., Mwisongo, A., Pollard, A. J., & Neuzil, K. M. (2021). Typhoid Conjugate Vaccines: Advancing the Research and Public Health Agendas. *J Infect Dis, 224*(12 Suppl 2), S781-s787. doi:10.1093/infdis/jiab449

Bode, L. (2012). Human milk oligosaccharides: every baby needs a sugar mama. *Glycobiology, 22*(9), 1147-1162. doi:10.1093/glycob/cws074

Boltin, D., & Niv, Y. (2014). Pharmacological and alimentary alteration of the gastric barrier. *Best Practice & Research Clinical Gastroenterology, 28*(6), 981-994. doi:https://doi.org/10.1016/j.bpg.2014.09.001

Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., . . . Caporaso, J. G. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol, 37*(8), 852-857. doi:10.1038/s41587-019-0209-9

Bonder, M. J., Kurilshikov, A., Tigchelaar, E. F., Mujagic, Z., Imhann, F., Vila, A. V., . . . Zhernakova, A. (2016). The effect of host genetics on the gut microbiome. *Nat Genet, 48*(11), 1407-1412. doi:10.1038/ng.3663

Bondue, P., Crèvecoeur, S., Brose, F., Daube, G., Seghaye, M.-C., Griffiths, M. W., . . . Delcenserie, V. (2016). Cell-Free Spent Media Obtained from Bifidobacterium bifidum and Bifidobacterium crudilactis Grown in Media Supplemented with 3′-Sialyllactose Modulate Virulence Gene Expression in Escherichia coli O157:H7 and Salmonella Typhimurium. *Front Microbiol, 7*(1460). doi:10.3389/fmicb.2016.01460

Boyce, M. R., Katz, R., & Standley, C. J. (2019). Risk Factors for Infectious Diseases in Urban Environments of Sub-Saharan Africa: A Systematic Review and Critical Appraisal of Evidence. *Trop Med Infect Dis, 4*(4). doi:10.3390/tropicalmed4040123

Brenner, H., Rothenbacher, D., Bode, G., & Adler, G. (1997). Relation of smoking and alcohol and coffee consumption to active Helicobacter pylori infection: cross sectional study. *Bmj, 315*(7121), 1489-1492. doi:10.1136/bmj.315.7121.1489

Brook, I. (2010). The role of anaerobic bacteria in bacteremia. *Anaerobe, 16*(3), 183-189. doi:https://doi.org/10.1016/j.anaerobe.2009.12.001

Brooks, A. W., Priya, S., Blekhman, R., & Bordenstein, S. R. (2018). Gut microbiota diversity across ethnicities in the United States. *PLoS Biol, 16*(12), e2006842. doi:10.1371/journal.pbio.2006842

Bunker, Jeffrey J., Flynn, Theodore M., Koval, Jason C., Shaw, Dustin G., Meisel, M., McDonald, Benjamin D., . . . Bendelac, A. (2015). Innate and Adaptive Humoral Responses Coat Distinct Commensal Bacteria with Immunoglobulin A. *Immunity, 43*(3), 541-553. doi:https://doi.org/10.1016/j.immuni.2015.08.007

Burnett, E., Jonesteller, C. L., Tate, J. E., Yen, C., & Parashar, U. D. (2017). Global Impact of Rotavirus Vaccination on Childhood Hospitalizations and Mortality From Diarrhea. *J Infect Dis, 215*(11), 1666-1672. doi:10.1093/infdis/jix186

Butera, E., Mukabutera, A., Nsereko, E., Munyanshongore, C., Rujeni, N., Mwikarago, I. E., . . . Manasse, M. N. (2019). Prevalence and risk factors of intestinal parasites among children under two years of age in a rural area of Rutsiro district, Rwanda - a cross-sectional study. *The Pan African medical journal, 32*, 11-11. doi:10.11604/pamj.2019.32.11.15949

Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J., & Holmes, S. P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods, 13*(7), 581-583. doi:10.1038/nmeth.3869

Carter, M. M., Olm, M. R., & Sonnenburg, E. D. (2022). Microbiome assembly in The Gambia. *Nat Microbiol, 7*(1), 18-19. doi:10.1038/s41564-021-01036-1

Carvalho, M. F., & Gill, D. (2019). Rotavirus vaccine efficacy: current status and areas for improvement. *Hum Vaccin Immunother, 15*(6), 1237-1250. doi:10.1080/21645515.2018.1520583

Cassivi, A., Tilley, E., Waygood, E. O. D., & Dorea, C. (2020). Trends in access to water and sanitation in Malawi: progress and inequalities (1992-2017). *J Water Health, 18*(5), 785-797. doi:10.2166/wh.2020.069

Chaguza, C., Nyaga, M. M., Mwenda, J. M., Esona, M. D., & Jere, K. C. (2020). Using genomics to improve preparedness and response of future epidemics or pandemics in Africa. *Lancet Microbe, 1*(7), e275-e276. doi:10.1016/s2666-5247(20)30169-5

Champagne-Jorgensen, K., Mian, M. F., Kay, S., Hanani, H., Ziv, O., McVey Neufeld, K.-A., . . . Bienenstock, J. (2020). Prenatal low-dose penicillin results in long-term sex-specific changes to murine behaviour, immune regulation, and gut microbiota. *Brain, Behavior, and Immunity, 84*, 154-163. doi:https://doi.org/10.1016/j.bbi.2019.11.020

Chan, P. P., Lin, B. Y., Mak, A. J., & Lowe, T. M. (2019). tRNAscan-SE 2.0: Improved Detection and Functional Classification of Transfer RNA Genes. *bioRxiv*, 614032. doi:10.1101/614032

Chen, C., Yang, X., & Shen, X. (2019). Confirmed and Potential Roles of Bacterial T6SSs in the Intestinal Ecosystem. *Front Microbiol, 10*, 1484. doi:10.3389/fmicb.2019.01484

Chen, Y., Zhou, J., & Wang, L. (2021). Role and Mechanism of Gut Microbiota in Human Disease. *Front Cell Infect Microbiol, 11*, 625913. doi:10.3389/fcimb.2021.625913

Cheung, Y. B., Xu, Y., Mangani, C., Fan, Y.-M., Dewey, K. G., Salminen, S. J., . . . Ashorn, P. (2016). Gut microbiota in Malawian infants in a nutritional supplementation trial. *Tropical Medicine & International Health, 21*(2), 283-290. doi:https://doi.org/10.1111/tmi.12650

Chirambo AC, N. T., Jambo N et al. (2020). Performance of molecular methods for the detection of Salmonella in human stool specimens [version 1; peer review: awaiting peer review]. *Wellcome Open Research*(5:237). doi:https://doi.org/10.12688/wellcomeopenres.16305.1

Cho, I., & Blaser, M. J. (2012). The human microbiome: at the interface of health and disease. *Nat Rev Genet, 13*(4), 260-270. doi:10.1038/nrg3182

Chow, J., & Mazmanian, S. K. (2010). A pathobiont of the microbiota balances host colonization and intestinal inflammation. *Cell Host Microbe, 7*(4), 265-276. doi:10.1016/j.chom.2010.03.004

Chunxi, L., Haiyue, L., Yanxia, L., Jianbing, P., & Jin, S. (2020). The Gut Microbiota and Respiratory Diseases: New Evidence. *Journal of immunology research, 2020*, 2340670-2340670. doi:10.1155/2020/2340670

Clarke, S. F., Murphy, E. F., O'Sullivan, O., Lucey, A. J., Humphreys, M., Hogan, A., . . . Cotter, P. D. (2014). Exercise and associated dietary extremes impact on gut microbial diversity. *Gut, 63*(12), 1913-1920. doi:10.1136/gutjnl-2013-306541

Colston, J. M., Faruque, A. S. G., Hossain, M. J., Saha, D., Kanungo, S., Mandomando, I., . . . Kosek, M. N. (2020). Associations between Household-Level Exposures and All-Cause Diarrhea and Pathogen-Specific Enteric Infections in Children Enrolled in Five Sentinel Surveillance Studies. *Int J Environ Res Public Health, 17*(21). doi:10.3390/ijerph17218078

Conlon, M. A., & Bird, A. R. (2015). The Impact of Diet and Lifestyle on Gut Microbiota and Human Health. *Nutrients, 7*(1), 17-44. doi:10.3390/nu7010017

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I., & Knight, R. (2009). Bacterial community variation in human body habitats across space and time. *Science, 326*(5960), 1694-1697. doi:10.1126/science.1177486

Dahl, C., Stigum, H., Valeur, J., Iszatt, N., Lenters, V., Peddada, S., . . . Eggesbø, M. (2018). Preterm infants have distinct microbiomes not explained by mode of delivery, breastfeeding duration or antibiotic exposure. *International journal of epidemiology, 47*(5), 1658-1669. doi:10.1093/ije/dyy064

Dave, M., Higgins, P. D., Middha, S., & Rioux, K. P. (2012). The human gut microbiome: current knowledge, challenges, and future directions. *Transl Res, 160*(4), 246-257. doi:10.1016/j.trsl.2012.05.003

de Jong, S. E., Olin, A., & Pulendran, B. (2020). The Impact of the Microbiome on Immunity to Vaccination in Humans. *Cell Host Microbe, 28*(2), 169-179. doi:10.1016/j.chom.2020.06.014

Desai, M. S., Seekatz, A. M., Koropatkin, N. M., Kamada, N., Hickey, C. A., Wolter, M., . . . Martens, E. C. (2016). A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell, 167*(5), 1339-1353.e1321. doi:10.1016/j.cell.2016.10.043

Diacovich, L., Lorenzi, L., Tomassetti, M., Méresse, S., & Gramajo, H. (2017). The infectious intracellular lifestyle of Salmonella enterica relies on the adaptation to nutritional conditions within the Salmonella-containing vacuole. *Virulence, 8*(6), 975-992. doi:10.1080/21505594.2016.1270493

Dinan, T. G., & Cryan, J. F. (2017). Gut instincts: microbiota as a key regulator of brain development, ageing and neurodegeneration. *J Physiol, 595*(2), 489-503. doi:10.1113/jp273106
Ding, T., & Schloss, P. D. (2014). Dynamics and associations of microbial community types across the human body. *Nature, 509*(7500), 357-360. doi:10.1038/nature13178
Dixon, P. (2003). VEGAN, a package of R functions for community ecology. *Journal of Vegetation Science, 14*(6), 927-930.
Dominguez-Bello, M. G., Costello, E. K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., & Knight, R. (2010). Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc Natl Acad Sci U S A, 107*(26), 11971-11975. doi:10.1073/pnas.1002601107
Donaldson, G. P., Lee, S. M., & Mazmanian, S. K. (2016). Gut biogeography of the bacterial microbiota. *Nat Rev Microbiol, 14*(1), 20-32. doi:10.1038/nrmicro3552
Dong, T., Chen, T., White, R. A., 3rd, Wang, X., Hu, W., Liang, Y., . . . Xia, Y. (2018). Meconium microbiome associates with the development of neonatal jaundice. *Clinical and translational gastroenterology, 9*(9), 182-182. doi:10.1038/s41424-018-0048-x
Ducarmon, Q. R., Zwittink, R. D., Hornung, B. V. H., van Schaik, W., Young, V. B., & Kuijper, E. J. (2019). Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. *Microbiol Mol Biol Rev, 83*(3). doi:10.1128/mmbr.00007-19
Dunn, J., & Grider, M. (2020). StatPearls. In: Treasure Island: StatPearls Publishing) Physiology, Adenosine Triphosphate.
Duranti, S., Longhi, G., Ventura, M., van Sinderen, D., & Turroni, F. (2020). Exploring the Ecology of Bifidobacteria and Their Genetic Adaptation to the Mammalian Gut. *Microorganisms, 9*(1). doi:10.3390/microorganisms9010008
Eberl, C., Weiss, A. S., Jochum, L. M., Durai Raj, A. C., Ring, D., Hussain, S., . . . Stecher, B. (2021). E. coli enhance colonization resistance against Salmonella Typhimurium by competing for galactitol, a context-dependent limiting carbon source. *Cell Host Microbe, 29*(11), 1680-1692.e1687. doi:10.1016/j.chom.2021.09.004
Eggesbø, M., Baird, D., Moen, B., Sekelja, M., Rudi, K., Peddada, S., . . . Midtvedt, T. (2011). Development of gut microbiota in infants not exposed to medical interventions. *APMIS, 119*(1), 17-35. doi:10.1111/j.1600-0463.2010.02688.x
Enomoto, T., Sowa, M., Nishimori, K., Shimazu, S., Yoshida, A., Yamada, K., . . . Xiao, J.-z. (2014). Effects of Bifidobacterial Supplementation to Pregnant Women and Infants in the Prevention of Allergy Development in Infants and on Fecal Microbiota. *Allergology International, 63*(4), 575-585. doi:https://doi.org/10.2332/allergolint.13-OA-0683
Eyamo, T., Girma, M., Alemayehu, T., & Bedewi, Z. (2019). Soil-Transmitted Helminths And Other Intestinal Parasites Among Schoolchildren In Southern Ethiopia. *Research and reports in tropical medicine, 10*, 137-143. doi:10.2147/RRTM.S210200
Fan, W., Huo, G., Li, X., Yang, L., & Duan, C. (2014). Impact of Diet in Shaping Gut Microbiota Revealed by a Comparative Study in Infants During the First Six Months of Life. *Journal of microbiology and biotechnology*(2), 133.

Fan, Y., & Pedersen, O. (2021). Gut microbiota in human metabolic health and disease. *Nat Rev Microbiol, 19*(1), 55-71. doi:10.1038/s41579-020-0433-9
Fan, Y. M., Oikarinen, S., Lehto, K. M., Nurminen, N., Juuti, R., Mangani, C., . . . Ashorn, P. (2019). High prevalence of selected viruses and parasites and their predictors in Malawian children. *Epidemiology and infection, 147*, e90-e90. doi:10.1017/S0950268819000025
Fanning, S., Hall, L. J., & van Sinderen, D. (2012). Bifidobacterium breve UCC2003 surface exopolysaccharide production is a beneficial trait mediating commensal-host interaction through immune modulation and pathogen protection. *Gut Microbes, 3*(5), 420-425. doi:10.4161/gmic.20630
Farrell, S. H., Truscott, J. E., & Anderson, R. M. (2017). The importance of patient compliance in repeated rounds of mass drug administration (MDA) for the elimination of intestinal helminth transmission. *Parasites & Vectors, 10*(1), 291. doi:10.1186/s13071-017-2206-5
Feasey, N. A., Dougan, G., Kingsley, R. A., Heyderman, R. S., & Gordon, M. A. (2012). Invasive non-typhoidal salmonella disease: an emerging and neglected tropical disease in Africa. *Lancet, 379*(9835), 2489-2499. doi:10.1016/s0140-6736(11)61752-2
Feasey, N. A., Everett, D., Faragher, E. B., Roca-Feltrer, A., Kang'ombe, A., Denis, B., . . . Heyderman, R. S. (2015). Modelling the Contributions of Malaria, HIV, Malnutrition and Rainfall to the Decline in Paediatric Invasive Non-typhoidal Salmonella Disease in Malawi. *PLoS Negl Trop Dis, 9*(7), e0003979. doi:10.1371/journal.pntd.0003979
Fijan, S., Sulc, D., & Steyer, A. (2018). Study of the In Vitro Antagonistic Activity of Various Single-Strain and Multi-Strain Probiotics against Escherichia coli. *Int J Environ Res Public Health, 15*(7). doi:10.3390/ijerph15071539
Fraher, M. H., O'Toole, P. W., & Quigley, E. M. (2012). Techniques used to characterize the gut microbiota: a guide for the clinician. *Nat Rev Gastroenterol Hepatol, 9*(6), 312-322. doi:10.1038/nrgastro.2012.44
Fujimura, K. E., Sitarik, A. R., Havstad, S., Lin, D. L., Levan, S., Fadrosh, D., . . . Lynch, S. V. (2016). Neonatal gut microbiota associates with childhood multisensitized atopy and T cell differentiation. *Nat Med, 22*(10), 1187-1191. doi:10.1038/nm.4176
Gantois, I., Ducatelle, R., Pasmans, F., Haesebrouck, F., Hautefort, I., Thompson, A., . . . Van Immerseel, F. (2006). Butyrate specifically down-regulates salmonella pathogenicity island 1 gene expression. *Appl Environ Microbiol, 72*(1), 946-949. doi:10.1128/aem.72.1.946-949.2006
García, C., Tebbji, F., Daigneault, M., Liu, N. N., Köhler, J. R., Allen-Vercoe, E., & Sellam, A. (2017). The Human Gut Microbial Metabolome Modulates Fungal Growth via the TOR Signaling Pathway. *mSphere, 2*(6). doi:10.1128/mSphere.00555-17
GBD, C. (2019a). The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis*, S1473-3099(1419)30418-30419. doi:10.1016/S1473-3099(19)30418-9
GBD, C. (2019b). The global burden of typhoid and paratyphoid fevers: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infect Dis*. doi:10.1016/s1473-3099(18)30685-6
GBDa. (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of diarrhoea in 195 countries: a systematic analysis for the

Global Burden of Disease Study 2016. *Lancet Infect Dis, 18*(11), 1211-1228. doi:10.1016/s1473-3099(18)30362-1
GBDb. (2018). Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet, 392*(10159), 1789-1858. doi:10.1016/s0140-6736(18)32279-7
Gibson, G. R., Hutkins, R., Sanders, M. E., Prescott, S. L., Reimer, R. A., Salminen, S. J., . . . Reid, G. (2017). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics (ISAPP) consensus statement on the definition and scope of prebiotics. *Nature reviews Gastroenterology & hepatology, 14*(8), 491-502. doi:10.1038/nrgastro.2017.75
Gilchrist, J. J., & MacLennan, C. A. (2019). Invasive Nontyphoidal Salmonella Disease in Africa. *EcoSal Plus, 8*(2). doi:10.1128/ecosalplus.ESP-0007-2018
Gill, S. R., Pop, M., Deboy, R. T., Eckburg, P. B., Turnbaugh, P. J., Samuel, B. S., . . . Nelson, K. E. (2006). Metagenomic analysis of the human distal gut microbiome. *Science, 312*(5778), 1355-1359. doi:10.1126/science.1124234
Gillis, C. C., Hughes, E. R., Spiga, L., Winter, M. G., Zhu, W., Furtado de Carvalho, T., . . . Winter, S. E. (2018). Dysbiosis-Associated Change in Host Metabolism Generates Lactate to Support Salmonella Growth. *Cell Host Microbe, 23*(1), 54-64.e56. doi:10.1016/j.chom.2017.11.006
Godfrey, K. M., Reynolds, R. M., Prescott, S. L., Nyirenda, M., Jaddoe, V. W. V., Eriksson, J. G., & Broekman, B. F. P. (2017). Influence of maternal obesity on the long-term health of offspring. *The Lancet Diabetes & Endocrinology, 5*(1), 53-64. doi:https://doi.org/10.1016/S2213-8587(16)30107-3
Gopinath, S., Carden, S., & Monack, D. (2012). Shedding light on Salmonella carriers. *Trends Microbiol, 20*(7), 320-327. doi:10.1016/j.tim.2012.04.004
Gordon, M. A. (2008). Salmonella infections in immunocompromised adults. *The Journal of infection, 56*(6), 413-422. doi:10.1016/j.jinf.2008.03.012
Gordon, M. A. (2011). Invasive nontyphoidal Salmonella disease: epidemiology, pathogenesis and diagnosis. *Curr Opin Infect Dis, 24*(5), 484-489. doi:10.1097/QCO.0b013e32834a9980
Gosalbes, M. J., Llop, S., Valles, Y., Moya, A., Ballester, F., & Francino, M. P. (2013). Meconium microbiota types dominated by lactic acid or enteric bacteria are differentially associated with maternal eczema and respiratory problems in infants. *Clin Exp Allergy, 43*(2), 198-211. doi:10.1111/cea.12063
Graf, J., Henderson, W., Cong, X., Henderson, W. A., Graf, J., & McGrath, J. M. (2015). Early Life Experience and Gut Microbiome The Brain-Gut-Microbiota Signaling System. *ADVANCES IN NEONATAL CARE, 15*(5), 314-323.
Groeger, D., O'Mahony, L., Murphy, E. F., Bourke, J. F., Dinan, T. G., Kiely, B., . . . Quigley, E. M. M. (2013). Bifidobacterium infantis 35624 modulates host inflammatory processes beyond the gut. *Gut Microbes, 4*(4), 325-339. doi:10.4161/gmic.25487
Grönlund, M. M., Gueimonde, M., Laitinen, K., Kociubinski, G., Grönroos, T., Salminen, S., & Isolauri, E. (2007). Maternal breast-milk and intestinal bifidobacteria guide the compositional development of the Bifidobacterium microbiota in infants at risk of allergic disease. *Clin Exp Allergy, 37*(12), 1764-1772. doi:10.1111/j.1365-2222.2007.02849.x
Grześkowiak, Ł., Collado, M. C., Mangani, C., Maleta, K., Laitinen, K., Ashorn, P., . . . Salminen, S. (2012). Distinct gut microbiota in southeastern African and

northern European infants. *J Pediatr Gastroenterol Nutr, 54*(6), 812-816. doi:10.1097/MPG.0b013e318249039c
Haak, B. W., de Jong, H. K., Kostidis, S., Giera, M., Maude, R. R., Samad, R., . . . Wiersinga, W. J. (2020). Altered Patterns of Compositional and Functional Disruption of the Gut Microbiota in Typhoid Fever and Nontyphoidal Febrile Illness. *Open Forum Infect Dis, 7*(7), ofaa251. doi:10.1093/ofid/ofaa251
Hadfield, J., Croucher, N. J., Goater, R. J., Abudahab, K., Aanensen, D. M., & Harris, S. R. (2018). Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics (Oxford, England), 34*(2), 292-293. doi:10.1093/bioinformatics/btx610
Hajela, N., Chattopadhyay, S., Nair, G. B., & Ganguly, N. K. (2020). Intestinal microbiota and vaccine efficacy in children from resource poor settings - potential impact for the usefulness of probiotics? *Benef Microbes, 11*(4), 319-328. doi:10.3920/bm2019.0155
Han, D., Gao, P., Li, R., Tan, P., Xie, J., Zhang, R., & Li, J. (2020). Multicenter assessment of microbial community profiling using 16S rRNA gene sequencing and shotgun metagenomic sequencing. *J Adv Res, 26*, 111-121. doi:10.1016/j.jare.2020.07.010
Hansen, R., Scott, K. P., Khan, S., Martin, J. C., Berry, S. H., Stevenson, M., . . . Hold, G. L. (2015). First-Pass Meconium Samples from Healthy Term Vaginally-Delivered Neonates: An Analysis of the Microbiota. *PLoS One, 10*(7), e0133320. doi:10.1371/journal.pone.0133320
Hapfelmeier, S., Lawson, M. A. E., Slack, E., Kirundi, J. K., Stoel, M., Heikenwalder, M., . . . Macpherson, A. J. (2010). Reversible Microbial Colonization of Germ-Free Mice Reveals the Dynamics of IgA Immune Responses. *Science, 328*(5986), 1705-1709. doi:10.1126/science.1188454
Hardwick, R. J., Werkman, M., Truscott, J. E., & Anderson, R. M. (2019). Stochastic challenges to interrupting helminth transmission. *medRxiv*, 2019.2012.2017.19013490. doi:10.1101/2019.12.17.19013490
Harris, V., Ali, A., Fuentes, S., Korpela, K., Kazi, M., Tate, J., . . . de Vos, W. M. (2018). Rotavirus vaccine response correlates with the infant gut microbiota composition in Pakistan. *Gut Microbes, 9*(2), 93-101. doi:10.1080/19490976.2017.1376162
Harris, V. C., Armah, G., Fuentes, S., Korpela, K. E., Parashar, U., Victor, J. C., . . . de Vos, W. M. (2017). Significant Correlation Between the Infant Gut Microbiome and Rotavirus Vaccine Response in Rural Ghana. *J Infect Dis, 215*(1), 34-41. doi:10.1093/infdis/jiw518
Hasegawa, M., Kamada, N., Jiao, Y., Liu, M. Z., Nunez, G., & Inohara, N. (2012). Protective role of commensals against Clostridium difficile infection via an IL-1beta-mediated positive-feedback loop. *J Immunol, 189*(6), 3085-3091. doi:10.4049/jimmunol.1200821
Hedges, A. J., Shannon, R., & Hobbs, R. P. (1978). Comparison of the precision obtained in counting viable bacteria by the spiral plate maker, the droplette and the Miles & Misra methods. *J Appl Bacteriol, 45*(1), 57-65. doi:10.1111/j.1365-2672.1978.tb04198.x
Henrick, B. M., Rodriguez, L., Lakshmikanth, T., Pou, C., Henckel, E., Arzoomand, A., . . . Brodin, P. (2021). Bifidobacteria-mediated immune system imprinting early in life. *Cell, 184*(15), 3884-3898.e3811. doi:10.1016/j.cell.2021.05.030
Hercberg, S., Galan, P., Preziosi, P., Roussel, A. M., Arnaud, J., Richard, M. J., . . . Favier, A. (1998). Background and rationale behind the SU.VI.MAX Study, a

prevention trial using nutritional doses of a combination of antioxidant vitamins and minerals to reduce cardiovascular diseases and cancers. SUpplementation en VItamines et Minéraux AntioXydants Study. *International journal for vitamin and nutrition research. Internationale Zeitschrift fur Vitamin- und Ernahrungsforschung. Journal international de vitaminologie et de nutrition, 68*(1), 3-20.

Hill, C., Guarner, F., Reid, G., Gibson, G. R., Merenstein, D. J., Pot, B., . . . Salminen, S. (2014). Expert consensus document: The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nature reviews Gastroenterology & hepatology, 11*(8), 506-514.

Hill, C., Guarner, F., Reid, G., Gibson, G. R., Merenstein, D. J., Pot, B., . . . Sanders, M. E. (2014). Expert consensus document. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol, 11*(8), 506-514. doi:10.1038/nrgastro.2014.66

Hill, M. J. (1997). Intestinal flora and endogenous vitamin synthesis. *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation (ECP), 6 Suppl 1*, S43-45.

Holohan, D. R., Van Gool, F., & Bluestone, J. A. (2019). Thymically-derived Foxp3+ regulatory T cells are the primary regulators of type 1 diabetes in the non-obese diabetic mouse model. *PLoS One, 14*(10), e0217728. doi:10.1371/journal.pone.0217728

Hooper, L. V., Stappenbeck, T. S., Hong, C. V., & Gordon, J. I. (2003). Angiogenins: a new class of microbicidal proteins involved in innate immunity. *Nat Immunol, 4*(3), 269-273. doi:10.1038/ni888

Hori, S., Nomura, T., & Sakaguchi, S. (2003). Control of Regulatory T Cell Development by the Transcription Factor Foxp3. *Science, 299*(5609), 1057-1061. doi:10.1126/science.1079490

Hosomi, K., & Kunisawa, J. (2020). Impact of the intestinal environment on the immune responses to vaccination. *Vaccine, 38*(44), 6959-6965. doi:10.1016/j.vaccine.2020.08.079

Hotez, P. J., Brindley, P. J., Bethony, J. M., King, C. H., Pearce, E. J., & Jacobson, J. (2008). Helminth infections: the great neglected tropical diseases. *The Journal of clinical investigation, 118*(4), 1311-1321. doi:10.1172/JCI34261

Hou, Y.-P., He, Q.-Q., Ouyang, H.-M., Peng, H.-S., Wang, Q., Li, J., . . . Yin, A.-H. (2017). Human Gut Microbiota Associated with Obesity in Chinese Children and Adolescents. *BioMed Research International, 2017*, 7585989. doi:10.1155/2017/7585989

Hu, J., Nomura, Y., Bashir, A., Fernandez-Hernandez, H., Itzkowitz, S., Pei, Z., . . . Peter, I. (2013). Diversified microbiota of meconium is affected by maternal diabetes status. *PLoS One, 8*(11), e78257. doi:10.1371/journal.pone.0078257

Huh, S. Y., Rifas-Shiman, S. L., Zera, C. A., Edwards, J. W. R., Oken, E., Weiss, S. T., & Gillman, M. W. (2012). Delivery by caesarean section and risk of obesity in preschool age children: a prospective cohort study. *Archives of disease in childhood, 97*(7), 610-616. doi:10.1136/archdischild-2011-301141

Hyman, R. W., Fukushima, M., Diamond, L., Kumm, J., Giudice, L. C., & Davis, R. W. (2005). Microbes on the human vaginal epithelium. *Proc Natl Acad Sci U S A, 102*(22), 7952-7957. doi:10.1073/pnas.0503236102

iNTS-GBD-Collaborators, G. N.-T. S. I. D. (2019). The global burden of non-typhoidal salmonella invasive disease: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet Infectious Diseases, 19*(12), 1312 - 1324. doi:10.1016/S1473-3099(19)30418-9

Iqbal, N. T., Syed, S., Kabir, F., Jamil, Z., Akhund, T., Qureshi, S., . . . Ali, A. (2019). Pathobiome driven gut inflammation in Pakistani children with Environmental Enteric Dysfunction. *PLoS One, 14*(8), e0221095. doi:10.1371/journal.pone.0221095

Iroh Tam, P. Y., Musicha, P., Kawaza, K., Cornick, J., Denis, B., Freyne, B., . . . Heyderman, R. (2019). Emerging Resistance to Empiric Antimicrobial Regimens for Pediatric Bloodstream Infections in Malawi (1998-2017). *Clin Infect Dis, 69*(1), 61-68. doi:10.1093/cid/ciy834

Ishizuka, S., Iwama, A., Dinoto, A., Suksomcheep, A., Maeta, K., Kasai, T., . . . Yokota, A. (2009). Synbiotic promotion of epithelial proliferation by orally ingested encapsulated Bifidobacterium breve and raffinose in the small intestine of rats. *Molecular Nutrition & Food Research, 53*(S1), S62-S67. doi:10.1002/mnfr.200800041

Iturriza-Gómara, M., Jere, K. C., Hungerford, D., Bar-Zeev, N., Shioda, K., Kanjerwa, O., . . . Cunliffe, N. A. (2019). Etiology of Diarrhea Among Hospitalized Children in Blantyre, Malawi, Following Rotavirus Vaccine Introduction: A Case-Control Study. *J Infect Dis, 220*(2), 213-218. doi:10.1093/infdis/jiz084

J Gregory Caporaso, J. K., Jesse Stombaugh, Kyle Bittinger, Frederic D Bushman, Elizabeth K Costello, Noah Fierer, Antonio Gonzalez Pena, Julia K Goodrich, Jeffrey I Gordon, Gavin A Huttley, Scott T Kelley, Dan Knights, Jeremy E Koenig, Ruth E Ley, Catherine A Lozupone, Daniel McDonald, Brian D Muegge, Meg Pirrung, Jens Reeder, Joel R Sevinsky, Peter J Turnbaugh, William A Walters, Jeremy Widmann, Tanya Yatsunenko, Jesse Zaneveld and Rob Knight (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature Methods*. doi:10.1038/nmeth.f.303

Jacobson, A., Lam, L., Rajendram, M., Tamburini, F., Honeycutt, J., Pham, T., . . . Monack, D. (2018). A Gut Commensal-Produced Metabolite Mediates Colonization Resistance to Salmonella Infection. *Cell Host Microbe, 24*(2), 296-307.e297. doi:10.1016/j.chom.2018.07.002

Jandhyala, S. M., Talukdar, R., Subramanyam, C., Vuyyuru, H., Sasikala, M., & Nageshwar Reddy, D. (2015). Role of the normal gut microbiota. *World J Gastroenterol, 21*(29), 8787-8803. doi:10.3748/wjg.v21.i29.8787

Jiang, F., Meng, D., Weng, M., Zhu, W., Wu, W., Kasper, D., & Walker, W. A. (2017). The symbiotic bacterial surface factor polysaccharide A on Bacteroides fragilis inhibits IL-1β-induced inflammation in human fetal enterocytes via toll receptors 2 and 4. *PLoS One, 12*(3), e0172738-e0172738. doi:10.1371/journal.pone.0172738

Johari, M. I., Besari, A. M., Wan Ghazali, W. S., & Yusof, Z. (2019). Disseminated Salmonella infection. *BMJ Case Rep, 12*(5). doi:10.1136/bcr-2018-226337

Jové Blanco, A., Gutiérrez Vélez, A., Solís-García, G., Salcedo Posadas, A., Bellón Alonso, S., & Rodríguez Cimadevilla, J. L. (2020). Comorbidities and course of lung function in patients with congenital esophageal atresia. *Arch Argent Pediatr, 118*(1), 25-30. doi:10.5546/aap.2020.eng.25

Jun, S., Si, F., Pugatch, R., & Scott, M. (2018). Fundamental principles in bacterial physiology-history, recent progress, and the future with focus on cell size

control: a review. *Reports on progress in physics. Physical Society (Great Britain), 81*(5), 056601-056601. doi:10.1088/1361-6633/aaa628
Kåhrström, C. T. (2012). Pro-angiogenic bacteria in the gut. *Nature Reviews Microbiology, 10*(5), 310-310. doi:10.1038/nrmicro2788
Kalumbi, L. R., Thaulo, C., MacPherson, E. E., & Morse, T. (2020). Perspectives and Practices on Water, Sanitation, and Hygiene from a Fishing Community along Lake Malombe, Southern Malawi. *Int J Environ Res Public Health, 17*(18). doi:10.3390/ijerph17186703
Kamada, N., Chen, G. Y., Inohara, N., & Nunez, G. (2013). Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol, 14*(7), 685-690. doi:10.1038/ni.2608
Kamada, N., Kim, Y. G., Sham, H. P., Vallance, B. A., Puente, J. L., Martens, E. C., & Nunez, G. (2012). Regulated virulence controls the ability of a pathogen to compete with the gut microbiota. *Science, 336*(6086), 1325-1329. doi:10.1126/science.1222195
Kamada, N., & Nunez, G. (2013). Role of the gut microbiota in the development and function of lymphoid cells. *J Immunol, 190*(4), 1389-1395. doi:10.4049/jimmunol.1203100
Kamada, N., & Nunez, G. (2014). Regulation of the immune system by the resident intestinal bacteria. *Gastroenterology, 146*(6), 1477-1488. doi:10.1053/j.gastro.2014.01.060
Kamada, N., Seo, S. U., Chen, G. Y., & Nunez, G. (2013). Role of the gut microbiota in immunity and inflammatory disease. *Nat Rev Immunol, 13*(5), 321-335. doi:10.1038/nri3430
Kamng'ona, A. W., Young, R., Arnold, C. D., Kortekangas, E., Patson, N., Jorgensen, J. M., . . . Dewey, K. G. (2019). The association of gut microbiota characteristics in Malawian infants with growth and inflammation. *Sci Rep, 9*(1), 12893. doi:10.1038/s41598-019-49274-y
Kamng'ona, A. W., Young, R., Arnold, C. D., Patson, N., Jorgensen, J. M., Kortekangas, E., . . . Dewey, K. G. (2020). Provision of Lipid-Based Nutrient Supplements to Mothers During Pregnancy and 6 Months Postpartum and to Their Infants from 6 to 18 Months Promotes Infant Gut Microbiota Diversity at 18 Months of Age but Not Microbiota Maturation in a Rural Malawian Setting: Secondary Outcomes of a Randomized Trial. *J Nutr, 150*(4), 918-928. doi:10.1093/jn/nxz298
Kaoutari, A. E., Armougom, F., Gordon, J. I., Raoult, D., & Henrissat, B. (2013). The abundance and variety of carbohydrate-active enzymes in the human gut microbiota. *Nature Reviews Microbiology, 11*(7), 497-504. doi:10.1038/nrmicro3050
Kariuki, S., Revathi, G., Kariuki, N., Kiiru, J., Mwituria, J., Muyodi, J., . . . Hart, C. A. (2006). Invasive multidrug-resistant non-typhoidal Salmonella infections in Africa: zoonotic or anthroponotic transmission? *J Med Microbiol, 55*(Pt 5), 585-591. doi:10.1099/jmm.0.46375-0
Karkey, A., Thwaites, G. E., & Baker, S. (2018). The evolution of antimicrobial resistance in Salmonella Typhi. *Curr Opin Gastroenterol, 34*(1), 25-30. doi:10.1097/mog.0000000000000406
Katayama, T. (2016). Host-derived glycans serve as selected nutrients for the gut microbe: human milk oligosaccharides and bifidobacteria. *Bioscience, Biotechnology, and Biochemistry, 80*(4), 621-632. doi:10.1080/09168451.2015.1132153

Keiser, J., & Utzinger, J. (2019). Community-wide soil-transmitted helminth treatment is equity-effective. *The Lancet, 393*(10185), 2011-2012. doi:10.1016/S0140-6736(18)32981-7

Khanam, F., Darton, T. C., Ross, A. G. P., Zaman, K., Pollard, A. J., Clemens, J. D., & Qadri, F. (2021). Case Report: Typhoid Fever Complicated by Ileal Perforation in an Urban Slum of Dhaka, Bangladesh. *Am J Trop Med Hyg, 104*(5), 1755-1757. doi:10.4269/ajtmh.20-1448

Kim, A. H., Hogarty, M. P., Harris, V. C., & Baldridge, M. T. (2020). The Complex Interactions Between Rotavirus and the Gut Microbiota. *Front Cell Infect Microbiol, 10*. doi:10.3389/fcimb.2020.586751

Kim, J. H., Cheong, H. K., & Jeon, B. H. (2018). Burden of Disease Attributable to Inadequate Drinking Water, Sanitation, and Hygiene in Korea. *J Korean Med Sci, 33*(46), e288. doi:10.3346/jkms.2018.33.e288

Kingsley, R. A., Kay, S., Connor, T., Barquist, L., Sait, L., Holt, K. E., . . . Dougan, G. (2013). Genome and transcriptome adaptation accompanying emergence of the definitive type 2 host-restricted Salmonella enterica serovar Typhimurium pathovar. *MBio, 4*(5), e00565-00513. doi:10.1128/mBio.00565-13

Kirolos, A., Thindwa, D., Khundi, M., Burke, R. M., Henrion, M. Y. R., Nakamura, I., . . . MacPherson, P. (2021). Tuberculosis case notifications in Malawi have strong seasonal and weather-related trends. *Sci Rep, 11*(1), 4621. doi:10.1038/s41598-021-84124-w

Klymiuk, I., Singer, G., Castellani, C., Trajanoski, S., Obermüller, B., & Till, H. (2021). Characterization of the Luminal and Mucosa-Associated Microbiome along the Gastrointestinal Tract: Results from Surgically Treated Preterm Infants and a Murine Model. *Nutrients, 13*(3). doi:10.3390/nu13031030

Koh, A., De Vadder, F., Kovatcheva-Datchary, P., & Bäckhed, F. (2016). From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites. *Cell, 165*(6), 1332-1345. doi:https://doi.org/10.1016/j.cell.2016.05.041

Koleva, P. T., Kim, J. S., Scott, J. A., & Kozyrskyj, A. L. (2015). Microbial programming of health and disease starts during fetal life. *Birth Defects Res C Embryo Today, 105*(4), 265-277. doi:10.1002/bdrc.21117

Korpela, K., Salonen, A., Vepsalainen, O., Suomalainen, M., Kolmeder, C., Varjosalo, M., . . . de Vos, W. M. (2018). Probiotic supplementation restores normal microbiota composition and function in antibiotic-treated and in caesarean-born infants. *Microbiome, 6*(1), 182. doi:10.1186/s40168-018-0567-4

Kortekangas, E., Kamng'ona, A. W., Fan, Y. M., Cheung, Y. B., Ashorn, U., Matchado, A., . . . Ashorn, P. (2020). Environmental exposures and child and maternal gut microbiota in rural Malawi. *Paediatr Perinat Epidemiol, 34*(2), 161-170. doi:10.1111/ppe.12623

Kotloff, K. L. (2017). The Burden and Etiology of Diarrheal Illness in Developing Countries. *Pediatr Clin North Am, 64*(4), 799-814. doi:10.1016/j.pcl.2017.03.006

Kotloff, K. L., Nataro, J. P., Blackwelder, W. C., Nasrin, D., Farag, T. H., Panchalingam, S., . . . Levine, M. M. (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. doi:10.1016/s0140-6736(13)60844-2

Koziolek, M., Grimm, M., Becker, D., Iordanov, V., Zou, H., Shimizu, J., . . . Weitschies, W. (2015). Investigation of pH and Temperature Profiles in the GI Tract of Fasted Human Subjects Using the Intellicap(®) System. *J Pharm Sci, 104*(9), 2855-2863. doi:10.1002/jps.24274

Kujawska, M., La Rosa, S. L., Roger, L. C., Pope, P. B., Hoyles, L., McCartney, A. L., & Hall, L. J. (2020). Succession of Bifidobacterium longum Strains in Response to a Changing Early Life Nutritional Environment Reveals Dietary Substrate Adaptations. *iScience, 23*(8), 101368. doi:10.1016/j.isci.2020.101368

Kurtz, J. R., Goggins, J. A., & McLachlan, J. B. (2017). Salmonella infection: Interplay between the bacteria and host immune system. *Immunol Lett, 190*, 42-50. doi:10.1016/j.imlet.2017.07.006

Lackey, K. A., Williams, J. E., Meehan, C. L., Zachek, J. A., Benda, E. D., Price, W. J., . . . McGuire, M. K. (2019). What's Normal? Microbiomes in Human Milk and Infant Feces Are Related to Each Other but Vary Geographically: The INSPIRE Study. *Frontiers in nutrition, 6*, 45-45. doi:10.3389/fnut.2019.00045

Lagesen, K., Hallin, P., Rødland, E. A., Staerfeldt, H.-H., Rognes, T., & Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic acids research, 35*(9), 3100-3108. doi:10.1093/nar/gkm160

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

Lau, C. S. M., & Chamberlain, R. S. (2016). Probiotics are effective at preventing Clostridium difficile-associated diarrhea: a systematic review and meta-analysis. *International Journal of General Medicine, 9*, 27-37. doi:10.2147/IJGM.S98280

Laudadio, I., Fulci, V., Palone, F., Stronati, L., Cucchiara, S., & Carissimi, C. (2018). Quantitative Assessment of Shotgun Metagenomics and 16S rDNA Amplicon Sequencing in the Study of Human Gut Microbiome. *Omics, 22*(4), 248-254. doi:10.1089/omi.2018.0013

Lawley, T. D., & Walker, A. W. (2013). Intestinal colonization resistance. *Immunology, 138*(1), 1-11. doi:10.1111/j.1365-2567.2012.03616.x

Lawson, M. A. E., O'Neill, I. J., Kujawska, M., Gowrinadh Javvadi, S., Wijeyesekera, A., Flegg, Z., . . . Hall, L. J. (2020). Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. *Isme j, 14*(2), 635-648. doi:10.1038/s41396-019-0553-2

LeBlanc, J. G., Milani, C., de Giori, G. S., Sesma, F., van Sinderen, D., & Ventura, M. (2013). Bacteria as vitamin suppliers to their host: a gut microbiota perspective. *Current Opinion in Biotechnology, 24*(2), 160-168. doi:https://doi.org/10.1016/j.copbio.2012.08.005

Lecuyer, E., Rakotobe, S., Lengline-Garnier, H., Lebreton, C., Picard, M., Juste, C., . . . Gaboriau-Routhiau, V. (2014). Segmented filamentous bacterium uses secondary and tertiary lymphoid tissues to induce gut IgA and specific T helper 17 cell responses. *Immunity, 40*(4), 608-620. doi:10.1016/j.immuni.2014.03.009

Lee, A., Gordon, J., & Dubos, R. (1968). Enumeration of the Oxygen Sensitive Bacteria usually Present in the Intestine of Healthy Mice. *Nature, 220*, 1137. doi:10.1038/2201137a0

Lee, S. A., Lim, J. Y., Kim, B.-S., Cho, S. J., Kim, N. Y., Kim, O. B., & Kim, Y. (2015). Comparison of the gut microbiota profile in breast-fed and formula-fed Korean infants using pyrosequencing. *Nutrition research and practice*(3), 242.

Legoux, F., Bellet, D., Daviaud, C., El Morr, Y., Darbois, A., Niort, K., . . . Lantz, O. (2019). Microbial metabolites control the thymic development of mucosal-associated invariant T cells. *Science, 366*(6464), 494-499. doi:10.1126/science.aaw2719

Lertsethtakarn, P., Silapong, S., Sakpaisal, P., Serichantalergs, O., Ruamsap, N., Lurchachaiwong, W., . . . Mason, C. J. (2018). Travelers' Diarrhea in Thailand: A Quantitative Analysis Using TaqMan® Array Card. *Clin Infect Dis, 67*(1), 120-127. doi:10.1093/cid/ciy040

Letunic, I., & Bork, P. (2016). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic acids research, 44*(W1), W242-W245. doi:10.1093/nar/gkw290

Levine, M. M., & Farag, T. H. (2011). Invasive salmonella infections and HIV in Northern Tanzania. *Clin Infect Dis, 52*(3), 349-351. doi:10.1093/cid/ciq109

Levine, M. M., & Robins-Browne, R. M. (2012). Factors that explain excretion of enteric pathogens by persons without diarrhea. *Clin Infect Dis, 55 Suppl 4*(Suppl 4), S303-311. doi:10.1093/cid/cis789

Levine, U. Y., Bearson, S. M., & Stanton, T. B. (2012). Mitsuokella jalaludinii inhibits growth of Salmonella enterica serovar Typhimurium. *Vet Microbiol, 159*(1-2), 115-122. doi:10.1016/j.vetmic.2012.03.027

Li, M., Zhou, X., Stanton, C., Ross, R. P., Zhao, J., Zhang, H., . . . Chen, W. (2021). Comparative Genomics Analyses Reveal the Differences between B. longum subsp. infantis and B. longum subsp. longum in Carbohydrate Utilisation, CRISPR-Cas Systems and Bacteriocin Operons. *Microorganisms, 9*(8). doi:10.3390/microorganisms9081713

Liang, S. C., Tan, X.-Y., Luxenberg, D. P., Karim, R., Dunussi-Joannopoulos, K., Collins, M., & Fouser, L. A. (2006). Interleukin (IL)-22 and IL-17 are coexpressed by Th17 cells and cooperatively enhance expression of antimicrobial peptides. *Journal of Experimental Medicine, 203*(10), 2271-2279. doi:10.1084/jem.20061308

Lim, H. J., & Shin, H. S. (2020). Antimicrobial and Immunomodulatory Effects of Bifidobacterium Strains: A Review. *J Microbiol Biotechnol, 30*(12), 1793-1800. doi:10.4014/jmb.2007.07046

Liu, D., Yan, J., Wang, H., Jiang, F., Song, P., Cai, Z., & Zhang, T. (2021). Microbial Biogeography along the Gastrointestinal Tract Segments of Sympatric Subterranean Rodents (Eospalax baileyi and Eospalax cansus). *Animals (Basel), 11*(11). doi:10.3390/ani11113297

Liu, J., Gratz, J., Amour, C., Kibiki, G., Becker, S., Janaki, L., . . . Houpt, E. R. (2013). A laboratory-developed TaqMan Array Card for simultaneous detection of 19 enteropathogens. *J Clin Microbiol, 51*(2), 472-480. doi:10.1128/jcm.02658-12

Liu, J., Kabir, F, Manneh, J et al. . (2014). Development and assessment of molecular diagnostic tests for 15 enteropathogens causing childhood diarrhoea: a multicentre study. *Lancet Infect Dis., 14*, 716–724.

Lkhagva, E., Chung, H. J., Hong, J., Tang, W. H. W., Lee, S. I., Hong, S. T., & Lee, S. (2021). The regional diversity of gut microbiome along the GI tract of male C57BL/6 mice. *BMC Microbiol, 21*(1), 44. doi:10.1186/s12866-021-02099-0

Lloyd-Price, J., Abu-Ali, G., & Huttenhower, C. (2016). The healthy human microbiome. *Genome medicine, 8*(1), 51-51. doi:10.1186/s13073-016-0307-y
Longhi, G., van Sinderen, D., Ventura, M., & Turroni, F. (2020). Microbiota and Cancer: The Emerging Beneficial Role of Bifidobacteria in Cancer Immunotherapy. *Front Microbiol, 11*, 575072. doi:10.3389/fmicb.2020.575072
Lowe, T. M., & Chan, P. P. (2016). tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. *Nucleic acids research, 44*(W1), W54-W57. doi:10.1093/nar/gkw413
Ly, A., & Hansen, D. S. (2019). Development of B Cell Memory in Malaria. *Front Immunol, 10*, 559. doi:10.3389/fimmu.2019.00559
MacLennan, C. A., Martin, L. B., & Micoli, F. (2014). Vaccines against invasive Salmonella disease: current status and future directions. *Hum Vaccin Immunother, 10*(6), 1478-1493. doi:10.4161/hv.29054
Mailhe, M., Ricaboni, D., Vitton, V., Gonzalez, J. M., Bachar, D., Dubourg, G., . . . Raoult, D. (2018). Repertoire of the gut microbiota from stomach to colon using culturomics and next-generation sequencing. *BMC Microbiol, 18*(1), 157. doi:10.1186/s12866-018-1304-7
Makino, H. (2018). Bifidobacterial strains in the intestines of newborns originate from their mothers. *Bioscience of microbiota, food and health, 37*(4), 79-85. doi:10.12938/bmfh.18-011
MAL-ED, N. I. (2014). The MAL-ED study: a multinational and multidisciplinary approach to understand the relationship between enteric pathogens, malnutrition, gut physiology, physical growth, cognitive development, and immune responses in infants and children up to 2 years of age in resource-poor environments. *Clin Infect Dis, 59 Suppl 4*, S193-206. doi:10.1093/cid/ciu653
Manko-Prykhoda, A., Allain, T., Motta, J. P., Cotton, J. A., Feener, T., Oyeyemi, A., . . . Buret, A. G. (2020). Giardia spp. promote the production of antimicrobial peptides and attenuate disease severity induced by attaching and effacing enteropathogens via the induction of the NLRP3 inflammasome. *Int J Parasitol, 50*(4), 263-275. doi:10.1016/j.ijpara.2019.12.011
Mathew, S., Smatti, M. K., Al Ansari, K., Nasrallah, G. K., Al Thani, A. A., & Yassine, H. M. (2019). Mixed Viral-Bacterial Infections and Their Effects on Gut Microbiota and Clinical Illnesses in Children. *Sci Rep, 9*(1), 865. doi:10.1038/s41598-018-37162-w
Mattarelli, P., & Biavati, B. (2018). Chapter 2 - Species in the Genus Bifidobacterium. In P. Mattarelli, B. Biavati, W. H. Holzapfel, & B. J. B. Wood (Eds.), *The Bifidobacteria and Related Organisms* (pp. 9-48): Academic Press.
Mattarelli, P., Bonaparte, C., Pot, B., & Biavati, B. (2008). Proposal to reclassify the three biotypes of Bifidobacterium longum as three subspecies: Bifidobacterium longum subsp. longum subsp. nov., Bifidobacterium longum subsp. infantis comb. nov. and Bifidobacterium longum subsp. suis comb. nov. *Int J Syst Evol Microbiol, 58*(Pt 4), 767-772. doi:10.1099/ijs.0.65319-0
Mazmanian, S. K., Liu, C. H., Tzianabos, A. O., & Kasper, D. L. (2005). An immunomodulatory molecule of symbiotic bacteria directs maturation of the host immune system. *Cell, 122*(1), 107-118. doi:10.1016/j.cell.2005.05.007
McMurry, T. L., McQuade, E. T. R., Liu, J., Kang, G., Kosek, M. N., Lima, A. A. M., . . . Platts-Mills, J. A. (2020). Duration of post-diarrheal enteric pathogen carriage in young children in low-resource settings. *Clin Infect Dis*. doi:10.1093/cid/ciaa1528

McNamara, J. F., Righi, E., Wright, H., Hartel, G. F., Harris, P. N. A., & Paterson, D. L. (2018). Long-term morbidity and mortality following bloodstream infection: A systematic literature review. *The Journal of infection, 77*(1), 1-8. doi:10.1016/j.jinf.2018.03.005

Mekonnen, H. S., & Ekubagewargies, D. T. (2019). Prevalence and factors associated with intestinal parasites among under-five children attending Woreta Health Center, Northwest Ethiopia. *BMC Infect Dis, 19*(1), 256-256. doi:10.1186/s12879-019-3884-8

Metsälä, J., Lundqvist, A., Virta, L. J., Kaila, M., Gissler, M., & Virtanen, S. M. (2015). Prenatal and post-natal exposure to antibiotics and risk of asthma in childhood. *Clin Exp Allergy, 45*(1), 137-145. doi:10.1111/cea.12356

Milani, C., Alessandri, G., Mancabelli, L., Mangifesta, M., Lugli, G. A., Viappiani, A., . . . Ventura, M. (2020). Multi-omics Approaches To Decipher the Impact of Diet and Host Physiology on the Mammalian Gut Microbiome. *Appl Environ Microbiol, 86*(23). doi:10.1128/aem.01864-20

Milani, C., Duranti, S., Bottacini, F., Casey, E., Turroni, F., Mahony, J., . . . Ventura, M. (2017). The First Microbial Colonizers of the Human Gut: Composition, Activities, and Health Implications of the Infant Gut Microbiota. *Microbiol Mol Biol Rev, 81*(4). doi:10.1128/mmbr.00036-17

Milligan, R., Paul, M., Richardson, M., & Neuberger, A. (2018). Vaccines for preventing typhoid fever. *Cochrane Database Syst Rev, 5*(5), Cd001261. doi:10.1002/14651858.CD001261.pub4

Milliken, S., Allen, R. M., & Lamont, R. F. (2019). The role of antimicrobial treatment during pregnancy on the neonatal gut microbiome and the development of atopy, asthma, allergy and obesity in childhood. *Expert Opin Drug Saf, 18*(3), 173-185. doi:10.1080/14740338.2019.1579795

Million, M., Tidjani Alou, M., Khelaifia, S., Bachar, D., Lagier, J.-C., Dione, N., . . . Raoult, D. (2016). Increased Gut Redox and Depletion of Anaerobic and Methanogenic Prokaryotes in Severe Acute Malnutrition. *Sci Rep, 6*(1), 26051. doi:10.1038/srep26051

Mitsuyama, M., Ohara, R., Amako, K., Nomoto, K., Yokokura, T., & Nomoto, K. (1986). Ontogeny of macrophage function to release superoxide anion in conventional and germfree mice. *Infect Immun, 52*(1), 236-239.

Moore, R. E., & Townsend, S. D. (2019). Temporal development of the infant gut microbiome. *Open Biol, 9*(9), 190128. doi:10.1098/rsob.190128

Moore, W. E., & Holdeman, L. V. (1974). Human fecal flora: the normal flora of 20 Japanese-Hawaiians. *Appl Microbiol, 27*(5), 961-979.

Moossavi, S., Sepehri, S., Robertson, B., Bode, L., Goruk, S., Field, C. J., . . . Azad, M. B. (2019). Composition and Variation of the Human Milk Microbiota Are Influenced by Maternal and Early-Life Factors. *Cell Host Microbe, 25*(2), 324-335.e324. doi:10.1016/j.chom.2019.01.011

Mørland, B., & Midtvedt, T. (1984). Phagocytosis, peritoneal influx, and enzyme activities in peritoneal macrophages from germfree, conventional, and ex-germfree mice. *Infection and Immunity, 44*(3), 750-752.

Msefula, C. L., Olgemoeller, F., Jambo, N., Segula, D., Van Tan, T., Nyirenda, T. S., . . . Heyderman, R. S. (2019). Ascertaining the burden of invasive Salmonella disease in hospitalised febrile children aged under four years in Blantyre, Malawi. *PLoS Negl Trop Dis, 13*(7), e0007539. doi:10.1371/journal.pntd.0007539

Mueller, N. T., Whyatt, R., Hoepner, L., Oberfield, S., Dominguez-Bello, M. G., Widen, E. M., . . . Rundle, A. (2015). Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int J Obes (Lond), 39*(4), 665-670. doi:10.1038/ijo.2014.180

Murphy, K., Curley, D., O'Callaghan, T. F., O'Shea, C.-A., Dempsey, E. M., O'Toole, P. W., . . . Stanton, C. (2017). The Composition of Human Milk and Infant Faecal Microbiota Over the First Three Months of Life: A Pilot Study. *Sci Rep, 7*, 40597-40597. doi:10.1038/srep40597

Murray, L. J., Lane, A. J., Harvey, I. M., Donovan, J. L., Nair, P., & Harvey, R. F. (2002). Inverse relationship between alcohol consumption and active Helicobacter pylori infection: the Bristol Helicobacter project. *The American Journal of Gastroenterology, 97*(11), 2750-2755. doi:https://doi.org/10.1016/S0002-9270(02)05481-3

Musicha, P., Cornick, J. E., Bar-Zeev, N., French, N., Masesa, C., Denis, B., . . . Feasey, N. A. (2017). Trends in antimicrobial resistance in bloodstream infection isolates at a large urban hospital in Malawi (1998-2016): a surveillance study. *Lancet Infect Dis, 17*(10), 1042-1052. doi:10.1016/s1473-3099(17)30394-8

Network Investigators, M.-E. (2018). Early childhood cognitive development is affected by interactions among illness, diet, enteropathogens and the home environment: findings from the MAL-ED birth cohort study. *BMJ Glob Health, 3*(4), e000752. doi:10.1136/bmjgh-2018-000752

Neuzil, K. M., Pollard, A. J., & Marfin, A. A. (2019). Introduction of Typhoid Conjugate Vaccines in Africa and Asia. *Clin Infect Dis, 68*(Supplement_1), S27-s30. doi:10.1093/cid/ciy878

Niu, S. Y., Yang, J., McDermaid, A., Zhao, J., Kang, Y., & Ma, Q. (2018). Bioinformatics tools for quantitative and functional metagenome and metatranscriptome data analysis in microbes. *Brief Bioinform, 19*(6), 1415-1429. doi:10.1093/bib/bbx051

Noverr, M. C., & Huffnagle, G. B. (2004). Does the microbiota regulate immune responses outside the gut? *Trends Microbiol, 12*(12), 562-568. doi:10.1016/j.tim.2004.10.008

Nwafia, I. N., Ohanu, M. E., Ebede, S. O., & Ozumba, U. C. (2019). Molecular detection and antibiotic resistance pattern of extended-spectrum beta-lactamase producing Escherichia coli in a Tertiary Hospital in Enugu, Nigeria. *Annals of clinical microbiology and antimicrobials, 18*(1), 41. doi:10.1186/s12941-019-0342-9

Nyirenda, T. S. (2015). *Natural Immunity to Salmonella in Humans.* (Doctor of Philosophy), University of Liverpool,

Nyirenda, T. S., Gilchrist, J. J., Feasey, N. A., Glennie, S. J., Bar-Zeev, N., Gordon, M. A., . . . Heyderman, R. S. (2014). Sequential Acquisition of T Cells and Antibodies to Nontyphoidal Salmonella in Malawian Children. *The Journal of Infectious Diseases, 210*(1), 56-64. doi:10.1093/infdis/jiu045

O'Neill, I., Schofield, Z., & Hall, L. J. (2017). Exploring the role of the microbiota member Bifidobacterium in modulating immune-linked diseases. *Emerging Topics in Life Sciences, 1*(4), 333-349. doi:10.1042/etls20170058

O'Toole, P. W., & Claesson, M. J. (2010). Review: Gut microbiota: Changes throughout the lifespan from infancy to elderly. *International Dairy Journal, 20*, 281-291. doi:10.1016/j.idairyj.2009.11.010

Office, N. S. (2019). *2018 Malawi Population and housing census Main Report*. Retrieved from http://populationmalawi.org/wp1/wp-content/uploads/2019/10/2018-Malawi-Population-and-Housing-Census-Main-Report-1.pdf.

Ondov, B. D., Treangen, T. J., Melsted, P., Mallonee, A. B., Bergman, N. H., Koren, S., & Phillippy, A. M. (2016). Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol, 17*(1), 132. doi:10.1186/s13059-016-0997-x

Orban, J. I., & Patterson, J. A. (2000). Modification of the phosphoketolase assay for rapid identification of bifidobacteria. *J Microbiol Methods, 40*(3), 221-224. doi:https://doi.org/10.1016/S0167-7012(00)00133-0

Östman, S., Rask, C., Wold, A. E., Hultkrantz, S., & Telemo, E. (2006). Impaired regulatory T cell function in germ-free mice. *European Journal of Immunology, 36*(9), 2336-2346. doi:10.1002/eji.200535244

Ottman, N., Smidt, H., de Vos, W., & Belzer, C. (2012). The function of our microbiota: who is out there and what do they do? *Frontiers in Cellular and Infection Microbiology, 2*(104). doi:10.3389/fcimb.2012.00104

Owen, N., Sparling, P. B., Healy, G. N., Dunstan, D. W., & Matthews, C. E. (2010). Sedentary behavior: emerging evidence for a new health risk. *Mayo Clin Proc, 85*(12), 1138-1141. doi:10.4065/mcp.2010.0444

Palleja, A., Mikkelsen, K. H., Forslund, S. K., Kashani, A., Allin, K. H., Nielsen, T., . . . Pedersen, O. (2018). Recovery of gut microbiota of healthy adults following antibiotic exposure. *Nat Microbiol, 3*(11), 1255-1265. doi:10.1038/s41564-018-0257-9

Panek, M., Čipčić Paljetak, H., Barešić, A., Perić, M., Matijašić, M., Lojkić, I., . . . Verbanac, D. (2018). Methodology challenges in studying human gut microbiota - effects of collection, storage, DNA extraction and next generation sequencing technologies. *Sci Rep, 8*(1), 5143. doi:10.1038/s41598-018-23296-4

Parker, E. P. K., Praharaj, I., Zekavati, A., Lazarus, R. P., Giri, S., Operario, D. J., . . . Grassly, N. C. (2018). Influence of the intestinal microbiota on the immunogenicity of oral rotavirus vaccine given to infants in south India. *Vaccine, 36*(2), 264-272. doi:10.1016/j.vaccine.2017.11.031

Pascual, V., Alonso, N., Simó, M., Ortiz, G., Garcia, M. C., Xercavins, M., . . . Calbo, E. (2016). Bloodstream infections caused by Escherichia coli producing AmpC β-lactamases: epidemiology and clinical features. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology, 35*(12), 1997-2003. doi:10.1007/s10096-016-2752-3

Patel, P. D., Patel, P., Liang, Y., Meiring, J. E., Misiri, T., Mwakiseghile, F., . . . Neuzil, K. M. (2021). Safety and Efficacy of a Typhoid Conjugate Vaccine in Malawian Children. *N Engl J Med, 385*(12), 1104-1115. doi:10.1056/NEJMoa2035916

Pichon, M., & Burucoa, C. (2019). Impact of the Gastro-Intestinal Bacterial Microbiome on Helicobacter-Associated Diseases. *Healthcare (Basel), 7*(1). doi:10.3390/healthcare7010034

Platts-Mills, J. A., Babji, S., Bodhidatta, L., Gratz, J., Haque, R., Havt, A., . . . Houpt, E. R. (2015). Pathogen-specific burdens of community diarrhoea in developing countries: a multisite birth cohort study (MAL-ED). *Lancet Glob Health, 3*(9), e564-575. doi:10.1016/s2214-109x(15)00151-5

Platts-Mills, J. A., Liu, J., Rogawski, E. T., Kabir, F., Lertsethtakarn, P., Siguas, M., . . . Houpt, E. R. (2018). Use of quantitative molecular diagnostic methods to assess the aetiology, burden, and clinical characteristics of diarrhoea in children in low-resource settings: a reanalysis of the MAL-ED cohort study. *Lancet Glob Health, 6*(12), e1309-e1318. doi:10.1016/s2214-109x(18)30349-8

Platts-Mills, J. A., Taniuchi, M., Uddin, M. J., Sobuz, S. U., Mahfuz, M., Gaffar, S. A., . . . Ahmed, T. (2017). Association between enteropathogens and malnutrition in children aged 6-23 mo in Bangladesh: a case-control study. *Am J Clin Nutr, 105*(5), 1132-1138. doi:10.3945/ajcn.116.138800

Ponzo, V., Ferrocino, I., Zarovska, A., Amenta, M. B., Leone, F., Monzeglio, C., . . . Bo, S. (2019). The microbiota composition of the offspring of patients with gestational diabetes mellitus (GDM). *PLoS One, 14*(12), e0226545. doi:10.1371/journal.pone.0226545

Pope, J. L., Tomkovich, S., & Jobin, C. (2020). Chapter 9 - Influence of Commensal Microbiota and Metabolite for Mucosal Immunity. In H. Kiyono & D. W. Pascual (Eds.), *Mucosal Vaccines (Second Edition)* (pp. 143-164): Academic Press.

Post, A. S., Diallo, S. N., Guiraud, I., Lompo, P., Tahita, M. C., Maltha, J., . . . Jacobs, J. (2019). Supporting evidence for a human reservoir of invasive non-Typhoidal Salmonella from household samples in Burkina Faso. *PLoS Negl Trop Dis, 13*(10), e0007782-e0007782. doi:10.1371/journal.pntd.0007782

Praharaj, I., Parker, E. P. K., Giri, S., Allen, D. J., Silas, S., Revathi, R., . . . Kang, G. (2019). Influence of Nonpolio Enteroviruses and the Bacterial Gut Microbiota on Oral Poliovirus Vaccine Response: A Study from South India. *J Infect Dis, 219*(8), 1178-1186. doi:10.1093/infdis/jiy568

Pritchard, L. a. G., Rachel H. and Humphris, Sonia and Elphinstone, John G. and Toth, Ian K. (2016). Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Anal. Methods, 8*(1), 12-24. doi:10.1039/C5AY02550H

Proctor, L. M. (2016). The National Institutes of Health Human Microbiome Project. *Seminars in fetal & neonatal medicine, 21*(6), 368-372. doi:10.1016/j.siny.2016.05.002

Qiime2. (2021). Training feature classifiers with q2-feature-classifier. Retrieved from https://docs.qiime2.org/2021.11/tutorials/feature-classifier/

Qu, N., Xu, M., Mizoguchi, I., Furusawa, J.-i., Kaneko, K., Watanabe, K., . . . Yoshimoto, T. (2013). Pivotal Roles of T-Helper 17-Related Cytokines, IL-17, IL-22, and IL-23, in Inflammatory Diseases. *Clinical and Developmental Immunology, 2013*, 968549. doi:10.1155/2013/968549

Rakoff-Nahoum, S., Paglino, J., Eslami-Varzaneh, F., Edberg, S., & Medzhitov, R. (2004). Recognition of commensal microflora by toll-like receptors is required for intestinal homeostasis. *Cell, 118*(2), 229-241. doi:10.1016/j.cell.2004.07.002

Ranjan, R., Rani, A., Metwally, A., McGee, H. S., & Perkins, D. L. (2016). Analysis of the microbiome: Advantages of whole genome shotgun versus 16S amplicon sequencing. *Biochem Biophys Res Commun, 469*(4), 967-977. doi:10.1016/j.bbrc.2015.12.083

Rea, M. C., Clayton, E., O'Connor, P. M., Shanahan, F., Kiely, B., Ross, R. P., & Hill, C. (2007). Antimicrobial activity of lacticin 3,147 against clinical Clostridium

difficile strains. *J Med Microbiol, 56*(Pt 7), 940-946. doi:10.1099/jmm.0.47085-0
Reyman, M., van Houten, M. A., van Baarle, D., Bosch, A. A. T. M., Man, W. H., Chu, M. L. J. N., . . . Bogaert, D. (2019). Impact of delivery mode-associated gut microbiota dynamics on health in the first year of life. *Nature communications, 10*(1), 4997-4997. doi:10.1038/s41467-019-13014-7
Riba, A., Hassani, K., Walker, A., van Best, N., von Zeschwitz, D., Anslinger, T., . . . Hornef, M. W. (2020). Disturbed gut microbiota and bile homeostasis in Giardia-infected mice contributes to metabolic dysregulation and growth impairment. *Sci Transl Med, 12*(565). doi:10.1126/scitranslmed.aay7019
Riley, L. W. (2014). Pandemic lineages of extraintestinal pathogenic Escherichia coli. *Clinical microbiology and infection : the official publication of the European Society of Clinical Microbiology and Infectious Diseases, 20*(5), 380-390. doi:10.1111/1469-0691.12646
Rinninella, E., Raoul, P., Cintoni, M., Franceschi, F., Miggiano, G. A. D., Gasbarrini, A., & Mele, M. C. (2019). What is the Healthy Gut Microbiota Composition? A Changing Ecosystem across Age, Environment, Diet, and Diseases. *Microorganisms, 7*(1). doi:10.3390/microorganisms7010014
Rissman, L., Deavenport-Saman, A., Corden, M. H., Zipkin, R., & Espinoza, J. (2021). A pilot project: handwashing educational intervention decreases incidence of respiratory and diarrheal illnesses in a rural Malawi orphanage. *Glob Health Promot, 28*(3), 14-22. doi:10.1177/1757975920963889
Riva, A., Borgo, F., Lassandro, C., Verduci, E., Morace, G., Borghi, E., & Berry, D. (2017). Pediatric obesity is associated with an altered gut microbiota and discordant shifts in Firmicutes populations. *Environmental Microbiology, 19*(1), 95-105. doi:10.1111/1462-2920.13463
Rivière, A., Selak, M., Lantin, D., Leroy, F., & De Vuyst, L. (2016). Bifidobacteria and Butyrate-Producing Colon Bacteria: Importance and Strategies for Their Stimulation in the Human Gut. *Front Microbiol, 7*, 979-979. doi:10.3389/fmicb.2016.00979
Robertson, C., Savva, G. M., Clapuci, R., Jones, J., Maimouni, H., Brown, E., . . . Clarke, P. (2020). Incidence of necrotising enterocolitis before and after introducing routine prophylactic Lactobacillus and Bifidobacterium probiotics. *Arch Dis Child Fetal Neonatal Ed, 105*(4), 380-386. doi:10.1136/archdischild-2019-317346
Rogawski, E. T., Liu, J., Platts-Mills, J. A., Kabir, F., Lertsethtakarn, P., Siguas, M., . . . Houpt, E. R. (2018). Use of quantitative molecular diagnostic methods to investigate the effect of enteropathogen infections on linear growth in children in low-resource settings: longitudinal analysis of results from the MAL-ED cohort study. *Lancet Glob Health, 6*(12), e1319-e1328. doi:10.1016/s2214-109x(18)30351-6
Rogers, Y.-H., & Zhang, C. (2016). Genomic technologies in medicine and health: past, present, and future. In *Medical and Health Genomics* (pp. 15-28): Elsevier.
Rothschild, D., Weissbrod, O., Barkan, E., Kurilshikov, A., Korem, T., Zeevi, D., . . . Segal, E. (2018). Environment dominates over host genetics in shaping human gut microbiota. *Nature, 555*, 210. doi:10.1038/nature25973
https://www.nature.com/articles/nature25973#supplementary-information
Rowland, I., Gibson, G., Heinken, A., Scott, K., Swann, J., Thiele, I., & Tuohy, K. (2018). Gut microbiota functions: metabolism of nutrients and other food

components. *European Journal of Nutrition, 57*(1), 1-24. doi:10.1007/s00394-017-1445-8
Rutayisire, E., Huang, K., Liu, Y., & Tao, F. (2016). The mode of delivery affects the diversity and colonization pattern of the gut microbiota during the first year of infants' life: a systematic review. *BMC Gastroenterology, 16*(1), 86. doi:10.1186/s12876-016-0498-0
Sana, T. G., Flaugnatti, N., Lugo, K. A., Lam, L. H., Jacobson, A., Baylot, V., . . . Monack, D. M. (2016). Salmonella Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. *Proceedings of the National Academy of Sciences, 113*(34), E5044-E5051. doi:10.1073/pnas.1608858113
Sanschagrin, S., & Yergeau, E. (2014). Next-generation sequencing of 16S ribosomal RNA gene amplicons. *J Vis Exp*(90). doi:10.3791/51709
Sassone-Corsi, M., & Raffatellu, M. (2015). No Vacancy: How Beneficial Microbes Cooperate with Immunity To Provide Colonization Resistance to Pathogens. *The Journal of Immunology, 194*(9), 4081-4087. doi:10.4049/jimmunol.1403169
Scepanovic, P., Hodel, F., Mondot, S., Partula, V., Byrd, A., Hammer, C., . . . Fellay, J. (2019). A comprehensive assessment of demographic, environmental, and host genetic associations with gut microbiome diversity in healthy individuals. *Microbiome, 7*(1), 130. doi:10.1186/s40168-019-0747-x
Schriefer, A. E., Cliften, P. F., Hibberd, M. C., Sawyer, C., Brown-Kennerly, V., Burcea, L., . . . Head, R. D. (2018). A multi-amplicon 16S rRNA sequencing and analysis method for improved taxonomic profiling of bacterial communities. *J Microbiol Methods, 154*, 6-13. doi:10.1016/j.mimet.2018.09.019
Schülke, S. (2018). Induction of Interleukin-10 Producing Dendritic Cells As a Tool to Suppress Allergen-Specific T Helper 2 Responses. *Front Immunol, 9*, 455. Retrieved from http://europepmc.org/abstract/MED/29616018
https://doi.org/10.3389/fimmu.2018.00455
https://europepmc.org/articles/PMC5867300
https://europepmc.org/articles/PMC5867300?pdf=render
doi:10.3389/fimmu.2018.00455
Seboxa, T., Amogne, W., Abebe, W., Tsegaye, T., Azazh, A., Hailu, W., . . . Henriksen, T.-H. (2015). High Mortality from Blood Stream Infection in Addis Ababa, Ethiopia, Is Due to Antimicrobial Resistance. *PLoS One, 10*(12), e0144944-e0144944. doi:10.1371/journal.pone.0144944
Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics (Oxford, England), 30*(14), 2068-2069. doi:10.1093/bioinformatics/btu153
Selma-Royo, M., Tarrazo, M., Garcia-Mantrana, I., Gomez-Gallego, C., Salminen, S., & Collado, M. C. (2019). Shaping Microbiota During the First 1000 Days of Life. *Adv Exp Med Biol, 1125*, 3-24. doi:10.1007/5584_2018_312
Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol, 14*(8), e1002533. doi:10.1371/journal.pbio.1002533
Sevelsted, A., Stokholm, J., Bønnelykke, K., & Bisgaard, H. (2015). Cesarean section and chronic immune disorders. *Pediatrics, 135*(1), e92-98. doi:10.1542/peds.2014-0596
Shakya, M., Voysey, M., Theiss-Nyland, K., Colin-Jones, R., Pant, D., Adhikari, A., . . . Pollard, A. J. (2021). Efficacy of typhoid conjugate vaccine in Nepal: final

results of a phase 3, randomised, controlled trial. *Lancet Glob Health, 9*(11), e1561-e1568. doi:10.1016/s2214-109x(21)00346-6
Sherman, S. B., Sarsour, N., Salehi, M., Schroering, A., Mell, B., Joe, B., & Hill, J. W. (2018). Prenatal androgen exposure causes hypertension and gut microbiota dysbiosis. *Gut Microbes, 9*(5), 400-421. doi:10.1080/19490976.2018.1441664
Sirilun, S., Takahashi, H., Boonyaritichaikij, S., Chaiyasut, C., Lertruangpanya, P., Koga, Y., & Mikami, K. (2015). Impact of maternal bifidobacteria and the mode of delivery on Bifidobacterium microbiota in infants. *Benef Microbes, 6*(6), 767-774. doi:10.3920/bm2014.0124
Slack, M. P., & Wheldon, D. B. (1978). A simple and safe volumetric alternative to the method of Miles, Misra and Irwin for counting viable bacteria. *J Med Microbiol, 11*(4), 541-545. doi:10.1099/00222615-11-4-541
Smith, M. I., Yatsunenko, T., Manary, M. J., Trehan, I., Mkakosya, R., Cheng, J., . . . Gordon, J. I. (2013). Gut Microbiomes of Malawian Twin Pairs Discordant for Kwashiorkor. *Science, 339*(6119), 548-554. doi:10.1126/science.1229000
Socha-Banasiak, A., Pawłowska, M., Czkwianianc, E., & Pierzynowska, K. (2021). From Intrauterine to Extrauterine Life-The Role of Endogenous and Exogenous Factors in the Regulation of the Intestinal Microbiota Community and Gut Maturation in Early Life. *Frontiers in nutrition, 8*, 696966. doi:10.3389/fnut.2021.696966
Spellberg, B., & Edwards, J. E., Jr. (2001). Type 1/Type 2 Immunity in Infectious Diseases. *Clinical Infectious Diseases, 32*(1), 76-102. doi:10.1086/317537
Stacy, A., Andrade-Oliveira, V., McCulloch, J. A., Hild, B., Oh, J. H., Perez-Chaparro, P. J., . . . Belkaid, Y. (2021). Infection trains the host for microbiota-enhanced resistance to pathogens. *Cell, 184*(3), 615-627.e617. doi:10.1016/j.cell.2020.12.011
Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics (Oxford, England), 30*(9), 1312-1313. doi:10.1093/bioinformatics/btu033
Stanislawski, M. A., Dabelea, D., Wagner, B. D., Sontag, M. K., Lozupone, C. A., & Eggesbø, M. (2017). Pre-pregnancy weight, gestational weight gain, and the gut microbiota of mothers and their infants. *Microbiome, 5*(1), 113-113. doi:10.1186/s40168-017-0332-0
Statista. (2021). The 20 countries with the lowest gross domestic product (GDP) per capita in 2020. Retrieved from https://www.statista.com/statistics/256547/the-20-countries-with-the-lowest-gdp-per-capita/
Stecher, B. (2015). The Roles of Inflammation, Nutrient Availability and the Commensal Microbiota in Enteric Pathogen Infection. *Microbiol Spectr, 3*(3). doi:10.1128/microbiolspec.MBP-0008-2014
Stelter, C., Käppeli, R., König, C., Krah, A., Hardt, W. D., Stecher, B., & Bumann, D. (2011). Salmonella-induced mucosal lectin RegIIIβ kills competing gut microbiota. *PLoS One, 6*(6), e20749. doi:10.1371/journal.pone.0020749
Streit, W. R., & Schmitz, R. A. (2004). Metagenomics – the key to the uncultured microbes. *Curr Opin Microbiol, 7*(5), 492-498. doi:https://doi.org/10.1016/j.mib.2004.08.002
Subramanian, S., Huq, S., Yatsunenko, T., Haque, R., Mahfuz, M., Alam, M. A., . . . Gordon, J. I. (2014). Persistent gut microbiota immaturity in malnourished Bangladeshi children. *Nature, 510*(7505), 417-421. doi:10.1038/nature13421
Taft, D., Ho, S., Tancredi, D., Stephensen, C., Hinde, K., von Mutius, E., . . . Mills, D. (2019). Population Duration of Breastfeeding and Prevalence of

Bifidobacterium Longum Subspecies Infantis (OR01-01-19). *Current Developments in Nutrition, 3*(Suppl 1), nzz040.OR001-001-019. doi:10.1093/cdn/nzz040.OR01-01-19
Taft, D. H., Liu, J., Maldonado-Gomez, M. X., Akre, S., Huda, M. N., Ahmad, S. M., . . . Mills, D. A. (2018). Bifidobacterial Dominance of the Gut in Early Life and Acquisition of Antimicrobial Resistance. *mSphere, 3*(5). doi:10.1128/mSphere.00441-18
Tanaka, S., Kobayashi, T., Songjinda, P., Tateyama, A., Tsubouchi, M., Kiyohara, C., . . . Nakayama, J. (2009). Influence of antibiotic exposure in the early postnatal period on the development of intestinal microbiota. *FEMS Immunol Med Microbiol, 56*(1), 80-87. doi:10.1111/j.1574-695X.2009.00553.x
Taniuchi, M., Platts-Mills, J. A., Begum, S., Uddin, M. J., Sobuz, S. U., Liu, J., . . . Houpt, E. R. (2016). Impact of enterovirus and other enteric pathogens on oral polio and rotavirus vaccine performance in Bangladeshi infants. *Vaccine, 34*(27), 3068-3075. doi:10.1016/j.vaccine.2016.04.080
Tarracchini, C., Milani, C., Lugli, G. A., Mancabelli, L., Fontana, F., Alessandri, G., . . . Ventura, M. (2021). Phylogenomic disentangling of the Bifidobacterium longum subsp. infantis taxon. *Microb Genom, 7*(7). doi:10.1099/mgen.0.000609
Temkin, E., Fallach, N., Almagor, J., Gladstone, B. P., Tacconelli, E., Carmeli, Y., & Consortium, D.-A. (2018). Estimating the number of infections caused by antibiotic-resistant Escherichia coli and Klebsiella pneumoniae in 2014: a modelling study. *Lancet Glob Health, 6*(9), e969-e979. doi:10.1016/S2214-109X(18)30278-X
Tennant, S. M., & Levine, M. M. (2015). Live attenuated vaccines for invasive Salmonella infections. *Vaccine, 33 Suppl 3*(0 3), C36-41. doi:10.1016/j.vaccine.2015.04.029
Tennant, S. M., MacLennan, C. A., Simon, R., Martin, L. B., & Khan, M. I. (2016). Nontyphoidal salmonella disease: Current status of vaccine research and development. *Vaccine, 34*(26), 2907-2910. doi:10.1016/j.vaccine.2016.03.072
Tenthani, L., van Oosterhout, J. J., Haas, A. D., Msukwa, M., Phiri, N., Chimbwandira, F., . . . Keiser, O. (2018). Patterns of caesarean section in HIV infected and non-infected women in Malawi: is caesarean section used for PMTCT? *BMC Pregnancy Childbirth, 18*(1), 95. doi:10.1186/s12884-018-1722-4
The Human Microbiome Project, C. (2012). Structure, function and diversity of the healthy human microbiome. *Nature, 486*, 207. doi:10.1038/nature11234
https://www.nature.com/articles/nature11234#supplementary-information
Thindwa, D., Chipeta, M. G., Henrion, M. Y. R., & Gordon, M. A. (2019). Distinct climate influences on the risk of typhoid compared to invasive non-typhoid Salmonella disease in Blantyre, Malawi. *Sci Rep, 9*(1), 20310. doi:10.1038/s41598-019-56688-1
Thomson, P., Medina, D. A., & Garrido, D. (2018). Human milk oligosaccharides and infant gut bifidobacteria: Molecular strategies for their utilization. *Food Microbiol, 75*, 37-46. doi:10.1016/j.fm.2017.09.001
Thursby, E., & Juge, N. (2017). Introduction to the human gut microbiota. *Biochem J, 474*(11), 1823-1836. doi:10.1042/bcj20160510
Tidjani Alou, M., Lagier, J.-C., & Raoult, D. (2016). Diet influence on the gut microbiota and dysbiosis related to nutritional disorders. *Human Microbiome Journal, 1*, 3-11. doi:https://doi.org/10.1016/j.humic.2016.09.001

Timmerman, H. M., Rutten, N., Boekhorst, J., Saulnier, D. M., Kortman, G. A. M., Contractor, N., . . . Rijkers, G. T. (2017). Intestinal colonisation patterns in breastfed and formula-fed infants during the first 12 weeks of life reveal sequential microbiota signatures. *Sci Rep, 7*(1), 8327. doi:10.1038/s41598-017-08268-4

Tizifa, T. A., Kabaghe, A. N., McCann, R. S., Nkhono, W., Mtengula, S., Takken, W., . . . van Vugt, M. (2021). Incidence of clinical malaria, acute respiratory illness, and diarrhoea in children in southern Malawi: a prospective cohort study. *Malar J, 20*(1), 473. doi:10.1186/s12936-021-04013-5

Toro-Londono, M. A., Bedoya-Urrego, K., Garcia-Montoya, G. M., Galvan-Diaz, A. L., & Alzate, J. F. (2019). Intestinal parasitic infection alters bacterial gut microbiota in children. *PeerJ, 7*, e6200-e6200. doi:10.7717/peerj.6200

Toshitaka, O., Kumiko, K., Hirosuke, S., Nanami, H., Sachiko, T., Jin-zhong, X., . . . Ro, O. (2016). Age-related changes in gut microbiota composition from newborn to centenarian: a cross-sectional study. *BMC Microbiology, 16*, 1-12. doi:10.1186/s12866-016-0708-5

Tremblay, M. S., Colley, R. C., Saunders, T. J., Healy, G. N., & Owen, N. (2010). Physiological and health implications of a sedentary lifestyle. *Appl Physiol Nutr Metab, 35*(6), 725-740. doi:10.1139/h10-079

Tsugawa, H., Kabe, Y., Kanai, A., Sugiura, Y., Hida, S., Taniguchi, S., . . . Suematsu, M. (2020). Short-chain fatty acids bind to apoptosis-associated speck-like protein to activate inflammasome complex to prevent Salmonella infection. *PLoS Biol, 18*(9), e3000813. doi:10.1371/journal.pbio.3000813

Tun, H. M., Mandhane, P. J., Kozyrskyj, A. L., Konya, T., Brook, J. R., Scott, J. A., . . . Takaro, T. (2017). Exposure to household furry pets influences the gut microbiota of infants at 3-4 months following various birth scenarios. *Microbiome, 5*(1). doi:10.1186/s40168-017-0254-x

Turroni, F., Milani, C., Duranti, S., Ferrario, C., Lugli, G. A., Mancabelli, L., . . . Ventura, M. (2018). Bifidobacteria and the infant gut: an example of co-evolution and natural selection. *Cell Mol Life Sci, 75*(1), 103-118. doi:10.1007/s00018-017-2672-0

Turroni, F., Milani, C., Ventura, M., & van Sinderen, D. (2021). The human gut microbiota during the initial stages of life: insights from bifidobacteria. *Curr Opin Biotechnol, 73*, 81-87. doi:10.1016/j.copbio.2021.07.012

Turroni, F., Peano, C., Pass, D. A., Foroni, E., Severgnini, M., Claesson, M. J., . . . Ventura, M. (2012). Diversity of bifidobacteria within the infant gut microbiota. *PLoS One, 7*(5), e36957-e36957. doi:10.1371/journal.pone.0036957

Ursell, L. K., Clemente, J. C., Rideout, J. R., Gevers, D., Caporaso, J. G., & Knight, R. (2012). The interpersonal and intrapersonal diversity of human-associated microbiota in key body sites. *J Allergy Clin Immunol, 129*(5), 1204-1208. doi:10.1016/j.jaci.2012.03.010

Ursell, L. K., Metcalf, J. L., Parfrey, L. W., & Knight, R. (2012). Defining the human microbiome. *Nutrition reviews, 70 Suppl 1*(Suppl 1), S38-S44. doi:10.1111/j.1753-4887.2012.00493.x

van den Broek, N. R., Jean-Baptiste, R., & Neilson, J. P. (2014). Factors associated with preterm, early preterm and late preterm birth in Malawi. *PLoS One, 9*(3), e90128. doi:10.1371/journal.pone.0090128

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., & Thermes, C. (2018). The Third Revolution in Sequencing Technology. *Trends Genet, 34*(9), 666-681. doi:10.1016/j.tig.2018.05.008

Ventura, M., Milani, C., Turroni, F., & van Sinderen, D. (2021). Envisioning emerging frontiers on human gut microbiota and its applications. *Microbial biotechnology, 14*(1), 12-17. doi:10.1111/1751-7915.13671
Versloot, C. J., Attia, S., Bourdon, C., Richardson, S. E., Potani, I., Bandsma, R. H. J., & Voskuijl, W. (2018). Intestinal pathogen clearance in children with severe acute malnutrition is unrelated to inpatient morbidity. *Clin Nutr ESPEN, 24*, 109-113. doi:10.1016/j.clnesp.2018.01.004
Walker, S. P., Wachs, T. D., Gardner, J. M., Lozoff, B., Wasserman, G. A., Pollitt, E., & Carter, J. A. (2007). Child development: risk factors for adverse outcomes in developing countries. *Lancet, 369*(9556), 145-157. doi:10.1016/s0140-6736(07)60076-2
Wang, B., Yao, M., Lv, L., Ling, Z., & Li, L. (2017). The Human Microbiota in Health and Disease. *Engineering, 3*(1), 71-82. doi:https://doi.org/10.1016/J.ENG.2017.01.008
Wang, H. X., & Wang, Y. P. (2016). Gut Microbiota-brain Axis. *Chin Med J (Engl), 129*(19), 2373-2380. doi:10.4103/0366-6999.190667
Wang, J., Thingholm, L. B., Skieceviciene, J., Rausch, P., Kummen, M., Hov, J. R., . . . Franke, A. (2016). Genome-wide association analysis identifies variation in vitamin D receptor and other host factors influencing the gut microbiota. *Nat Genet, 48*(11), 1396-1406. doi:10.1038/ng.3695
Wang, Y., Yin, Y., Chen, X., Zhao, Y., Wu, Y., Li, Y., . . . Xiang, C. (2019). Induction of Intestinal Th17 Cells by Flagellins From Segmented Filamentous Bacteria. *Front Immunol, 10*, 2750-2750. doi:10.3389/fimmu.2019.02750
Watts, G. S., Youens-Clark, K., Slepian, M. J., Wolk, D. M., Oshiro, M. M., Metzger, G. S., . . . Hurwitz, B. L. (2017). 16S rRNA gene sequencing on a benchtop sequencer: accuracy for identification of clinically important bacteria. *J Appl Microbiol, 123*(6), 1584-1596. doi:10.1111/jam.13590
Werlang, I. C. R., Mueller, N. T., Pizoni, A., Wisintainer, H., Matte, U., Costa, S., . . . Goldani, H. A. S. (2018). Associations of birth mode with cord blood cytokines, white blood cells, and newborn intestinal bifidobacteria. *PLoS One, 13*(11), e0205962. doi:10.1371/journal.pone.0205962
WHO. (2018a, 19 February 2018). Preterm birth. Retrieved from https://www.who.int/news-room/fact-sheets/detail/preterm-birth
WHO. (2018b). Salmonella (non-typhoidal). *Fact sheets -detail.* Retrieved from https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal)
WHO. (2019). The Global Health Observatory - Malawi. Retrieved from https://www.who.int/data/gho/data/countries/country-details/GHO/malawi?countryProfileId=3d12e768-34b8-4076-bdc7-061df3ce5f10
WHO. (2020). Antimicrobial resistance. *Antimicrobial resistance* Retrieved from https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance
WHO. (2021). Malawi Overview: Development news, research, data. Retrieved from https://www.worldbank.org/en/country/malawi/overview
Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol, 13*(6), e1005595. doi:10.1371/journal.pcbi.1005595
Wolf, K. J., & Lorenz, R. G. (2012). Gut Microbiota and Obesity. *Current obesity reports, 1*(1), 1-8. doi:10.1007/s13679-011-0001-8

World-Medical-Association. (2013). World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. *Jama, 310*(20), 2191-2194. doi:10.1001/jama.2013.281053

Y.E., C., Fischbach, M. A., & Belkaid, Y. (2018). Skin microbiota-host interactions. *Nature, 553*(7689), 427-436. doi:10.1038/nature25177

Yan, J., & Charles, J. F. (2017). Gut Microbiome and Bone: to Build, Destroy, or Both? *Current osteoporosis reports, 15*(4), 376-384. doi:10.1007/s11914-017-0382-z

Yang, A., Zhang, W., Wang, J., Yang, K., Han, Y., & Zhang, L. (2020). Review on the Application of Machine Learning Algorithms in the Sequence Data Mining of DNA. *Frontiers in bioengineering and biotechnology, 8*, 1032-1032. doi:10.3389/fbioe.2020.01032

Yasmin, F., Dinu, I., Kozyrskyj, A. L., Tun, H. M., Mandhane, P. J., Konya, T. B., . . . Sears, M. R. (2017). Cesarean Section, Formula Feeding, and Infant Antibiotic Exposure: Separate and Combined Impacts on Gut Microbial Changes in Later Infancy. *FRONTIERS IN PEDIATRICS, 5*.

Yatsunenko, T., Rey, F. E., Manary, M. J., Trehan, I., Dominguez-Bello, M. G., Contreras, M., . . . Gordon, J. I. (2012). Human gut microbiome viewed across age and geography. *Nature, 486*(7402), 222-227. doi:10.1038/nature11053

Yoon, S. H., Ha, S. M., Lim, J., Kwon, S., & Chun, J. (2017). A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek, 110*(10), 1281-1286. doi:10.1007/s10482-017-0844-4

Zhou, X., Brown, C. J., Abdo, Z., Davis, C. C., Hansmann, M. A., Joyce, P., . . . Forney, L. J. (2007). Differences in the composition of vaginal microbial communities found in healthy Caucasian and black women. *Isme j, 1*(2), 121-133. doi:10.1038/ismej.2007.12

Zou, Z.-H., Liu, D., Li, H.-D., Zhu, D.-P., He, Y., Hou, T., & Yu, J.-L. (2018). Prenatal and postnatal antibiotic exposure influences the gut microbiota of preterm infants in neonatal intensive care units. *Annals of clinical microbiology and antimicrobials, 17*(1), 9-9. doi:10.1186/s12941-018-0264-y